INTERNATIONAL DOCTORAL THESIS / DOCTORAL PROGRAM IN EDUCATION

INNOVATIVE TEACHING UNITS AND ACTIVITY TRACKERS FOR THE PROMOTION OF HEALTHY PHYSICAL ACTIVITY HABITS IN PHYSICAL EDUCATION

CAROLINA CASADO ROBLES

Departamento de Educación Física y Deportiva
Facultad de Ciencias del Deporte
Universidad de Granada

2021

UNIVERSIDAD DE GRANADA

INNOVATIVE TEACHING UNITS AND ACTIVITY TRACKERS FOR THE PROMOTION OF HEALTHY PHYSICAL ACTIVITY HABITS IN PHYSICAL EDUCATION

DOCTORAL THESIS

INNOVATIVE TEACHING UNITS AND ACTIVITY TRACKERS FOR THE PROMOTION OF HEALTHY PHYSICAL ACTIVITY HABITS IN PHYSICAL EDUCATION

Editor: Universidad de Granada. Tesis doctorales

Cover design: Damián Fernández Muñoz

TESIS DOCTORAL INTERNACIONAL/ INTERNATIONAL DOCTORAL THESIS

PROGRAMA DE DOCTORADO EN CIENCIAS DE LA EDUCACIÓN

UNIVERSIDAD DE GRANADA

UNIDADES DIDÁCTICAS INNOVADORAS Y USO DE MONITORES PORTÁTILES DE FITNESS PARA LA PROMOCIÓN DE HÁBITOS DE ACTIVIDAD FÍSICA SALUDABLES EN EDUCACIÓN FÍSICA

INNOVATIVE TEACHING UNITS AND ACTIVITY TRACKERS FOR THE PROMOTION OF HEALTHY PHYSICAL ACTIVITY HABITS IN PHYSICAL EDUCATION

DEPARTAMENTO DE EDUCACIÓN FÍSICA Y DEPORTIVA

FACULTAD DE CIENCIAS DEL DEPORTE

UNIVERSIDAD DE GRANADA

CAROLINA CASADO ROBLES

2021

DEPARTAMENTO DE EDUCACIÓN FÍSICA Y DEPORTIVA
FACULTAD DE CIENCIAS DEL DEPORTE
UNIVERSIDAD DE GRANADA

UNIDADES DIDÁCTICAS INNOVADORAS Y USO DE MONITORES PORTÁTILES DE FITNESS PARA LA PROMOCIÓN DE HÁBITOS DE ACTIVIDAD FÍSICA SALUDABLES EN EDUCACIÓN FÍSICA

INNOVATIVE TEACHING UNITS AND ACTIVITY TRACKERS FOR THE PROMOTION OF HEALTHY PHYSICAL ACTIVITY HABITS IN PHYSICAL EDUCATION

CAROLINA CASADO ROBLES

Director de Tesis [Thesis Supervisor]
Dr. Jesús Viciana Ramírez
Catedrático de Universidad
Universidad de Granada
España

En Granada, 15 de octubre de 2021

Editor: Universidad de Granada. Tesis Doctorales
Autor: Carolina Casado Robles
ISBN: 978-84-1117-209-7
URI: http://hdl.handle.net/10481/72362

A todos los que me han acompañado en esta aventura,

especialmente a mi familia, amigos y Damián…

Cree en ti mismo y volarás muy alto

Walt Disney

TABLE OF CONTENTS [TABLA DE CONTENIDOS]

RESEARCH PROJECTS AND FUNDING

[PROYECTOS DE INVESTIGACIÓN Y FINANCIACIÓN]

RESEARCH PROJECTS AND FUNDING [PROYECTOS DE INVESTIGACIÓN Y FINANCIACIÓN]

La presente Tesis Doctoral ha sido desarrollada bajo los siguientes proyectos de investigación:

- Contrato predoctoral para la Formación del Profesorado Universitario (FPU16/03314). Proyecto: Unidades didácticas intermitentes, alternadas, irregulares y reforzadas en Educación Física para la promoción de hábitos de actividad física saludables: Diseño controlado aleatorizado por grupos naturales (FPU16/03314). Entidad financiadora: Ministerio de Universidades, Gobierno de España. Duración: Desde el 31/10/2017 hasta el 30/10/2021.

- Proyecto: Aplicación de monitores de FITness PERSONales en Educación Física: Validez de medida y efecto para promocionar la actividad física habitual en escolares. Estudio FIT-PERSON (Referencia: PID2019-110179GB-I00). Modalidad: Programa Estatal de I+D+i orientados a retos de la sociedad. Entidad financiadora: Ministerio de Ciencia e Innovación. Duración: Desde el 01/06/2020 hasta el 31/05/2023.

- Proyecto: Monitorización y promoción de hábitos saludables de actividad física y bienestar mediante monitores portátiles de fitness en escolares y adultos: Estudio WearFit (Referencia: B-SEJ-029-UGR18). Modalidad: Programa I+D+i en el marco del Programa Operativo FEDER Andalucía 2014-2020. Entidad financiadora: Unión Europea y la Junta de Andalucía. Duración: Desde el 01/01/2020 hasta el 31/12/2021.

- Proyecto: Validez de medida de las PULSeras de fitness y efecto de una intervención escolar en Educación Física sobre la actividad física habitual y conducta sedentaria en Educación Secundaria (PULSEF). Modalidad: Plan Propio de Investigación 2019 Programa de Proyectos de Investigación Precompetitivos para Jóvenes Investigadores. Entidad financiadora: Vicerrectorado de Investigación y Transferencia de la Universidad de Granada. Duración: Desde el 01/01/2020 hasta el 31/12/2020.

Además, la presente Tesis Doctoral ha sido posible gracias a las siguientes becas y subvenciones:

- Ayuda de movilidad para estancias breves en el extranjero a beneficiarios del subprograma de Formación del Profesorado Universitario (FPU) (EST18/00315). Centro: Faculty of Physical Culture, Palacký University Olomouc, Olomouc, República Checa. Entidad financiadora: Ministerio de Ciencia, Innovación y Universidades, Gobierno de España. Duración: Desde el 17/06/2019 hasta el 16/09/2019.

- Ayuda de movilidad ERASMUS+ de movilidad internacional para estudiantes de doctorado (Modalidad KA103). Centro: Faculty of Physical Culture, Palacký University Olomouc, Olomouc, República Checa. Entidad financiadora: Programa Erasmus+ de la Unión Europea, Ministerio de Educación, Cultura y Deporte, y la Junta de Andalucía. Duración: Desde el 13/09/2018 hasta el 12/12/2018.

- Beca de Iniciación a la Investigación para estudiantes de máster universitario. Entidad financiadora: Vicerrectorado de Investigación y Transferencia de la Universidad de Granada. Duración: Desde el 01/09/2017 hasta el 30/06/2018.

PUBLICATIONS

[PUBLICACIONES]

PUBLICATIONS [PUBLICACIONES]

La presente Tesis Doctoral está compuesta por las siguientes publicaciones:

I. **Casado-Robles, C.,** Mayorga-Vega, D., Guijarro-Romero, S., & Viciana, J. (2020). ¿Es efectivo el Modelo de Educación Deportiva para mejorar la deportividad y actividad física en escolares? *RICYDE. Revista Internacional de Ciencias del Deporte, 16*(60), 180-198. https://doi.org/10.5232/ricyde2020.06005. SJR: 0.276; Physical therapy, Sports Therapy and Rehabilitation: Q3

II. **Casado-Robles, C.,** Guijarro-Romero, S., & Mayorga-Vega, D. (2019). Planificación en Educación Física mediante unidades didácticas innovadoras para incrementar los niveles de actividad física habitual de los escolares. En S. Alonso García, J. M. Romero Rodríguez, C. Rodríguez-Jiménez, & J. M. Sola Reche (Eds.), *Investigación, Innovación docente y TIC. Nuevos Horizontes Educativos* (pp. 283-296). Dykinson S.L. ISBN: 978-84-1324-492-1. SPI: 28; Education: 12/33. SPI: 456; General: 13/104 (Q1).

III. **Casado-Robles, C.,** Mayorga-Vega, D., Guijarro-Romero, S., & Viciana, J. (2020). Sport education-based irregular teaching unit and students' physical activity during school recess. *The Journal of Educational Research, 113*(4), 262-274. https://doi.org/10.1080/00220671.2020.1806014. JCR: 1.871; Education and Educational Research: Q3.

IV. **Casado-Robles, C.,** Mayorga-Vega, D., Guijarro-Romero, S., & Viciana, J. Is it more fun to do Physical Education outside the school center? Under review in *The Journal of Educational Research*. JCR: 1.871; Education and Educational Research: Q3.

V. **Casado-Robles, C.,** Viciana, J., Guijarro-Romero, S., & Mayorga-Vega, D. (2021). Effect of an inside-outside school alternated teaching unit of knowledge of the environment for practicing physical activity: A cluster randomized control trial. *Journal of Teaching in Physical Education*. Advance online publication. https://doi.org/10.1123/jtpe.2020-0132. JCR: 4.155; Education and Educational Research: Q1.

VI. **Casado-Robles, C.,** Guijarro-Romero, S., & Mayorga-Vega, D. (In press). Promoción de la actividad física de los escolares mediante unidades didácticas intermitentes en Educación Física. En *Innovación en la docencia e investigación de las Ciencias Sociales y de la Educación*. Dykinson S.L. ISBN: 978-84-1377-590-6. SPI: 28; Education: 12/33. SPI: 456; General: 13/104 (Q1).

VII. **Casado-Robles, C.,** Viciana, J., Guijarro-Romero, S., & Mayorga-Vega, D. Effects of consumer-wearable activity tracker-based programs on objectively measured daily physical activity and sedentary behavior among school-aged children: A systematic review and meta-analysis. Under review in *Sports Medicine-Open*. JCR: 4.830; Sport Sciences: Q1.

VIII. **Casado-Robles, C.,** Mayorga-Vega, D., Guijarro-Romero, S., & Viciana, J. (2021). Conocimiento del entorno para la práctica de actividad física en escolares (CEPAF): Desarrollo y validación de una prueba escrita objetiva de elección múltiple. *Journal of Sport and Health Research, 13*(2), 223-244. SJR: 0.225; Education: Q3.

IX. **Casado-Robles, C.,** Viciana, J., Guijarro-Romero, S., Blanco-Vega, H., & Mayorga-Vega, D. Adaptation of the Sport Satisfaction Instrument to the Physical Education lesson in Spanish secondary students: Psychometric properties and factor invariance. *Draft.*

X. **Casado-Robles, C.,** Mayorga-Vega, D., Guijarro-Romero, S., & Viciana, J. Validity of Xiaomi Mi Band 2, 3, 4 and 5 wristbands for assessing physical activity in 12-to-18-year-old adolescents. *Draft.*

XI. Viciana, J., **Casado-Robles, C.,** Guijarro-Romero, S., & Mayorga-Vega, D. Are activity wristbands and mobile apps valid for assessing physical activity in adolescents? *Draft.*

*JCR: Journal Citation Reports impact factor – ISI Web of Knowledge™; Subject category: Quartile
*SJR: SCImago Journal Rank impact factor – Scopus; Subject category: Quartile
*SPI: Scholarly Publishers Indicators – Books in Humanities and Social Sciences: ICEE; Subject category: Position (Quartile)

ABSTRACT

[RESUMEN]

ABSTRACT

Currently, there is strong evidence that both physical activity and sedentary behavior are powerful indicators of health and quality of life among students. Unfortunately, worldwide, about 80% of students do not meet the daily physical activity recommendations, and they also spend more than two-thirds of their daily waking time engaged in sedentary behavior. Therefore, school, especially Physical Education subject, has been considered a key setting for promoting students' healthy physical activity and sedentary behavior levels. However, to implement effective physical activity interventions for students in the school context, it is necessary to: (a) understand the determinants and factors influencing the acquisition of healthy physical activity habits; (b) consider what kind of teaching intervention is the most appropriate during Physical Education lessons; and (c) correctly distribute Physical Education lessons throughout the school year to achieve authentic outcomes. Nevertheless, Physical Education teachers have to face numerous limitations in the educational context, which make the design and development of adequate physical activity promotion programs for students a complex task. Consequently, it is necessary to seek feasible and effective strategies to overcome these constraints achieving significant results in the promotion of students' healthy physical activity and sedentary behavior levels.

The overall objective of the present Doctoral Thesis was to examine the effectiveness of innovative programs to increase students' habitual physical activity and reduce sedentary behavior from the Physical Education setting. The main findings from the present Doctoral Thesis suggest that: (1) a short-term Sport Education-based program carried out during Physical Education lessons only improves students' desire and willingness to participate in sports competitions, but it is not enough to improve students' self-reported nor objective habitual physical activity and sedentary behavior levels; (2) the innovative teaching units (intermittent, alternated and irregular) could help Physical Education teachers to design effective and viable programs to increase habitual physical activity levels and decrease sedentary behavior among students, overcoming some limitations related to planning in Physical Education; (3) an irregular teaching unit based on the Sport Education Model during Physical Education lessons and using school recess for developing the formal competition phase is effective for increasing objectively measured high-school students' physical activity levels and reduce sedentary behavior levels during this extracurricular period; (4) alternated inside-outside school teaching units are effective for improving high-school students' knowledge about the environment for practicing physical activity, perceived autonomy, enjoyment in Physical Education, autonomous motivation towards physical activity, the intention to be physically active, and self-reported physical activity, although they did not improve objectively measured physical activity nor sedentary behavior levels; (5) consumer-wearable activity tracker-based programs seem to be effective in promoting students' daily total steps and moderate-to-vigorous physical activity levels; (6) The designed CEPAF test is a valid and reliable measuring instrument to gather information about the knowledge of the immediate environment for the high-school students' practice of physical activity; (7) the adapted versions of the Sport Satisfaction Instruments referred to both Physical Education subject and only one Physical Education class are valid and reliable scales to be applied among high-school students; (8) the mobile apps Pedometer, Pacer, Google Fit and Apple Health, and the wristbands Samsung Watch Active 2, Apple Watch Series 5, and Xiaomi Mi Band 2, 3, 4, and 5 seem valid activity trackers for monitoring high-school students' steps under free-living conditions, although caution is warranted when using these wristbands for assessing students' total and moderate-to-vigorous physical activity levels. Furthermore, wristbands allow obtaining accurate information about students' compliance with the physical activity recommendations, offering a feasible alternative to the research-grade accelerometers, including for being used by teachers in the Physical Education setting during health promotion programs to provide accurate feedback to students.

In conclusion, the innovative teaching units effectively improve some factors influencing the acquisition of students' healthy physical activity and sedentary behavior habits, overcoming some limitations related to planning in Physical Education. However, other strategies such as proposed longer programs, complementing them with other extracurricular physical activity plans for leisure time, structured programs for recess, or including consumer-wearable activity trackers for providing students real-time feedback seem necessary to improve students' objective habitual physical activity and reduce sedentary behavior levels. This knowledge could help Physical Education teachers to design effective and feasible intervention programs to increase students' daily physical activity levels and reduce sedentary behavior from the Physical Education setting.

RESUMEN

Actualmente, existe una fuerte evidencia científica de que tanto la actividad física como la conducta sedentaria son potentes indicadores de salud y calidad de vida de los estudiantes. Desafortunadamente, a nivel mundial, aproximadamente el 80% de los estudiantes no cumplen con las recomendaciones diarias de actividad física, y además pasan más de dos tercios de su tiempo despierto en conducta sedentaria. La escuela, y especialmente la asignatura de Educación Física, ha sido considerada un contexto clave para la promoción de unos niveles de actividad física y conducta sedentaria saludables entre los estudiantes. Sin embargo, para llevar a cabo intervenciones en el contexto escolar que sean efectivas para promover la práctica de actividad física de los escolares, es necesario: (a) conocer los factores que influyen en la adquisición de hábitos saludables de actividad física; (b) considerar qué tipo de intervención docente es la más adecuada durante las clases de Educación Física; y (c) distribuir correctamente las clases de Educación Física a lo largo del curso escolar para conseguir resultados auténticos. Sin embargo, los profesores de Educación Física tienen que enfrentarse a numerosas limitaciones en el contexto educativo, que hacen que el diseño y el desarrollo de programas efectivos para la promoción de la práctica actividad física de los escolares sea una tarea compleja. Consecuentemente, es necesario buscar estrategias viables y efectivas para superar estas limitaciones logrando resultados significativos en la promoción de unos niveles de actividad física y conducta sedentaria saludables entre los escolares.

El objetivo general de la presente Tesis Doctoral fue examinar el efecto de programas de intervención innovadores para incrementar los niveles de actividad física habitual y reducir la conducta sedentaria de los escolares desde el contexto de la Educación Física. Los principales hallazgos de la presente Tesis Doctoral sugieren que: (1) un programa de corta duración siguiendo el Modelo de Educación Deportiva durante las clases de Educación Física sólo mejora el deseo y la voluntad del alumnado para participar en competiciones deportivas, pero no es suficiente para mejorar los niveles auto-reportados ni objetivos de actividad física habitual y de conducta sedentaria de los escolares; (2) las unidades didácticas innovadoras (intermitentes, alternadas, irregulares y reforzadas) podrían ayudar a los profesores de Educación Física a diseñar programas efectivos y viables para incrementar los niveles de actividad física habitual y disminuir la conducta sedentaria de los escolares,

superando algunas limitaciones relacionadas con la planificación en Educación Física; (3) una unidad didáctica irregular siguiendo el Modelo de Educación Deportiva durante las clases de Educación Física, y utilizando el recreo escolar para desarrollar la fase de competición formal, es eficaz para incrementar los niveles objetivos de actividad física y reducir la conducta sedentaria del alumnado de Educación Secundaria durante este periodo extraescolar; (4) las unidades didácticas alternadas dentro-fuera del centro escolar son efectivas para mejorar el conocimiento de los estudiantes de Educación Secundaria sobre el entorno para la práctica de la actividad física, la autonomía percibida, el disfrute durante la Educación Física, la motivación autónoma hacia la actividad física, la intención de ser físicamente activo y la actividad física auto-reportada, aunque no mejoran los niveles de actividad física ni de conducta sedentaria medidos objetivamente; (5) los programas de intervención que incluyen monitores portátiles de fitness parecen ser eficaces para incrementar los pasos diarios y los niveles de actividad física moderada-vigorosa de los escolares; (6) la prueba CEPAF es un instrumento de medida válido y fiable para recoger información sobre el conocimiento que poseen los escolares de Educación Secundaria de su entorno próximo para la práctica de actividad física; (7) las versiones adaptadas del *Sport Satisfaction Instrument*, referidas tanto a la asignatura de Educación Física en general como a una única clase de Educación Física, son escalas válidas y fiables para ser aplicadas con estudiantes de Educación Secundaria; (8) las aplicaciones móviles Pedometer, Pacer, Google Fit y Apple Health, y las pulseras de fitness Samsung Watch Active 2, Apple Watch Series 5 y Xiaomi Mi Band 2, 3, 4 y 5 parecen ser dispositivos válidos para monitorizar los pasos realizados por los estudiantes de Educación Secundaria en condiciones de vida libre, aunque se debe tener cuidado al utilizar estas pulseras de fitness para evaluar los niveles de actividad física total y actividad física moderada-vigorosa de los estudiantes. Además, estas pulseras de fitness permiten obtener información precisa sobre el cumplimiento de las recomendaciones de actividad física, ofreciendo una alternativa viable a los acelerómetros de investigación, incluso para ser utilizados por los profesores en la asignatura de Educación Física durante los programas de promoción de hábitos saludables para proporcionar información precisa a los estudiantes.

En conclusión, las unidades didácticas innovadoras son eficaces para mejorar algunos factores que influyen en la adquisición de hábitos saludables de actividad física y conducta sedentaria de los escolares, superando algunas limitaciones relacionadas con la planificación en Educación Física. Sin embargo, parece necesario adoptar otras estrategias, tales como llevar a cabo programas más largos, complementarlos con otros planes de actividad física extraescolar para el tiempo libre, programas estructurados para el recreo escolar, o la inclusión de monitores portátiles de fitness que proporcionen a los estudiantes información en tiempo real para incrementar los niveles habituales de actividad física y reducir los niveles de conducta sedentaria de los escolares medidos objetivamente. Este conocimiento podría ayudar a los profesores de Educación Física a diseñar programas de intervención eficaces y viables para incrementar los niveles de actividad física diarios y reducir la conducta sedentaria de los escolares desde el contexto de la Educación Física.

ABBREVIATIONS

[ABREVIATURAS]

ABBREVIATIONS [ABREVIATURAS]

Abbreviation	**Definition**
AF	*Actividad física*
AFMV	*Actividad física moderada-vigorosa*
AGFI	Adjusted goodness-of-fit index
AIC	Akaike information criterion
ANCOVA	Analysis of covariance
ANOVA	Analysis of variance
AVE	Average variance extracted
BREQ-3	Behavioral regulation in exercise questionnaire
CCI	*Coeficiente de correlación intraclase*
CEPAF	*Conocimiento del entorno para la práctica de actividad física*
CFI	Comparative fit index
CG	Control group
CMD	*Cuestionario multidimensional de deportividad*
CMIN/DF	Chi-squared fit index over degrees of freedom
CRCT	Cluster-randomized controlled trial
CS	*Conducta sedentaria*
CT	Controlled trial
DE	*Desviación estándar*
EF	*Educación Física*
EG	Experimental group
ES	Effect size
GFI	Goodness-of fit index
ICC	Intraclass correlation coefficient
IG	Innovative group
k	Kappa coefficient
LOA	Limits of agreement
MAE	Mean absolute error
MAPE	Mean absolute percentage error
MED	*Modelo de Educación Deportiva*
MIFA	Measurement of intention to be physically active questionnaire
MLM	Multilevel linear model with participants nested within classes

MVPA	Moderate-to-vigorous physical activity
NFI	Normalized fit index
NNFI	Non-normalized fit index
OMS	*Organización Mundial de la Salud*
P	Proportion of agreements
PA	Physical activity
PACE	Physician-based assessment and counseling for exercise questionnaire
PASSES	Perceived autonomy support scale for exercise settings
PE	Physical Education
RCT	Randomized controlled trial
RIQ	*Rango intercuartil*
RMSEA	Root mean square error of approximation
SB	Sedentary behavior
SCT	Social Cognitive Theory
SD	Standard deviation
SDT	Self-Determination Theory
SE	Sport Education
SR	School recess
SRMR	Standardized root mean square residual
SSI-PE	Sport satisfaction instrument adapted to Physical Education subject
SSI-PE-lesson	Sport satisfaction instrument adapted to one Physical Education lesson
TAD	*Teoría de la Autodeterminación*
TCM	Trans-Contextual Model
TG	Traditional group
TLI	Tucker Lewis index
WHO	World Health Organization
YAP	Youth activity profile questionnaire

INTRODUCTION

[INTRODUCCIÓN]

INTRODUCTION

1. Physical activity and sedentary behavior: Implications for students' health, guidelines, and current prevalence

The practice of physical activity (PA) and sedentary behavior (SB) are two crucial movement-related behaviors that characterize students' lifestyle during their waking time (Tremblay et al., 2016). On the one hand, PA is defined as any waking behavior consisting of bodily movement that requires energy expenditure, which can be categorized into a continuum from light to vigorous intensity (Caspersen et al., 1985; World Health Organization, WHO, 2020). On the other hand, SB is defined as any waking behavior characterized by a low energy expenditure (i.e., ≤1.5 METs; Tremblay et al., 2017; WHO, 2020).

Currently, there is strong evidence that both PA and SB are powerful indicators of health and quality of life among students (Chaput et al., 2020; Poitras et al., 2016). Specifically, regular practice of moderate-to-vigorous PA (MVPA) confers benefits not only for physical health (e.g., improved physical fitness and cardiometabolic health), but also for mental health (e.g., reduced anxiety or depression symptoms), and cognitive outcomes (e.g., academic performance) in young people (Sriram et al., 2021; WHO, 2020). Furthermore, students' total PA levels (i.e., light-to-vigorous PA or number of daily steps) have also recently been demonstrated to be potential health indicators (Poitras et al., 2016). However, independently of students' habitual PA levels, long periods of SB are related to health problems among students (Carson et al., 2016; WHO, 2020). Specifically, it is associated with unfavorable body composition (i.e., increased adiposity), higher cardiovascular disease risk, lower fitness, unfavorable measures of pro-social behavior, lower self-confidence, and reduced sleep duration among school-aged children (Carson et al., 2016; WHO, 2020).

Therefore, to achieve these benefits derived from regular PA practice, as well as to avoid those potential health problems derived from large amounts of time involved in SB, the WHO (2020) has established international PA and SB recommendations for school-aged children. Following these guidelines, school-aged children should achieve, on average, at least 60 minutes daily of MVPA across the week. Specifically, its practice should be mainly based on aerobic activities and include vigorous activities that strengthen muscles and bones (WHO, 2020). Regarding SB, although there is no evidence about a specific threshold, these guidelines also indicate that school-aged children should limit the amount of time spent being sedentary, particularly the amount of recreational screen time (WHO, 2020).

Unfortunately, worldwide, about 80% of students do not meet the daily PA recommendations (i.e., they are physically inactive; Guthold et al., 2020). Furthermore, they also spend more than two-thirds of their daily waking time engaged in SB (Ruiz et al., 2011). These levels are worrying considering that physical inactivity is one of the leading risk factors for non-communicable diseases and global mortality (International Sport and Culture Association, 2015; WHO, 2008), and the above-mentioned health-related problems caused by high levels of SB (WHO, 2020). Therefore, the promotion of adequate PA and SB levels has been considered an important public health priority, and a global action plan on PA has been developed to reverse these current trends (Gillis et al., 2013; WHO, 2018).

1.1. How to assess students' physical activity levels and sedentary behavior?

However, to check the effectiveness of these global policies and monitor their progress, it is necessary to accurately measure students' daily PA and SB levels across time (Brooke et al., 2014; Sirard & Pate, 2001; WHO, 2018). The instruments to assess PA and SB levels under free-living conditions are mainly divided into self-reported or objective methods (Adamo et al., 2009; Sirard & Pate, 2001). Self-reported measures, such as questionnaires, diaries, surveys, or interviews, are the most common method of measuring PA and SB levels because they provide important contextual information about PA and SB, their low-cost and their feasibility to be applied to large-scale studies (Dyrstad et al., 2014; Slootmaker et al., 2009). However, they may include misinterpretations due to they are based on students' perceptions and their ability to recall their behavior, and therefore, they should be used with caution among young people due to their limited validity and reliability (Ainsworth et al., 2015; Ekelund et al., 2011; Prince et al., 2020; Shephard 2003).

On the contrary, objective methods overcome some of these limitations, improving the validity of PA and SB measurement compared to self-reported measures. For example, removing recall bias, monitoring students' PA and SB consistently across time, and providing a more accurate measurement of PA and SB (Reilly et al., 2008; Shephard & Tudor-Locke, 2016; Troiano et al., 2014). Furthermore, systematic reviews about the effectiveness of interventions on students' PA and SB levels highlighted that it is necessary to objectively measure high-school students' PA levels across time instead of using self-reported methods (Brooke et al., 2014; Metcalf et al., 2012; Neil-Sztramko et al., 2021). Among the different objective methods are accelerometers, which are small devices that provide information about the intensity, frequency, and volume of PA (Dhurandhar et al., 2015; Garriguet & Colley, 2014). Furthermore, they have been highlighted as the most common, valid, and reliable method for objectively assessing students' PA and SB levels during free-living conditions, although they do not provide any contextual information about the type or domain of PA participation (Colley et al., 2019; Romanzini et al., 2014; Santos-Lozano et al., 2013; Shephard & Tudor-Locke, 2016; Trost et al., 2011). Therefore, due to previous evidence has demonstrated that students' self-reported measures are poorly correlated with objectively measured PA and SB levels (Prince et al., 2020; Shephard, 2003), some studies recommended using both methods to obtain the most complete information about students' PA and SB levels during intervention studies (Colley et al., 2019; Garriguet & Colley, 2014; Skender et al., 2016).

2. The role of Physical Education subject in the school setting for the promotion of students' healthy habits

Regarding the promotion of adequate PA and SB levels among school-aged children, the school context, as the institution where children and adolescents spend most of their time during the first two decades of life, has been considered a key setting (Hills et al., 2015; Neil-Sztramko et al., 2021; WHO, 2018b). In fact, the Spanish Ministry of Education and Vocational Training (2020) entrusts schools with the task of encouraging and consolidating students' healthy lifestyle habits during the Secondary Education stage. In particular, Physical Education (PE) subject plays a key role due to its mandatory nature in the educational curriculum and the fact that PE teachers have the knowledge and skills necessary to promote healthy PA and SB habits effectively (Association for PE, 2020; Viciana et al., 2015). Moreover, the Society of Health and Physical Educators (2014) highlights that PE plays an essential role in improving students' physical literacy, suggesting PE teachers to improve students' motivation, physical competence and the knowledge needed to achieve and maintain an active lifestyle.

Therefore, the transference of learning from PE lessons to students' daily life is another priority objective of the PE curriculum, which would also help to promote students' healthy PA and SB habits in the out-of-school setting (European Commission/EACEA/Eurydice, 2013; Society of Health and Physical Educators, 2014; Spanish Ministry of Education, Culture and Sport, 2015). Furthermore, for most students, PE is the only regular opportunity for PA practice, which contributes significantly to their daily PA levels and, thus, to their compliance with the daily MVPA recommendations (Association for PE, 2020; Mayorga-Vega et al., 2018; Uddin et al., 2020). Its importance is also reflected within the PE curricular framework, being the promotion of students' lifelong PA and reduction of SB one of their fundamental aims in most countries (e.g., Society of Health and Physical Educators, 2014; Spanish Ministry of Education, Culture and Sport, 2015). In this line, numerous school-based intervention programs have been implemented for the promotion of students' healthy PA and SB habits within the school context (Borde et al., 2017; Metcalf et al., 2012; Neil-Sztramko et al., 2021) and specifically in the PE subject (Hollis et al., 2017; Lonsdale et al., 2013).

Unfortunately, these previous programs have shown small or non-significant effectiveness on increasing high-school students' PA levels and/or reducing their SB levels (Altenburg et al., 2016; Borde et al., 2017; Friedrich et al., 2014; Neil-Sztramko et al., 2021; Russ et al., 2015). Furthermore, there is a wide variety of school-based interventions (e.g., PE-based programs, programs using additional opportunities for PA before school or during school recess, or multi-component programs), making it very difficult to draw conclusions about the most effective strategies of interventions that can elicit positive changes in students' PA and SB levels (Neil-Sztramko et al., 2021). In addition, most studies have been carried out with Primary students, with less evidence among high-school students (Neil-Sztramko et al., 2021). Therefore, it is necessary to design and implement other innovative school-based intervention programs to increase daily PA levels and reduce SB among high-school students (Borde et al., 2017; Rhodes et al., 2017).

3. Factors influencing the acquisition of students' healthy physical activity and sedentary behavior habits

In order to implement effective PA interventions for students in the school context, it is necessary to understand the determinants and factors influencing PA behavior (Sheeran et al., 2017). Firstly, the Social Cognitive Theory (SCT, Bandura, 2004) includes health-related knowledge as a core determinant on the design of PA promotion interventions, which has also been stood out as an enabling factor for PA practice in other PA promotion models (e.g., Welk, 1999). According to Bandura (2004), this health-related knowledge can be conveyed readily to students and supposes a first step toward generating a behavior change. Previous studies have shown that the student's knowledge about PA practice leads to a change in PA behavior and voluntary participation in PA (Ennis, 2015; Wang & Chen, 2019). Therefore, increasing students' knowledge about the environment for practicing PA could influence the amount of students' PA practice in their leisure time and, thus, their compliance with daily PA and SB recommendations because it is an enabling factor for PA practice (Bandura, 2004). Moreover, it would be meeting with the goal of the transferability of learning from the PE class to students' daily life (Spanish Ministry of Education, Culture and Sport, 2015; Viciana & Mayorga-Vega, 2018), making students capable of using the resources offered by the environment and structuring their own PA program during out-of-school time autonomously. In this sense, having valid and reliable tools to assess the acquisition of this learning (i.e., knowledge of the environment for practicing PA) seems necessary to check whether the PE subject is providing students with the tools needed to be active autonomously outside school, as well as to check the effectiveness of intervention programs to promote PA practice. (*Paper VIII*)

Furthermore, the Self-Determination Theory (SDT, Ryan & Deci, 2020) is a theoretical framework widely used to explain the antecedents and consequences of students' motivation to engage in PA. The SDT considers motivation as a multidimensional term with different levels along a continuum according to the degree of autonomy, ranging from more self-determined (i.e., autonomous) to less self-determined (i.e., controlled) forms of behavioral regulations (Ryan & Deci, 2020). On the one hand, autonomous motivation is related to the inherent pleasure and satisfaction provided by an activity, so when students are autonomously motivated, their reasons for engaging in PA are regulated by choice and volition. Conversely, students experience controlled motivation when their reasons for engaging in PA are pressured either internally (e.g., shame or guilt) or externally (e.g., tangible rewards). In this line, Ryan & Deci (2020) highlighted students' enjoyment as a relevant consequence associated with their autonomous motivation. Furthermore, enjoyment is considered a positive predictor of students' intention to engage in PA, as well as to reduce SB (Bai et al., 2018; Garcia et al., 2016; Pulido et al., 2014; Sánchez-Oliva et al., 2014). Specifically, the pupils' enjoyment during PE lessons would affect their PA levels in out-of-school contexts, tending to look for more opportunities to be physically active in out-of-school contexts, those with higher levels of enjoyment during PE lessons (Cox et al., 2008). For these reasons, the students' enjoyment should be considered as a crucial factor when designing intervention programs to promote healthy PA and SB habits from the PE setting. Consequently, it is also important to measure students' enjoyment during PE lessons using valid and reliable scales. (*Paper IX*)

Moreover, the SDT highlights the concept of basic psychological needs and their satisfaction as a fundamental factor for understanding students' motivation. Specifically, in order to enhance autonomous forms of motivation, the following three basic psychological needs should be satisfied: autonomy (referring to the need for initiative and ownership in one's own behavior), competence (referring to the need to feel capable of carrying out a certain behavior effectively), and relatedness (referring to the need of feeling connected and supported by significant others). According to the SDT, autonomy support is an important social context for encouraging higher levels of students' more self-determined regulations (Deci & Ryan, 2008). Specifically, teachers who support students' autonomy are favoring the satisfaction of the three psychological needs (i.e., autonomy, relatedness, and competence), and, therefore, promoting the students' most self-determined forms of motivation, while in contexts where autonomy support is not provided, these needs are thwarted (Ryan & Deci, 2020). This autonomy support in educational contexts is characterized by an environment that supports students' choice, initiation, and understanding, where students feel that they can participate in their own learning (Deci & Ryan, 2008). According to this framework, the use of an autonomy-supportive teaching style and the improvement of students' autonomous motivation are considered key determining factors related to the acquisition and maintenance of students' healthy PA habits (Teixeira et al., 2012). Consequently, school-based interventions with the aim of promoting students' lifelong PA should be based on psychological theories that highlight motivation and autonomy as core points of its tenets.

Besides, the Trans-Contextual Model (TCM, Hagger & Chatzisarantis, 2016) could be an ideal framework on which to base PE interventions focused on promoting students' motivation toward PA, and their engagement in PA practice out-of-school. Hagger & Chatzisarantis (2016) proposed three empirically testable propositions to explain how teachers can encourage students' PA participation during out-of-school contexts. The first proposition hypothesizes that students' perceived autonomy support from teachers can develop students' self-determined motivation toward in-class activities, and it will predict self-determined motivation toward similar activities within the educational context. The second one hypothesizes that self-determined motivation within the educational context will predict self-determined motivation toward similar activities in out-of-school contexts. Finally, the third proposition hypothesizes that self-determined motivation toward activities in an out-of-school context will predict future intention to engage in similar activities as well as actual behavioral engagement.

These propositions have received empirical support across multiple studies conducted in the PE setting (e.g., Barkoukis et al., 2010; Hagger et al., 2009; Hagger & Chatzisarantis, 2016; Yli-Piipari et al., 2018). These studies have found significant relationships between perceived autonomy support and self-determined motivation in PE (i.e., first proposition), between self-determined motivation in PE and in PA during leisure time (i.e., second proposition), and between self-determined motivation and intention to be physically active and actual PA engagement (i.e., third proposition). Therefore, according to this model, a key target is for teachers' autonomy-supportive behaviors and active styles of teaching in PE to be transferred to an increase of self-determined motivation, future intention, and actual participation in PA during out-of-school contexts (Chatzisarantis et al., 2008; González-Cutre et al., 2014; Wang & Chen, 2019).

To sum up, according to these theories, the use of a teaching methodology focused on motivational strategies and promoting students' autonomy seem necessary to carry out an effective program for promoting students' healthy PA and SB habits (Bandura, 2004; Hagger & Chatzisarantis, 2016; Ryan & Deci, 2020). Among others, these strategies may be based on developing students' knowledge for practicing PA, creating opportunities for students' choice, providing explanatory and meaningful rationales, proposing small-group activities, individualizing the tasks according to the students' level, or using an informational and non-controlling language (Cheon et al., 2018; Hagger & Chatzisarantis, 2016; Wang & Chen, 2019). Specifically, PE-based intervention studies using teachers' autonomy-supportive behaviors have shown to be an effective way to promote some determinants like students' enjoyment in PE lessons (Abós et al., 2017; Sevil et al., 2016), autonomous motivation towards PA (McDavid et al., 2012; Sevil et al., 2020), intention to be physically active (Chicote-López et al., 2018; Mavropoulou et al., 2018), and ultimately, actual PA practice during leisure time (González-Cutre et al., 2014; Wang, 2017).

3.1. Effectiveness of the Sport Education model for the promotion of students' healthy habits

In this sense, the PE teacher's intervention (e.g., methodology, teaching strategies, teaching styles or pedagogical models used) has an important influence on the promotion of students' PA levels (McKenzie & Lounsbery, 2014). Therefore, it have to be considered when organizing the teaching-learning process to achieve the educational aims pursued and guarantee high-quality PE lessons (Metzler, 2017; Scheuer et al., 2014; United Nations Educational, Scientific and Cultural Organization, UNESCO, 2015). Among others, an autonomy-supportive curricular model, which could be an appropriate curriculum for promoting a physically active lifestyle, is the Sport Education model proposed by Siedentop (1994). It is based on student-centered learning through a cooperative and constructivist pedagogy, which is facilitated by working in small groups in which each member has a role, and pursuing the students' autonomy during PE lessons and in out-of-school contexts (Siedentop et al., 2019).

Specifically, the Sport Education model emerged to provide positive, quality sporting experiences within PE lessons, centered on taking the competition to the educational context and taking advantage of all the educational values of sport such as respect, sportsmanship, or teamwork (Siedentop et al., 2019). This pedagogical approach is based on six main features: (a) the teaching unit is organized as a sport "season"; (b) students develop a "sense of affiliation" becoming members of a team (i.e., small group works); (c) the season has a "formal competition" phase, where small-sided games adapted to the students' level take place accordingly with the schedule of competitions; (d) students assume different roles other than players (e.g., coach, captain or referee); (e) a "record-keeping" of behaviors and results takes place throughout the formal competition to give feedback and establish goals; (f) and, the season is developed in a "festive" atmosphere, creating distinctive team elements, publishing photographs or celebrating achievements (Siedentop et al., 2019).

It is probably one of the most widely implemented and researched instructional approaches in the world, demonstrated by the numerous systematic reviews carried out internationally (e.g., Bessa et al., 2019; Chu & Zang, 2018; Evangelio et al., 2018; or Wallhead et al., 2013), and specifically in Spain (Guijarro, Rocamora, Evangelio et al., 2020). That previous evidence about the Sport Education model showed that it provides several positive effects in the factors mentioned above influencing students' PA practice, such as improvements in students' enjoyment of PE (Bessa et al., 2019), motivation towards PA practice (Burgueño et al., 2020; Chu & Zang, 2018) or the intention to be physically active (Cuevas et al., 2016; Martínez de Ojeda et al., 2016). Furthermore, it is also effective for increasing students' in-class MVPA levels (Guijarro, Rocamora, González-Víllora, et al., 2020; Perlman, 2012; Puente-Maxera et al., 2020; Rocamora et al., 2019) and improving students' self-reported MVPA levels during leisure time such as school recess or after-school time (Coolkens et al., 2018; Wallhead et al., 2010, 2014). However, to our knowledge, no previous studies are assessing the effect of a Sport Education season on the students' habitual PA and SB levels using objective instruments (i.e., accelerometers), which could provide further evidence on the actual behavioral changes resulting from the motivational impacts of the Sport Education model (Chu & Zang, 2018). (*Paper I*)

4. Innovative school-based programs to promote students' healthy physical activity and sedentary behavior habits

However, in addition to considering the determinants and factors influencing PA behavior (Sheeran et al., 2017) and the teacher intervention during PE lessons (McKenzie & Lounsbery, 2014), the temporal distribution of PE lessons into different teaching units throughout the school year must also be considered when organizing the teaching-learning process, to achieve the educational aims pursued and guarantee high-quality PE lessons (Viciana, 2002; Viciana & Mayorga-Vega, 2016).

4.1. Limitations related to planning in Physical Education

Nevertheless, PE teachers have to face numerous limitations in the educational context, which make the design and development of effective PA promotion programs for students a complex task. Firstly, the educational curriculum indicates that PE teachers have to develop a large volume of contents throughout the school year. Specifically, apart from the promotion of students' healthy PA and SB habits, they must also include: the development and maintenance of students' cardiorespiratory fitness; the knowledge of technical and tactical skills related to games and sport; gymnastics activities; dance and rhythmic activities; or outdoor and adventure tasks (e.g., Society of Health and Physical Educators, 2014; Spanish Ministry of Education, Culture and Sport, 2015). It is challenging considering that the PE subject has a very limited presence in the education system worldwide (European Commission/EACEA/Eurydice, 2013; Hardman et al., 2014; WHO, 2018b). Specifically, on average, in the European Union, about two hours per week are devoted to PE in Secondary Education, which would mean only two 60-minute lessons per week (WHO, 2018b). Nevertheless, it should be noted that, fortunately, Spain is gradually increasing the time devoted to PE subject in the curriculum incorporating a third weekly lesson, although only in some grades and Autonomous Communities in the last year (e.g., Department of Education and Sports of the Andalusian Regional Government, 2021; Regional Ministry of Education and Youth of the Government of Madrid, 2020). However, although increasing the number of compulsory PE hours in all grades of Secondary Education would play a vital role in helping students achieve international recommendations and create an active nation, this is not the current reality globally (Uddin et al., 2020; WHO, 2018b).

Furthermore, the distribution of PE classes is not based on PA-related criteria, but it is simply regulated by correctly adjusting the schedule of the school's teaching staff (Viciana et al., 2014). Therefore, it is possible that PE lessons were taught two consecutive days without leaving the necessary time for students' rest from the previous lesson, as well as leaving too much time between the last session of one week and the next one. Besides, the fact that PE classes can be constantly interrupted throughout the school year for different reasons is also a limitation (Viciana et al., 2014). For example, weather aspects (i.e., rain or extreme heat/cold), several holiday periods during the year (e.g., Christmas, Easter), academic activities (e.g., school trips or workshops) or coincidence of PE teachers' timetable sharing the same materials and facilities, that hinder the PE teaching-learning process (Guijarro-Romero et al., 2019; Viciana et al., 2014). Lastly, another limitation lies in the difficulty of controlling students' PA practice and SB outside the school when the PE teacher is not present. Thus, the teachers would need external aids to test the programs' effectiveness in increasing students' habitual PA and reducing SB in their leisure time.

Unfortunately, due to these PE-based planning limitations, PE teachers tend to plan short teaching units that are insufficient for achieving significant results regarding students' healthy PA and SB levels promotion (Viciana & Mayorga-Vega, 2016). That is, PE teachers consider the annual planning as a group of empty time segments to be filled with different contents (i.e., teaching units), basing their choice on personal criteria or available resources, instead of being really focused on achieving meaningful objectives for the subject. Moreover, the students' transference of significant learning between different PE contents or students' daily life does not occur due to the lack of coherence in the whole planning. Therefore, it is necessary to seek feasible and effective strategies to overcome these constraints achieving significant results in the promotion of students' PA practice (Borde et al., 2017; Viciana & Mayorga-Vega, 2016). (*Paper II*)

4.2. Innovative teaching units as a solution via modifying the curriculum management time in Physical Education

In this line, the innovative teaching units proposed by Viciana & Mayorga-Vega (2016) intend to give teachers the opportunity to organize teaching-learning periods better. These teaching units are based on the original principles of educational planning of flexibility, effectiveness and based on objectives (UNESCO, 1962) to propose an innovative distribution of time according to the educational aim pursued (Viciana & Mayorga-Vega, 2016). Therefore, using these alternative models may provide solutions to solve different problems related to the promotion of PA practice and the reduction of SB within the PE curriculum.

Firstly, the "Intermittent teaching units" are based on devoting only a specific period of the PE lesson (instead of the complete lesson) to a particular curricular objective, leaving the rest of the class for another aim (Viciana & Mayorga-Vega, 2016). That is, traditional lessons where the entire session is devoted to a single curricular objective are replaced by fragments (e.g., 5 to 15 minutes) in each class over an extended period of time (e.g., an academic term). In this way, students achieve progressive learning that is maintained and affixed as they work on the same objective over many lessons. Furthermore, it allows developing two or more related curricular goals during the same lesson, connecting students' learning between different contents of the PE subject, and facing the problem related to the limited time devoted to PE in the education system. An example of an intermittent teaching unit could consist of the promotion of students' healthy PA and SB levels only during the cool-down part of the PE lessons, dedicating the rest of the lesson to other different curricular content (e.g., the development and maintenance of students' cardiorespiratory fitness; or the knowledge of technical and tactical skills related to games and sport). (*Paper VI*)

Secondly, the "Alternated teaching units" are based on delivering PE lessons belonging to two different teaching units, about complementary contents, alternatively while putting them into practice (Viciana & Mayorga-Vega, 2016). Implementing two teaching units regarding complementary contents makes students aware that both contents are based on the same learning principles and avoid unconnected learning perceived by students in traditional and isolated units (Viciana & Mayorga-Vega, 2016). Furthermore, the transference of learning from the classroom to students' daily life could also be facilitated through the application of these alternated teaching units (Viciana & Mayorga-Vega, 2018). An example of alternated teaching units could consist of teaching students how to practice PA in-school (i.e., inside teaching unit) connected to how to use the environment that surrounds the school center and their particular community to practice PA (i.e., outside teaching unit), providing them authentic situational PA practices, and facilitating students a tool for practicing PA autonomously in their free time (Viciana & Mayorga-Vega, 2018). Besides, it allows facing the problem regarding teachers' timetable coincidence and the necessity of using the same facilities and materials. (*Papers IV-V*)

Lastly, the "Irregular teaching units" are based on the characteristics of flexible and dynamic planning at their maximum level to achieve authentic objectives and results like the improvement of daily PA levels or reduction of SB habits (Viciana & Mayorga-Vega, 2016). One of the main advantages is that it allows reinforcing and increasing the number of learning opportunities for students by using other extracurricular contexts beyond PE lessons, for example, school recess, after school time or weekends. Therefore, due to increasing the frequency and duration of PE lessons is not feasible given the curriculum demands, it is an excellent option to provide additional opportunities for PA practice before, during and after the school day. It allows increasing students' autonomy towards PA practice in their leisure time (González-Cutre et al., 2018) and helping students reach the recommended PA levels (Carson et al., 2014; Neil-Sztramko et al., 2021; WHO, 2018b). Furthermore, school-based interventions involving curricular time (e.g., PE lessons) and extracurricular time (e.g., school recess) have been considered one of the most promising approaches to increase students' PA levels (Kriemler et al., 2011; Murillo-Pardo et al., 2013). Specifically, school recess as a non-curricular time during school can be an ideal context in providing additional PA opportunities for high-school students (Parrish et al., 2013). Some organizations such as the Centers for Disease Control and Prevention and the Society of Health and Physical Educators (2017), and the Institute of Medicine (2013) highlight the advantage that the school recess offers daily practice opportunities outside of PE, which usually has a frequency of only two days a week. However, it is necessary to provide students with a certain autonomy and responsibility during PE lessons by using autonomy-supportive teaching styles, making them competent for practicing PA in their leisure time (González-Cutre et al., 2018; Sevil et al., 2020; Viciana & Mayorga-Vega, 2018). Therefore, an example of an irregular teaching unit to promote students' daily PA levels could consist of a Sport education-based season (Siedentop et al., 2019), using both compulsory PE lessons and extracurricular school recess periods where the competition phase of the model would take place. (*Paper III*)

5. Consumer-wearable activity trackers and students' physical activity and sedentary behavior habits

In addition to all the factors mentioned above, previous studies have recognized that self-monitoring behavior is another essential technique for PA practice promotion (Michie et al., 2009, 2013). For this reason, the use of consumer-wearable activity trackers (e.g., smartwatches, activity wristbands, or pedometers) could be ideal devices to track students' PA and SB levels, providing them this important real-time feedback (Strath & Rowley, 2018). These consumer-wearable activity trackers are electronic devices worn on the body as an accessory monitoring and recording daily PA and fitness-related metrics, and providing users real-time behavioral feedback (Ruiz & Goransson, 2015; Strath & Rowley, 2018). They usually integrate an accelerometer or pedometer to automatically track physical movements and their outputs are generally based on step counts, the amount and intensity of PA, energy expenditure, periods of inactivity, sleep time, or heart rate (Franssen et al., 2020; Mercer et al., 2016). Moreover, these devices can be synchronized with their specific smartphone apps, PCs, or websites to obtain more detailed feedback over days and weeks (Alley et al., 2016). Besides the real-time feedback, these devices often include other features that also may be facilitators of users' positive behavior change such as personalized goal-setting (i.e., a goal based on daily steps or minutes of PA), self-efficacy, peer comparison, or social support (Brickwood et al., 2019; Patel et al., 2015; Rich & Miah, 2016). Furthermore, these consumer-wearable activity trackers have become increasingly popular over recent years, being reflected in a significant increase in sales and strong demand from society worldwide every year (International Data Corporation, 2020).

5.1. Effectiveness of consumer-wearable activity trackers programs for the promotion of students' healthy habits

The popularity of these devices has led the scientific community to carry out studies that integrate consumer-wearable activity trackers for promoting health-enhancing students' PA and SB levels, obtaining good results for students' motivation and PA levels (e.g., Bronikowski et al., 2016; Duncan et al., 2012; Evans et al., 2017; Eyre et al., 2016; Grao-Cruces et al., 2016; or Hardman et al., 2011). However, every single primary study about the effectiveness of consumer-wearable activity trackers for promoting PA and/or reducing SB levels only constitutes a specific portion of the total evidence. Moreover, conflicting findings can often be found between different primary studies. For example, Eyre et al. (2016) found a positive effect, while Lubans et al. (2011) found a negative or null effect for improving PA levels, and Jago et al. (2006) found a positive effect in one wave but a null effect in the other wave for reducing SB levels. By contrast, systematic reviews and meta-analyses allow for an objective analysis of all the available evidence about a specific topic, making sense of the often conflicting results found and providing estimations with greater power and precision than each primary study (Cooper et al., 2019; Schmidt & Hunter, 2021). In this sense, systematic reviews and meta-analyses are considered the lens for appraising, synthesizing, and applying scientific evidence (Murad et al., 2016). Consequently, knowing the effects of consumer-wearable activity tracker-based programs and its associated behavior change strategies on daily objectively measured PA and SB among students would help PE teachers to design effective intervention programs to increase their daily PA levels and reduce SB. (*Paper VII*)

Concerning the educational setting, the use of these consumer-wearable activity trackers could solve some of the limitations facing the PE subject, such as, for example, the impossibility of controlling the students' PA practice outside the school when the teacher is not present. Therefore, using them from the PE setting to promote students' PA in their leisure time seems feasible and desirable. An excellent option for carrying out programs using consumer-wearable activity trackers is combined with the intermittent teaching units explained above (Viciana & Mayorga-Vega, 2016). Specifically, it could consist of the promotion of students' healthy PA and SB levels through an intermittent teaching unit during the cool-down part of the PE lessons (i.e., the last 10-15 minutes of the lessons) incorporating activity wristbands for the promotion of students' PA practice and the reduction of SB in their leisure time. Furthermore, in addition to the use of an intermittent teaching unit with consumer-wearable activity trackers for PA levels self-monitoring, the determinants and factors influencing students' PA behavior should be considered when designing interventions as they are vital factors in any behavioral change program (Lubans et al., 2016; Sheeran et al., 2017). Therefore, as mentioned above, the design of this proposal should be based on psychological theories such as SCT (Bandura, 2004), SDT (Ryan & Deci, 2020), or the TCM (Hagger & Chatzisarantis, 2016), as well as effective behavior change techniques such as goal-setting strategies or counseling sessions (Michie et al., 2013; Rhodes et al., 2017) to promote students' PA and SB behavioral change. Specifically, students' health-related knowledge (e.g., PA and SB recommendations, benefits or regular PA practice, or possibilities for healthy leisure time activities and sedentary activities to be avoided), motivation towards PA or satisfaction of basic psychological needs (i.e., autonomy, competence and relatedness) are influential variables to consider in this complex process of acquiring healthy active behaviors (Bandura, 2004; Hagger & Chatzisarantis, 2016; Ryan & Deci, 2020). (*Paper VI*)

5.2. Validity of consumer-wearable activity trackers for estimating students' physical activity and sedentary behavior

Nowadays, consumer-wearable activity trackers are increasingly being used in research as a motivational strategy during PA promotion programs and for monitoring students' PA and SB levels (Gorzelitz et al., 2020; Henriksen et al., 2018). Although, as previously mentioned, research-grade accelerometers have been highlighted as the most common and valid method for objectively assessing high-school students' PA levels during free-living conditions (Romanzini et al., 2014; Shephard & Tudor-Locke, 2016; Van Hecke et al., 2016), they are usually costly instruments, and in addition, they are unattractive and not interactive with the users (Actigraph Corporation, 2021), which make them not very useful to both purposes of monitoring and promoting students' PA practice (Šimůnek et al., 2019). On the contrary, these new consumer-wearable activity trackers share elements of research-based devices, and they are generally cheaper, more interactive, and more user-friendly (Gorzelitz et al., 2020; Henriksen et al., 2018). Therefore, they seem a feasible alternative to research-grade accelerometers for objectively assessing and promote students' higher PA and lower SB levels (Da Silva et al., 2015; Tudor-Locke et al., 2011). However, although the use of consumer-wearable activity trackers to monitor PA and SB levels is increasingly widespread, its validity has not been sufficiently studied, especially among school-aged children (Cosoli et al., 2020; Evenson et al., 2015; Fuller et al., 2020; Gorzelitz et al., 2020). Therefore, further research investigating their validity for measuring students' PA and SB levels is needed. (*Paper X-XI*)

AIMS

[OBJETIVOS]

AIMS

The overall aim of the present Doctoral Thesis was to examine the effectiveness of innovative programs to increase students' habitual PA and reduce SB levels from the PE setting.

Specific

1. To examine the effect of a Sport Education-based teaching unit in PE on the high-school students' intention to be physically active and habitual PA and SB levels. (*Paper I*)
2. To design a proposal of PE-based innovative teaching units to promote students' PA and reduce SB levels. (*Paper II*)
3. To assess the effect of a Sport Education-based irregular teaching unit on the high-school students' PA and SB levels during school recess. (*Paper III*)
4. To evaluate the effect of two alternated teaching units (i.e., inside and outside the school center) on high-school students' knowledge of their environment for practicing PA, perceptions of autonomy support, autonomous and controlled motivation towards PA, enjoyment of PE, intention to be physically active, and habitual PA and SB levels. (*Papers IV-V*)
5. To design an intermittent teaching unit combined with consumer-wearable activity trackers to promote students' healthy habitual PA and SB levels. (*Paper VI*)
6. To assess the effects of consumer-wearable activity tracker-based programs on daily objectively measured PA and SB among school-aged children. (*Paper VII*)
7. To develop and validate an *ad hoc* written test to evaluate the students' knowledge about the environment close to the educational center for the practice of PA (CEPAF) in high-school students. (*Paper VIII*)
8. To adapt and validate the Spanish version of the Sport Satisfaction Instrument adapted to PE (SSI-PE) to assess high-school students' enjoyment towards PE subject and the PE lesson that just finished. (*Paper IX*)
9. To assess the convergent validity of activity trackers to estimate high-school students' daily steps and PA levels under free-living conditions. (*Papers X-XI*)

OBJETIVOS

El objetivo general de la presente Tesis Doctoral fue examinar el efecto de programas de intervención innovadores para incrementar los niveles de AF habitual y reducir la CS de los escolares desde el contexto de la EF.

Específicos

1. Examinar el efecto de una unidad didáctica del Modelo de Educación Deportiva en EF sobre la intención de ser físicamente activo y los niveles habituales de AF y CS de los estudiantes de Educación Secundaria. (*Artículo I*)
2. Diseñar una propuesta de unidades didácticas innovadoras en EF para incrementar los niveles de AF y disminuir la CS en escolares. (*Artículo II*)
3. Evaluar el efecto de una unidad didáctica irregular basada en el Modelo de Educación Deportiva sobre los niveles de AF y CS de los alumnos de Educación Secundaria durante el recreo escolar. (*Artículo III*)
4. Comprobar el efecto de dos unidades didácticas alternadas (dentro y fuera del centro escolar) sobre el conocimiento de los estudiantes de Educación Secundaria sobre su entorno para la práctica de AF, la percepción de apoyo a la autonomía, la motivación hacia la AF, el disfrute durante las clases de EF, la intención de ser físicamente activo y los niveles de AF y CS habituales. (*Artículos IV-V*)
5. Diseñar una unidad didáctica intermitente en combinación con monitores de fitness para promover unos niveles habituales saludables de AF y CS en los estudiantes. (*Artículo VI*)
6. Evaluar el efecto de programas de intervención que utilizan monitores de fitness sobre los niveles diarios objetivos de AF y CS en escolares. (*Artículo VII*)
7. Construir y someter a un proceso de validación una prueba escrita *ad hoc* para la evaluación del conocimiento del entorno próximo al centro educativo para la práctica de AF (CEPAF) en estudiantes de Educación Secundaria (*Artículo VIII*)
8. Adaptar y validar la versión española del *Sport Satisfaction Instrument* adaptado a EF (SSI-EF) para evaluar el disfrute de los estudiantes de Educación Secundaria hacia la asignatura de EF y en la clase de EF que acaban de realizar. (*Artículo IX*)
9. Examinar la validez convergente de las pulseras de fitness para estimar los pasos y los niveles de AF de los estudiantes de Educación Secundaria en condiciones de vida real. (*Artículos X-XI*)

MATHERIAL AND METHODS

[MATERIAL Y MÉTODOS]

MATERIAL AND METHODS

The material and methods section of the present Doctoral Thesis is summarized in Table 1. This table includes the most relevant methodological features from the scientific studies that compose the present Doctoral Thesis. For further information of any study, please check the material and methods sections of the corresponding papers.

Table 1. Summary of the main methodological features of the studies of the present Doctoral Thesis

Paper	Design	Participants			Procedure	Me
		Number	Grade	Years		Ins
I. Is the Sport Education Model effective in improving sportsmanship and PA in students?	CRCT	114 (62 females/ 52 males)	Second and fourth grade high-school	13-16	Control group: 6 weeks, 2 lessons/week, 60 min, traditional methodology Experimental group: 6 weeks, 2 lessons/week, 60 min, Sport Education	CM PA MI phy GT (PA
III. Sport Education-based irregular teaching unit and students' PA during school recess	CRCT	165 (77 females/ 88 males)	Second and fourth grade high-school	13-16	Control group: 6 weeks, 2 lessons/week, 60 min, traditional methodology Traditional group: 6 weeks, 2 lessons/week, 60 min, Sport Education Irregular group: 6 weeks, 2 lessons/week, 60 min, Irregular Sport Education	GT (PA *Ad* (In par con rec
IV. Is it more fun to do PE outside the school center?	CRCT	179 (94 females/ 85 males)	Second and third grade high-school	13-15	Traditional group: 4 weeks, 2 sessions/week, 60 min, Inside Alternated group: 4 weeks, 2 sessions/week, 60 min, Inside-outside	SSI PE SSI (En par

Table 1. Continued

V. Effect of an inside-outside school alternated teaching unit of knowledge of the environment for practicing PA: A cluster randomized control trial	CRCT	179 (94 females/ 85 males)	Second and third grade high-school	13-15	Traditional group: 4 weeks, 2 sessions/week, 60 min, Inside Alternated group: 4 weeks, 2 sessions/week, 60 min, Inside-outside	CE the pra PA sup BR tow Int leis que to b YA GT (PA
VII. Effects of consumer-wearable activity tracker-based programs on objectively measured daily PA and SB among school-aged children: A systematic review and meta-analysis	Systematic review and meta-analysis	45 studies included in the systematic review (41 studies in the meta-analysis)			8 electronic bibliographic databases and 4 additional modes of searching	Da Mo PA Tot SB

Table 1. Continued

VIII. Knowledge about the environment to practice PA in schoolchildren (CEPAF). Development and validation of an objective written test of multiple-choice	Content validity	256 (128 females/ 128 males)	Second and third grade high-school	13-15	1. Construction and definition of the test 2. Content validation by experts 3. Pilot evaluation 4. Refinement and application of the final instrument 5. Evaluation of test-retest reliability 6. Evaluation of discriminant validity	CE the pra
IX. Adaptation of the Sport Satisfaction Instrument to the PE lesson in Spanish Secondary students: Psychometric properties and factor invariance	Construct validity	410 (204 females/ 206 males)	First to third grade high-school	11-17	1. Adaptation of the questionnaire 2. Content validation by experts 3. Confirmatory factor analysis and internal consistency 4. Factorial invariance between gender	SS PE SS (Er par

Table 1. Continued

X. Validity of Xiaomi Mi Band 2, 3, 4 and 5 wristbands for assessing PA in 12-to-18-year-old adolescents	Convergent validity	70 (35 females/ 35 males)	First to fourth grade high-school and first grade of baccalaureate	12-18	1-day students' PA assessment under free-living conditions using an accelerometer and wristbands	Xia and (St wG acc (St
XI. Are activity wristbands and mobile apps valid for assessing PA in adolescents?	Convergent validity	Study 1: 62 (30 females/ 32 males) Study 2: 61 (30 females/ 31 males)	First to fourth grade high-school and first grade of baccalaureate	12-18	Study 1: Assessment of students' steps count under structured free-living conditions using wristbands and mobile apps Study 2: 1-day students' PA assessment under free-living conditions using an accelerometer and wristbands	Pec Go He Sar 2, A and wri Vid wG acc (St

Note. Papers II and VI are not included in this table because they are theoretical studies. PA: Physical activity;
CMD: *Cuestionario Multidimensional de Deportividad*; PACE: Physician-based Assessment and Counseling for Exerc
Physically Active questionnaire; SB: Sedentary Behavior; ANOVA: Analysis of variance; MLM: Multilevel Lin
classes; ANCOVA: Analysis of covariance; CEPAF: *Conocimiento del entorno para la práctica de actividad física* que
Support Scale for Exercise Settings; BREQ-3: Behavioral Regulation in Exercise Questionnaire; YAP: Youth Act
Correlation Coefficient by a two-way mixed effects model with absolute agreement and multiple measurements
Satisfaction Instrument adapted to Physical Education subject; SSI-PE-lesson: Sport Satisfaction Instrument ad
lesson; LOA = Limits of Agreement; MAE = Mean Absolute Error; MAPE = Mean Absolute Percentage Error; IC
by a two-way random effects model with absolute agreement and single measurement; P = Proportion of agreement
[a]Two-way mixed nested ANOVA/ANCOVA implies both within-group and between-group factors.

MATERIAL Y MÉTODOS

El apartado de materiales y métodos de la presente Tesis Doctoral se resume en la Tabla 1. Esta tabla incluye las características metodológicas más relevantes de los estudios científicos que componen la presente Tesis Doctoral. Para obtener más información sobre cualquier estudio, consulte los apartados de material y métodos de los estudios correspondientes.

Tabla 1. Resumen de las principales características metodológicas de los estudios de la presente Tesis Doctoral

Estudio	Diseño	Participantes			Procedimiento
		Número	**Curso**	**Edad**	
I. ¿Es efectivo el Modelo de Educación Deportiva para mejorar la deportividad y AF en escolares?	CRCT	114 (62 mujeres/ 52 varones)	Segundo y cuarto curso de Educación Secundaria	13-16	Grupo control: 6 semanas, 2 sesiones/semana, 60 min, metodología tradicional Grupo experimental: 6 semanas, 2 sesiones/semana, 60 min, Educación Deportiva
III. Unidad didáctica irregular de Educación Deportiva y niveles de AF de los escolares durante el recreo	CRCT	165 (77 mujeres/ 88 varones)	Segundo y cuarto curso de Educación Secundaria	13-16	Grupo control: 6 semanas, 2 sesiones/semana, 60 min, metodología tradicional Grupo tradicional: 6 semanas, 2 sesiones/semana, 60 min, Educación Deportiva Grupo irregular: 6 semanas, 2 sesiones/semana, 60 min, Educación Deportiva irregular
IV. ¿Es más divertido realizar EF fuera del centro escolar?	CRCT	179 (94 mujeres/ 85 varones)	Segundo y tercer curso de Educación Secundaria	13-15	Grupo tradicional: 4 semanas, 2 sesiones/semana, 60 min, dentro Grupo alternado: 6 semanas, 2 sesiones/semana, 60 min, dentro-fuera

Tabla 1. Continuación

V. Efecto de una unidad didáctica alternada dentro-fuera del centro escolar para el conocimiento de los escolares del entorno para la práctica de AF: Un ensayo controlado y aleatorizado por grupos	CRCT	179 (94 mujeres/ 85 varones)	Segundo y tercer curso de Educación Secundaria	13-15	Grupo tradicional: 4 semanas, 2 sesiones/semana, 60 min, dentro Grupo alternado: 6 semanas, 2 sesiones/semana, 60 min, dentro-fuera
VII. Efecto de los programas que incorporan monitores portátiles de fitness sobre la AF y la CS diaria de los escolares medidas objetivamente: Una revisión sistemática y meta-análisis	Revisión sistemática y meta-análisis	45 estudios incluidos en la revisión sistemática (41 estudios incluidos en el meta-análisis)			8 bases de datos electrónicas y 4 modos adicionales de búsqueda

Tabla 1. Continuación

VIII. Conocimiento del entorno para la práctica de AF en escolares (CEPAF): Desarrollo y validación de una prueba escrita objetiva de elección múltiple	Validez de contenido	256 (128 mujeres/ 128 varones)	Segundo y tercer curso de Educación Secundaria	13-15	1. Construcción y definición de la prueba 2. Validación de contenido por expertos 3. Pilotaje 4. Depuración y aplicación del instrumento definitivo 5. Evaluación de la fiabilidad test-retest 6. Evaluación de la validez discriminante
IX. Adaptación del *Sport Satisfaction Instrument* a la clase de EF en estudiantes de Secundaria españoles: Propiedades psicométricas e invarianza factorial	Validez de constructo	410 (204 mujeres/ 206 varones)	De primer a tercer curso de Educación Secundaria	11-17	1. Adaptación del cuestionario 2. Validación de contenido por expertos 3. Análisis factorial confirmatorio y consistencia interna 4. Invarianza factorial por género

Tabla 1. Continuación

X. Validez de las pulseras Xiaomi Mi Band 2, 3, 4 y 5 para monitorizar la AF en adolescentes de 12-18 años	Validez convergente	70 (35 mujeres/ 35 varones)	De primer a cuarto curso de Educación Secundaria y primer curso de Bachillerato	12-18	Medición de la AF realizada por los escolares durante un día en condiciones de vida libre utilizando un acelerómetro y pulseras de fitness	P B (I A (I
XI. ¿Son válidos los monitores portátiles de fitness y las aplicaciones móviles para monitorizar la AF en adolescentes?	Validez convergente	Estudio 1: 62 (30 mujeres/ 32 varones) Estudio 2: 61 (30 mujeres/ 31 varones)	De primer a cuarto curso de Educación Secundaria y primer curso de Bachillerato	12-18	Estudio 1: Medición de los pasos realizados por los estudiantes en condiciones de vida libre estructurada con pulseras de fitness y aplicaciones móviles Estudio 2: Medición de la AF realizada por los escolares durante un día en condiciones de vida libre con acelerómetro y pulseras de fitness	A P y P W W M G (I A (I

Nota. Los artículos II y VI no están incluidos en esta tabla porque son estudios teóricos. AF: Actividad Física; C grupos naturales; CMD: Cuestionario Multidimensional de Deportividad; PACE: *Physician-based Assessment and* Intencionalidad para ser Físicamente Activo; CS: Conducta Sedentaria; ANOVA: Análisis de la varianza; MLM: M anidados dentro de las clases; ANCOVA: Análisis de la covarianza; CEPAF: Prueba de Conocimiento del entorno Escala del Apoyo a la Autonomía Percibido en Contextos de Ejercicio; BREQ-3: Cuestionario de Regulación de la *Youth Activity Profile*; ICC = Coeficiente de Correlación Intraclase con un modelo mixto de dos factores de tipo promedio; EF = Educación Física; SSI-EF: *Sport Satisfaction Instrument* adaptado a la asignatura de Educación Físic adaptado a una sesión particular de Educación Física; LOA = Límites de concordancia; MAE = Error absoluto medi ICC (2,1) = Coeficiente de Correlación Intraclase con un modelo aleatorio de dos factores de tipo acuerdo absoluto k = Coeficiente de Kappa.

[a]ANOVA/ANCOVA anidado mixto de dos factores implica ambos factores intra-grupo y entre-grupo.

RESULTS AND DISCUSSION

[RESULTADOS Y DISCUSIÓN]

RESULTS AND DISCUSSION

The results and discussion section of the present Doctoral Thesis is shown as a compilation of 11 scientific studies, which are classified within three sections:

Section 1: Effectiveness of innovative programs to promote students' healthy physical activity and sedentary behavior habits. (*Papers I-VII*)

Section 2: Validity of instruments for assessing factors influencing students' physical activity habits. (*Papers VIII-IX*)

Section 3: Validity of activity trackers for estimating students' daily physical activity and sedentary behavior levels. (*Papers X-XI*)

The full scientific studies are enclosed below in the language in which they were published or submitted.

RESULTADOS Y DISCUSIÓN

El apartado de resultados y discusión de la presente Tesis Doctoral se muestra como una compilación de 11 publicaciones, que se dividen en tres apartados:

Sección 1: Eficacia de programas innovadores para promover hábitos saludables de AF y CS en los estudiantes. (*Artículos I-VII*)

Sección 2: Validez de instrumentos para evaluar factores que influyen en los hábitos de AF de los estudiantes. (*Artículos VIII-IX*)

Sección 3: Validez de los monitores de fitness para estimar los niveles diarios de AF y CS de los estudiantes. (*Artículos X-XI*)

A continuación se adjuntan los estudios científicos completos en el idioma en que fueron publicados o sometidos.

SECTION 1

EFFECTIVENESS OF INNOVATIVE PROGRAMS TO PROMOTE STUDENTS' HEALTHY PHYSICAL ACTIVITY AND SEDENTARY BEHAVIOR HABITS

PAPERS I - VII

Paper I

RICYDE
Revista Internacional de Ciencias del Deporte
2020
Volumen 16(60)
180-198

¿ES EFECTIVO EL MODELO DE EDUCACIÓN DEPORTIVA PARA MEJORAR LA DEPORTIVIDAD Y ACTIVIDAD FÍSICA EN ESCOLARES?

Casado-Robles, C., Mayorga-Vega, D., Guijarro-Romero, S., & Viciana, J.

Ricyde. Revista Internacional de Ciencias del Deporte

2020, *16*(60), 180-198. https://doi.org/10.5232/ricyde2020.06005

Paper I

ABSTRACT

Objective: The main purpose of the study was to assess the effect of a Sport Education teaching unit in Physical Education on sportsmanship and habitual physical activity levels in Secondary Education students.

Methods: A total of 114 students (62 boys and 52 girls) from four classes (two from the second degree and two from the fourth degree) of compulsory Secondary Education were randomly assigned by classes to the Traditional group (methodology based on direct instruction) or the Sport Education group (Sport Education model). Both groups participated in 12 sessions of Physical Education (two sessions per week).

Results: The Multilevel Linear Model results showed that the Sport Education program improved the participation dimension in the students' sportsmanship. However, there were no significant differences in the rest of the dimensions of sportsmanship nor on the students' habitual physical activity levels.

Conclusions: The Sport Education Model improved the students' willingness to participate in sports competitions, but it did not improve the rest of the dimensions of sportsmanship or the students' physical activity levels.

KEYWORDS

Students. Fair-play. Sporting values. Sport Education Model. Educational Innovation.

Introducción

La práctica regular de actividad físico-deportiva es considerada idónea para la educación en valores y el desarrollo de comportamientos pro-sociales (Ruiz, Ponce, Sanz, y Valdemoros, 2015). En concreto, uno de los valores fundamentales que se desarrolla con la práctica deportiva es la deportividad, la cual implica valores como: (a) mostrar respeto por las reglas de juego y con el adversario; (b) rechazar la victoria a cualquier precio, o (c) mostrar una actitud digna en la victoria y la derrota (Bernal y Daniel-Huerta, 2016). Por desgracia, la práctica deportiva se asocia en numerosas ocasiones a valores completamente opuestos, especialmente si los responsables de organizar la práctica deportiva en jóvenes se centran en desarrollar modelos competitivos y orientados a la exaltación del resultado por encima de los valores educativos, llegando a generar comportamientos antideportivos que desvían el proceso de aprendizaje (Viciana, Mayorga-Vega, Ruiz, y Blanco, 2016).

Además del desarrollo de comportamientos pro-sociales, la práctica regular de actividad física (AF) supone numerosos beneficios para la salud y el bienestar de los jóvenes (Poitras y col., 2016). Cabe destacar el desarrollo de un sistema cardiovascular sano, la disminución de marcadores de riesgo cardiometabólico o la reducción de depresión y ansiedad (World Health Organization, 2014). Independientemente de los niveles de AF, la conducta sedentaria (CS) también está considerada como un potente indicador de salud cardiovascular en los jóvenes (Carson y col., 2016). Por ejemplo, una mayor cantidad de CS está relacionada con peores niveles de condición física, marcadores desfavorables de riesgo cardiovascular y peor composición corporal (Carson y col., 2016). Por ello, se recomienda que los jóvenes acumulen al menos 60 minutos diarios de AF moderada-vigorosa (World Health Organization, 2014) y eviten estar sentados durante periodos prolongados, limitando el tiempo de uso de pantalla a no más de dos horas diarias (Tremblay y col., 2016). Sin embargo, estudios previos han demostrado que la práctica de AF disminuye en la transición de la infancia a la adolescencia (Nader, Bradley, Houts, McRitchie, y O'Brien, 2008), así como posteriormente de la adolescencia a la edad adulta (Corder y col., 2017). De manera inversa, el comportamiento sedentario aumenta progresivamente durante esas etapas (Pearson, Haycraft, Johnston, y Atkin, 2017). Por ejemplo, mundialmente, alrededor del 80% de los jóvenes no cumplen con las recomendaciones diarias de AF (World Health Organization, 2014).

Desde el ámbito educativo, especialmente en la asignatura de Educación Física (EF), encontramos al profesor como un agente socializador fundamental para la creación de situaciones de aprendizaje centradas en la transmisión de valores éticos y morales, así como, para fomentar la práctica de AF y la adquisición de hábitos de vida saludables (Menéndez-Ferreira, Barguin, Maldonado, y Camacho, 2018). Además, esto se encuentra requerido desde el ámbito normativo en España. Por ejemplo, en el área de EF de la etapa de Educación Secundaria se destaca como objetivo de la materia ayudar a la adquisición y consolidación de hábitos responsables de AF regular, así como perseguir la mejora de la tolerancia y deportividad de los estudiantes a través de un estándar de aprendizaje a alcanzar (Ministerio de Educación, Cultura y Deporte, 2015).

Un modelo de enseñanza apropiado para promover el desarrollo de la deportividad, así como para fomentar la práctica de AF entre los adolescentes, podría ser el Modelo de Educación Deportiva (*Sport Education Model*) (MED) (Siedentop, Hastie, y Van der Mars, 2011). El MED surgió con el objetivo de proporcionar experiencias deportivas positivas y de calidad dentro de las clases de EF, trasladando la competición al contexto educativo y aprovechando así todos los valores educativos del deporte (Siedentop et al., 2011). Este modelo consiste en la integración de la dinámica de un club deportivo y sus diferentes roles en el aula de EF para intentar motivar e involucrar al alumnado en el proceso de enseñanza-aprendizaje (Siedentop y col., 2011). El MED está basado en seis rasgos característicos del deporte: (a) Los estudiantes desarrollan el "sentido de afiliación" al sentirse miembros de un equipo durante toda la temporada y asumir diferentes roles (por ejemplo, entrenador, preparador físico o árbitro); la "temporada" (es decir, la unidad didáctica) es más larga que una tradicional e incluye tiempo de práctica y partidos; (c) se celebra un "evento final" donde se desarrolla un torneo festivo y la entrega de premios a diferentes conductas y resultados obtenidos; (d) la "fase de competiciones" con situaciones reducidas de juego se celebra de acuerdo al calendario de competiciones e intercalada con periodos de práctica autónoma; (e) el "registro de datos" relacionado con el comportamiento de los alumnos y resultados de los partidos tiene lugar durante las competiciones para dar *feedback* a los alumnos y ayudarlos en la mejora de ese deporte; y (f) la naturaleza "festiva" del deporte está presente durante toda la temporada, utilizando elementos distintivos del equipo, publicitando fotografías y celebrando éxitos (Siedentop et al., 2011). Todas estas características se manifiestan a lo largo del modelo, el cual está dividido en tres fases según la metodología empleada en cada una de ellas: "pretemporada", la cual incluye una subfase introductoria al modelo, una subfase dirigida por el profesor de EF y la subfase de pretemporada dirigida por los alumnos-entrenadores; "temporada", destinada a la celebración de partidos; y "evento final" donde se realiza la clausura de la temporada con ambiente festivo (Siedentop et al., 2011). Son numerosas las intervenciones realizadas en EF utilizando esta metodología que han demostrado su potencial para la consecución de múltiples beneficios psicológicos, de aprendizaje técnico-táctico, rendimiento deportivo, disfrute de la AF o motivación intrínseca (Chu y Zang, 2018). En relación con la deportividad, estudios previos han obtenido mejoras en el desarrollo del juego limpio, imparcialidad, respeto a todos los participantes y participación activa (por ejemplo, Méndez-Giménez, Fernández-Río, y Méndez-Alonso, 2015; Vidoni y Ward 2009). Por otro lado, en relación con este modelo y el tiempo de práctica de AF, se han obtenido mejoras en la intención de ser físicamente activo durante el tiempo libre (Martínez de Ojeda, Méndez-Giménez y Pérez, 2016).

Sin embargo, los estudios previos presentan importantes limitaciones metodológicas que potencialmente podrían sesgar sus conclusiones: (a) la aplicación de un diseño no controlado (Wahl-Alexander y col., 2016); (b) el uso de instrumentos de medida como la observación sistemática sin cegar a los evaluadores (por ejemplo, Perlman y Goc Karp, 2010), o cuestionarios con una validez cuestionada (por ejemplo, la Escala Multidimensional de Orientaciones a la Deportividad o AF autoreportada) (Lamoneda, Huertas, Córdoba, y García, 2014; Wallhead, Garn, y Vidoni, 2014), y (c) el análisis estadístico ignorando que los estudiantes están anidados en clases (Méndez-Giménez y col., 2015). Por tanto, es conveniente llevar a cabo estudios con diseños cuasi-experimentales, controlados y aleatorizados por grupos naturales (Campbell, Piaggio, Elbourne, y Altman, 2012) y con un análisis estadístico realizado con un Modelo Lineal Multinivel, que considere que los estudiantes están anidados en clases (Li, Xiang, Chen, y Xie, 2017). En cuanto a las medidas, el reciente Cuestionario Multidimensional de Deportividad (CMD) es considerado una herramienta perfectamente válida para evaluar la deportividad (Iturbide-Luquin y Elosua-Oliden, 2017). El empleo de la medición objetiva de la AF habitual con acelerometría también aportaría al conocimiento científico un paso adelante respecto a las investigaciones anteriores (Baumgartner, Jackson, Mahar, y Rowe, 2015). Además, dado que la intención de ser físicamente activo es considerada el antecedente más importante para predecir el comportamiento de realizar AF (Ajzen, 1991), su estudio también parece necesario. Todos los aspectos anteriormente mencionados aportarían al conocimiento científico de esta línea de investigación un paso adelante respecto a las investigaciones anteriores realizadas. Consecuentemente, el objetivo principal del presente estudio fue examinar el efecto de una unidad didáctica del MED en EF sobre la deportividad y los niveles objetivos de AF habitual comparado con una metodología tradicional de los deportes en estudiantes de Educación Secundaria. El objetivo secundario fue evaluar el efecto del programa sobre los niveles de AF habitual percibidos y la intención de ser físicamente activos en estudiantes de Educación Secundaria.

Método

Diseño del estudio

El presente estudio está reportado de acuerdo a las directrices de CONSORT para estudios aleatorizados por grupo (Campbell y col., 2012). El protocolo seguido respeta la Declaración de Helsinki (64ª AMM, Brasil, octubre de 2013) y fue aprobado por el Comité de Ética en Investigación Humana de la Universidad de Granada.

Por razones prácticas y debido a la naturaleza del estudio (grupos naturales en el ámbito educativo), se empleó un diseño cuasi-experimental, controlado y aleatorizado por grupos naturales (*cluster-randomized controlled trial*) (Viciana, Casado-Robles, Pérez-Macías, y Mayorga-Vega ,2020). Las cuatro clases pre-establecidas por el centro educativo fueron distribuidas de forma aleatoria, balanceadas por curso, al grupo Tradicional y al grupo Educación Deportiva (ED) (una clase de segundo y una de cuarto para cada grupo).

Participantes

En primer lugar, se obtuvo la aprobación del centro. Posteriormente, se solicitó por escrito la firma del consentimiento informado por los padres, madres o tutores legales y el asentimiento informado de los estudiantes para participar en el estudio. La muestra inicial estaba compuesta por 114 estudiantes (62 mujeres y 52 varones) con edades comprendidas entre 13 y 16 años (media = 14,0 ± 1,1 años).

Todos los estudiantes pertenecían a cuatro clases diferentes de segundo y cuarto curso (dos clases por curso) de Educación Secundaria Obligatoria de un centro público de la provincia de Granada. Además, los estudiantes no tenían experiencia previa en el MED ni en ninguna otra experiencia innovadora desarrollada en el contexto educativo. El centro educativo estaba situado en el área urbana de la ciudad de Granada.

Los criterios de inclusión fueron: (a) estar matriculado en el segundo o cuarto curso de Educación Secundaria Obligatoria del centro educativo seleccionado (clases en las que se obtuvo permiso del centro escolar); (b) no padecer ninguna enfermedad o lesión que les impidieran realizar EF con normalidad; (c) presentar el consentimiento informado firmado por sus madres, padres o tutores legales, y (d) presentar el asentimiento informado firmado por los estudiantes. Los criterios de exclusión fueron: (a) no tener una asistencia al programa de intervención de al menos 10 sesiones, y (b) no tener un registro válido de ninguna medida.

Medidas

La evaluación se llevó a cabo durante las clases de EF al comienzo y al final del programa de intervención (pre-intervención y post-intervención, respectivamente). Cada evaluación fue realizada por los mismos evaluadores (primer y tercer autor), instrumentos y condiciones. Además, antes de comenzar el estudio se recogieron los datos sociodemográficos (es decir, edad, género y curso) de los informes del centro educativo y las medidas antropométricas. A continuación, se detalla el procedimiento de medida seguido con cada variable.

Antropometría

Las medidas antropométricas de los adolescentes fueron registradas siguiendo los Estándares Internacionales para la Evaluación Antropométrica (Stewart, Marfell-Jones, Olds, y De Ridder, 2011). En primer lugar, se midieron la masa corporal (Seca, Ltd., Hamburg, Alemania; precisión = 0,1 kg) y la talla de los participantes (Holtain Ltd., Crymmych, Pembs, Reino Unido; precisión = 0,1 cm) en pantalones cortos, camiseta y pies descalzos. Se realizaron dos mediciones de la masa corporal y la talla para posteriormente calcular el promedio de cada una de ellas. Posteriormente, se calculó el índice de masa corporal como la masa corporal (en kilogramos) dividida por talla al cuadrado (en metros).

Deportividad

La deportividad del alumnado se evaluó mediante el cuestionario CMD de Iturbide-Luquin y Elosua-Oliden (2017). Este cuestionario evalúa las orientaciones del alumnado hacia las diferentes categorías subyacentes del constructo de la deportividad como ideal de conducta ética. Este cuestionario posee un total de 21 ítems divididos en cinco dimensiones: (a) disfrute (por ejemplo, "Practico deporte porque me gusta"); (b) respeto (por ejemplo, "Muestro respeto hacia los árbitros"); (c) compromiso (por ejemplo, "Busco remedios para mejorar mis puntos débiles"); (d) juego limpio (por ejemplo, "Cuando el otro equipo juega duro, yo también lo hago"), y (e) participación (por ejemplo, "No me importa el resultado de la competición"). Todas las dimensiones poseen cuatro ítems, excepto la dimensión de disfrute que consta de cinco ítems. Los ítems estaban precedidos por la frase introductoria: "Respecto a tu comportamiento en el deporte…". La escala se adaptó a las calificaciones realizadas con estudiantes españoles con una escala tipo Likert que iba de 0 ("Totalmente en desacuerdo") a 10 ("Totalmente de acuerdo") (Viciana, Casado-Robles, Pérez-Macías, y Mayorga-Vega ,2020).

Actividad física habitual percibida

La AF habitual percibida de los participantes se evaluó mediante el cuestionario PACE (*Physician-based Assesment and Counseling for Exercise*), en su versión adaptada y validada al contexto español para adolescentes por Martínez-Gómez y col. (2009). Este cuestionario valora con dos preguntas cuántos días en la última semana y en una semana habitual el adolescente realiza al menos 60 min de AF (por ejemplo, "En los últimos siete días, ¿cuántos días hiciste AF 60 minutos o más?"). Las dos preguntas estaban precedidas por la frase introductoria: "A continuación te preguntamos sobre tus hábitos de AF…". Los estudiantes debían responder en una escala que iba desde 0 hasta 7 días. La versión española del cuestionario en adolescentes ha demostrado una correlación estadísticamente significativa con la AF habitual medida con acelerometría (r = 0,43) y una adecuada fiabilidad test-retest (CCI = 0,77) (Martínez-Gómez y col., 2009).

Intención de ser físicamente activo

La intención de los participantes de ser físicamente activos se evaluó mediante el cuestionario Medida de la Intencionalidad de ser Físicamente Activo (MIFA), en su versión adaptada y validada al contexto español por Moreno, Moreno, y Cervelló (2007). Este cuestionario está compuesto por una única dimensión de cinco ítems (por ejemplo, “Al margen de las clases de EF, me gusta practicar deporte”). Dichos ítems iban precedidos por la frase “Respecto a tu intención de practicar alguna actividad físico-deportiva...”. De acuerdo con estudios previos (por ejemplo, Viciana, Casado-Robles, Pérez-Macías, y Mayorga-Vega, 2020), para adaptar la escala a las calificaciones realizadas con estudiantes españoles, se utilizó una escala tipo Likert que iba de 0 (“Totalmente en desacuerdo”) a 10 (“Totalmente de acuerdo”). La versión española del cuestionario ha demostrado unas propiedades psicométricas adecuadas en adolescentes (CFI = 0,98; RMSEA = 0,056; α de Cronbach = 0,94) (Moreno y col., 2007).

Niveles objetivos de actividad física y conducta sedentaria

Los niveles de AF habitual y CS se evaluaron objetivamente mediante un acelerómetro GT3X (ActiGraph, LLC, Pensacola, FL, EE.UU.). El acelerómetro fue colocado en la cadera derecha de los estudiantes mediante un cinturón elástico. Se solicitó a los estudiantes que lo llevaran durante siete días consecutivos y para evitar sesgos de reactividad de los participantes, el primer día de registro se consideró como un día de familiarización y no se utilizó para los análisis estadísticos (es decir, seis días completos fueron evaluados) (Dössegger y col., 2014). Debido a que los patrones de comportamiento de los adolescentes están caracterizados por cortos periodos de actividad muy cambiante, se utilizó un *epoch* de un segundo (Cain, Sallis, Conway, Van Dyck, y Calhoon, 2013). El tiempo mínimo de uso del dispositivo para considerarlo como un día válido fue establecido en 600 min por día (Migueles y col., 2017). Los períodos de no uso del dispositivo se establecieron con una duración mínima de 60 min con un *epoch* de cero *counts* consecutivos con una tolerancia pico de hasta dos minutos (Oliver, Badland, Schofield, y Shepherd, 2011). Se estableció un criterio mínimo de dos días de semana y un día de fin de semana con tiempo válido (Autor, 2018b).

Para calcular el tiempo diario en CS, AF ligera y AF moderada-vigorosa se utilizaron los puntos de corte de Evenson (Evenson, Catellier, Gill, Ondrak, y McMurray, 2008): (a) 0-100 *counts*/ min para CS; (b) 101-2295 *counts*/ min para AF ligera; (c) ≥ 2295 *counts*/ min para AF moderada-vigorosa. La AF total fue calculada como el porcentaje de tiempo diario involucrado en AF de intensidad ligera a vigorosa (es decir, ≥ 101 *counts*/ min) y promedio del eje vertical diario (es decir, promedio *counts*/ min). De acuerdo al estudio de validación cruzada realizado por Trost, Loprinzi, Moore, y Pfeiffer (2011), estos puntos de corte han demostrado ser los más válidos para evaluar la AF en adolescentes con un *epoch* de un segundo. Para evitar sesgos potenciales en las diferentes duraciones de tiempo válido, todas las variables fueron expresadas en términos relativos al tiempo de registro de uso válido (Orme, Wijndaele, Sharp, Westgate, Ekelund, y Brage, 2014). Finalmente, los niveles de AF habitual y CS de los estudiantes fue calculada de la siguiente manera: ((5 × valor medio de los días de semana válidos) + (2 × valor medio de los días de fin de semana válidos))/ 7 (Parra-Saldías, Mayorga-Vega, López-Fernández, y Viciana, 2018). La iniciación, descarga y análisis de los datos se realizaron mediante el programa *ActiLife Lifestyle Monitoring System Software*, versión 6.13.3. Los valores de los acelerómetros ActiGraph han demostrado una alta validez para evaluar la AF y CS de los escolares (Trost y col., 2011).

Procedimiento

Los dos profesores de EF que llevaron a cabo la intervención (primer y tercer autor) tenían una experiencia de 2 años aplicando el MED en las clases de EF. En concreto, habían desarrollado previamente tres temporadas del MED para el aprendizaje del voleibol, *floorball* y fútbol sala con estudiantes de Educación Secundaria Obligatoria, con una duración de 12, 13 y 16 sesiones de EF, respectivamente. A pesar de contar con experiencia previa, los profesores de EF encargados de impartir las sesiones también recibieron tres sesiones formativas sobre las características específicas de los programas de intervención para garantizar la correcta implementación del modelo y diseñar exhaustivamente el plan semanal para toda la temporada junto con el equipo de investigación. Las sesiones fueron impartidas por los dos profesores de EF de forma balanceada. Es decir, un profesor impartió las sesiones del MED a un grupo de segundo curso y las sesiones de la intervención Tradicional a otro grupo de segundo curso, mientras que, el otro profesor impartió las sesiones del MED a un grupo de cuarto curso y las sesiones de la intervención Tradicional al otro grupo de cuarto curso.

Los estudiantes del grupo Tradicional realizaron un programa basado en la instrucción directa (Metzler, 2011) y los del grupo ED basado en el MED (Siedentop y col., 2011). Los programas de intervención fueron diseñados y analizados cuidadosamente de acuerdo a las características diferenciadoras de cada metodología de enseñanza. Además, un investigador externo con experiencia en la enseñanza de la EF observó la correcta aplicación de los programas de intervención durante las sesiones de EF, utilizando los puntos de referencia para la instrucción directa (Metlzer, 2011) y el MED siguiendo la lista de 19 ítems propuestos por Ko, Wallhead, y Ward (2006) para comprobar la correcta implementación del modelo de acuerdo a cinco categorías estructurales (afiliación, temporada, responsabilidad, competición formal y registro de datos). El investigador externo registró una fidelidad del 100% con el MED en las categorías de afiliación, temporada, competición formal y registro de datos durante las clases evaluadas (una clase de cada fase). En la categoría de responsabilidad, se obtuvo un 87,5% ya que el ítem "El maestro utiliza tareas para entrenar a los estudiantes en una comunicación verbal efectiva y *feedback*" no se cumplió correctamente. Los estudiantes de ambos grupos realizaron una unidad didáctica para la enseñanza del baloncesto (segundo curso) y el fútbol sala (cuarto curso) con una duración total de 12 sesiones de EF (de 60 minutos de duración programada; 50 minutos de tiempo útil) y una frecuencia de dos sesiones semanales.

En la Figura 1 podemos observar la intervención llevada a cabo por cada uno de los grupos. El grupo ED recibió la unidad didáctica siguiendo el MED. La intervención estaba diseñada de acuerdo al modelo original (Siedentop y col., 2011), diferenciando cinco fases en función de la metodología aplicada en cada una de ellas: (a) una sesión de fase introductoria, donde se explicó el modelo de enseñanza a aplicar y sus características principales (por ejemplo, fases del modelo o roles del alumnado) y se crearon los equipos de forma homogénea según su nivel inicial en el deporte. Cada clase estaba dividida en siete equipos y cada equipo estaba formado por cuatro alumnos; (b) dos sesiones de fase dirigida, donde el profesor propuso diferentes juegos y tareas para la enseñanza de los elementos técnico-tácticos y reglamentarios del deporte siguiendo el estilo de enseñanza de asignación de tareas; (c) dos sesiones de fase de pretemporada, donde los alumnos ya asumieron sus roles (es decir, entrenador, preparador físico, asistente de material o capitán) dentro del equipo y el alumno-entrenador dirigió la clase a sus compañeros previa planificación y consenso con el profesor; (d) seis sesiones de fase de temporada, donde se celebraron partidos de acuerdo al calendario de competición y un equipo realizó las funciones de árbitro y cronista (observando estadísticas y eventos destacados de los partidos), y (e) una sesión de evento final, donde se realizó la clausura de la temporada con la celebración de partidos amistosos y la entrega de premios elaborados por los estudiantes en un ambiente festivo.

El grupo Tradicional recibió la unidad didáctica siguiendo una metodología basada en la técnica de enseñanza de la instrucción directa, cuyas características principales son: (a) el rol del profesor es ser el protagonista del proceso de enseñanza-aprendizaje emitiendo un modelo de probado rendimiento y bien definido; (b) el rol del alumno es recibir y reproducir ese modelo con la mayor precisión posible y ajustando la respuesta al modelo; (c) las decisiones las toma el profesor (objetivos, contenidos, actividades y evaluación), quedando el papel del alumno en la toma de decisiones más limitada, y (d) la organización es formal y poco individualizada. Todas las sesiones estaban orientadas a la realización de juegos y tareas orientadas al aprendizaje y mejora técnico-táctica de los deportes con la siguiente estructura: un calentamiento de 10 minutos con ejercicios aeróbicos y juegos; 40 minutos de parte principal con tres-cinco tareas técnico-tácticas y finalizando con situaciones reducidas de juego 3x3 de cinco minutos de duración; y una vuelta a la calma de cinco minutos con ejercicios de flexibilidad comunes.

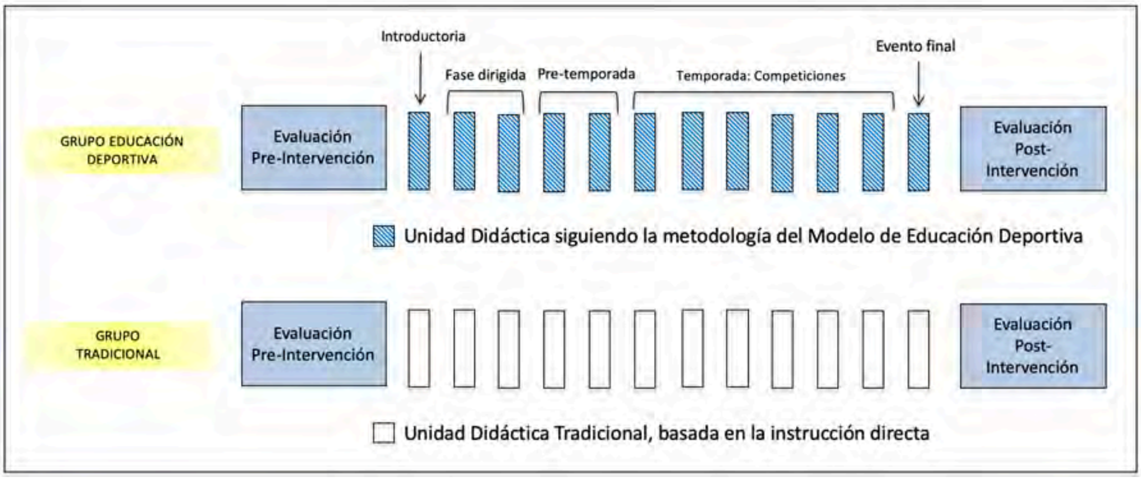


Figura 1. Programas de intervención. Unidad Didáctica Modelo de Educación Deportiva (grupo Educación Deportiva) y Unidad Didáctica Tradicional (grupo Tradicional).

Tamaño de muestra

El cálculo del tamaño de la muestra *a priori* se estimó con el programa *Optimal Design Plus Empirical Evidence* versión 3.01 para Windows. Los parámetros se establecieron de la siguiente manera: nivel de significación α = 0,05; número de participantes por grupo n = 25; tamaño del efecto δ = 0,50; coeficiente de correlación intraclase ρ = 0,01 y potencia estadística (1 - β) = 0,80. Se estimó un tamaño de muestra final de aproximadamente cuatro grupos (100 participantes).

Aleatorización

Antes de realizar la evaluación inicial de la intervención, un investigador externo y ajeno a los objetivos del estudio, realizó la asignación de las clases del centro educativo a los diferentes grupos de estudio. La asignación al azar se realizó en el nivel clase, utilizando un generador de números aleatorios computarizado. Sin embargo, a principio de curso los estudiantes ya habían sido previamente asignados de forma aleatoria, balanceado por género, a cada clase por el centro educativo.

Análisis estadístico

Se calcularon los estadísticos descriptivos de las características generales de los participantes y variables dependientes. Los supuestos de las pruebas estadísticas se verificaron mediante procedimientos comunes (por ejemplo, histogramas y diagramas Q-Q para la normalidad). Los análisis de varianza de un factor (ANOVA) (variables continuas) y la prueba de chi-cuadrado (variables categóricas) se llevaron a cabo para examinar las posibles diferencias en las características generales entre los dos grupos. Los coeficientes de correlación intraclase (CCI, también conocidos como coeficientes de correlación intragrupo) mostraron evidencia empírica de la independencia de las observaciones (CCI, media = 0,01; mediana = 0,00; mínimo = 0,00; máximo = 0,03; Wald *Z*, p > 0,05). Sin embargo, debido a que la unidad de intervención era la clase, el efecto del programa de intervención se estudió mediante un Modelo Lineal Multinivel con los participantes anidados dentro de las clases (Li y col., 2017). Según la recomendación de Field (2017), el enfoque comenzó a partir de modelos "básicos" en los que se fijaron todos los parámetros y luego se siguieron progresivamente los coeficientes aleatorios y la exploración de variables de confusión. Se utilizó el -2 log-verosimilitud (*-2 log-likelihood*) para comparar el ajuste de los modelos (es decir, comparar el cambio en la prueba de chi-cuadrado). De todas las posibles variables de confusión exploradas (es decir, género, edad, curso, masa corporal, altura, índice de masa corporal y asistencia a la intervención), no fue necesario utilizar ninguna covariable. Se utilizó el método de estimación de máxima verosimilitud. Los tamaños del efecto se estimaron mediante la ∂ de Cohen. Todos los análisis estadísticos se realizaron mediante el paquete estadístico SPSS versión 25.0 para Windows (IBM® SPSS® Statistics 25). El nivel de significación estadística se estableció en $p < 0{,}05$.

Resultados

La Figura 2 muestra el diagrama de flujo correspondiente a los participantes incluidos en el presente estudio. Aunque los 114 estudiantes invitados estuvieron de acuerdo en participar y cumplieron con los criterios de inclusión, únicamente 99 participantes superaron los criterios de exclusión y finalizaron el estudio. Sin embargo, para la variable de niveles objetivos de AF y CS, la muestra final fue de 45 participantes.

El reclutamiento se realizó en Diciembre de 2016, y la toma de datos se realizó de enero a marzo de 2017. La Tabla 1 muestra las características generales de los participantes analizados en el presente estudio. Los resultados de la prueba chi cuadrado mostraron que los grupos Tradicional y ED estaban balanceados en términos de grupo de edad, curso y género ($p > 0{,}05$). Además, los resultados del ANOVA de un factor no mostraron diferencias estadísticamente significativas entre los grupos estudiados en términos de masa corporal, talla e índice de masa corporal ($p > 0{,}05$). Para la muestra utilizada en el estudio de las variables niveles objetivos de AF y CS, las características generales de los participantes de ambos grupos también fueron similares. Los participantes del grupo ED obtuvieron una asistencia media del 97,96% y los participantes del grupo Tradicional del 88,46%.

Tabla 1. Características generales de los estudiantes analizados[a]

	Total (n = 99)	Tradicional (n = 44)	Educación Deportiva (n = 55)	χ^2/ F	p[b]
Edad (13-14/ 15-16 años)	55/ 44	26/ 18	29/ 26	0,401	0,527
Curso (2º/ 4º)	55/ 44	26/ 18	29/ 26	0,401	0,527
Género (mujeres/ varones)	45/ 54	17/ 27	28/ 27	1,485	0,223
	Media (DE)	Media (DE)	Media (DE)		
Masa corporal (kg)	57,01 (9,14)	56,11 (9,57)	57,73 (8,80)	0,768	0,383
Talla (cm)	163,22 (8,71)	162,47 (9,85)	163,82 (7,71)	0,590	0,444
Índice de masa corporal (kg/m^2)	21,37 (2,86)	21,23 (2,98)	21,49 (2,78)	0,191	0,663

Nota. DE = Desviación estándar; [a] Los datos están reportados como frecuencia o media (desviación estándar). [b] Nivel de significación para la prueba chi cuadrado (variables categóricas) y el ANOVA de un factor (variables continuas).

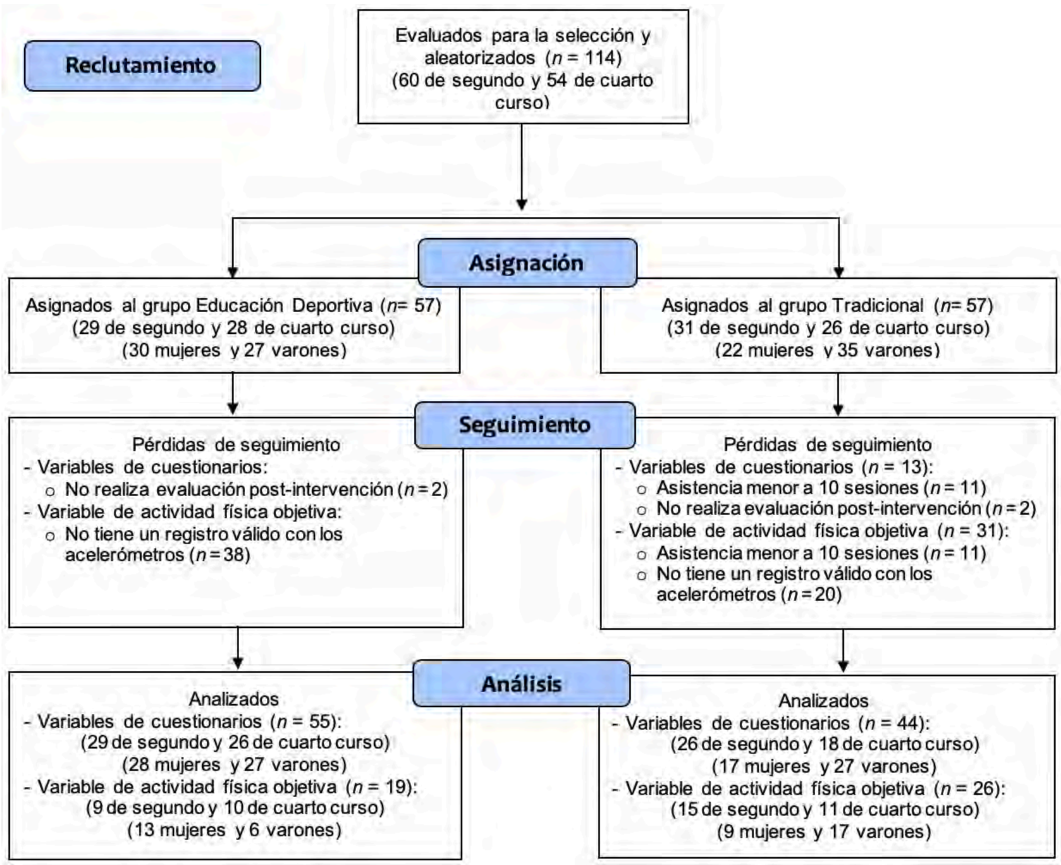


Figura 2. Diagrama de flujo correspondiente a los participantes incluidos en el estudio

Tabla 2. Efecto del programa de Educación Deportiva sobre la deportividad

		Pre-intervención	Post-intervención	Diferencia	Model
		Media (DE)	Media (DE)	Media (DE)	- 2LL
Deportividad					
Participación	Tradicional (*n* = 44)	7,30 (2,40)	6,91 (2,59)	- 0,39 (1,26)	403,92
	Educación Deportiva (*n* = 55)	6,32 (2,18)	7,03 (2,11)	0,71 (2,25)	
Disfrute	Tradicional (*n* = 44)	8,22 (2,30)	8,26 (2,35)	0,04 (1,26)	337,49
	Educación Deportiva (*n* = 55)	8,46 (1,74)	8,12 (1,76)	- 0,33 (1,41)	
Juego limpio	Tradicional (*n* = 44)	5,43 (2,09)	4,84 (2,24)	- 0,59 (2,04)	411,72
	Educación Deportiva (*n* = 55)	5,74 (1,72)	5,76 (2,16)	0,03 (1,89)	
Respeto	Tradicional (*n* = 44)	8,18 (2,00)	8,40 (1,66)	0,22 (1,13)	340,34
	Educación Deportiva (*n* = 55)	8,00 (1,45)	8,24 (1,63)	0,24 (1,52)	
Compromiso	Tradicional (*n* = 44)	8,12 (1,64)	8,34 (1,63)	0,22 (1,20)	348,82
	Educación Deportiva (*n* = 55)	8,26 (1,44)	8,14 (1,86)	- 0,13 (1,58)	

Nota. DE = Desviación estándar; - 2LL = -2 log-verosimilitud (*-2 log-likelihood*). [a] Tamaño del efecto con la *d* de

Tabla 3. Efecto del programa de Educación Deportiva sobre los niveles objetivos y percibidos de actividad física habitual e

		Pre-intervención	Post-intervención	Diferencia
		Media (DE)	Media (DE)	Media (DE)
Actividad física habitual/ conducta sedentaria objetiva				
Conducta sedentaria (%)	Tradicional (*n* = 26)	85,76 (4,18)	85,52 (4,43)	-0,24 (2,76)
	Educación Deportiva (*n* = 19)	86,84 (4,16)	86,32 (4,85)	-0,53 (4,54)
Ligera (%)	Tradicional (*n* = 26)	8,45 (2,27)	8,78 (2,60)	0,33 (1,87)
	Educación Deportiva (*n* = 19)	7,57 (2,43)	7,89 (2,97)	0,32 (2,20)
Moderada-vigorosa (%)	Tradicional (*n* = 26)	5,79 (2,42)	5,70 (2,53)	-0,09 (1,94)
	Educación Deportiva (*n* = 19)	5,59 (2,42)	5,80 (2,88)	0,21 (3,01)
Total (Ligera-vigorosa %)	Tradicional (*n* = 26)	14,24 (4,18)	14,47 (4,43)	0,24 (2,76)
	Educación Deportiva (*n* = 19)	13,16 (4,16)	13,68 (4,85)	0,53 (4,54)
Total (Eje vertical counts/ min)	Tradicional (*n* = 26)	340,95 (125,36)	343,08 (138,91)	2,13 (78,93)
	Educación Deportiva (*n* = 19)	336,40 (150,71)	370,85 (181,60)	34,45 (164,87)
Actividad física habitual percibida	Tradicional (*n* = 44)	3,16 (1,62)	3,17 (1,78)	0,01 (1,42)
	Educación Deportiva (*n* = 55)	2,92 (1,51)	3,48 (1,62)	0,56 (1,61)
Intención de ser físicamente activo	Tradicional (*n* = 44)	7,55 (2,23)	7,69 (2,36)	0,14 (1,32)
	Educación Deportiva (*n* = 55)	7,15 (1,99)	7,23 (2,37)	0,08 (1,55)

Nota. DE = Desviación estándar; - 2LL = -2 log-verosimilitud *(-2 log-likelihood*. [a] Tamaño del efecto con la *d* de Cohen.

Discusión

El objetivo principal del presente estudio fue examinar el efecto de una unidad didáctica del MED en EF sobre la deportividad y los niveles objetivos de AF en estudiantes de Educación Secundaria. El objetivo secundario fue evaluar el efecto del programa sobre los niveles de AF habitual percibida y la intención de ser físicamente activos en estudiantes de Educación Secundaria. De acuerdo con los resultados del presente estudio, un programa en EF de corta duración siguiendo el MED produjo mejoras en la dimensión participación de la deportividad del grupo ED en comparación con el modelo tradicional de iniciación deportiva. Sin embargo, no hubo diferencias significativas en el resto de las dimensiones de deportividad evaluadas ni tampoco sobre los niveles objetivos ni subjetivos de AF habitual, ni sobre la intención de ser físicamente activo.

Respecto a la mejora de la dimensión de deportividad mediante la aplicación del MED, también ha sido observada en otros estudios previos realizados durante las clases de EF con estudiantes de Educación Secundaria (por ejemplo, Méndez-Giménez y col., 2015; Perlman y Goc Karp, 2010). Sin embargo, son numerosos los factores que debemos tener en cuenta para comparar los resultados obtenidos en el presente estudio con la literatura previa. Por un lado, Méndez-Giménez y col. (2015) estudiaron el efecto del MED sobre los niveles de deportividad para la enseñanza del ultimate durante 12 sesiones. Aunque la deportividad fue medida de forma cuantitativa y obtuvieron mejoras significativas respecto al grupo tradicional como en el presente estudio, el tamaño del efecto fue muy pequeño ($\partial = 0,05$ para la dimensión de reglas y árbitros; $\partial = 0,06$ para la dimensión de convenciones sociales y $\partial = 0,10$ para la dimensión de adversarios) en comparación con la mejora obtenida en el presente estudio ($\partial = 0,48$). Además, dado que la intervención fue administrada a clases completas y no a estudiantes aislados, la unidad de análisis estadístico debería haber sido la clase, por lo que los resultados podrían ser erróneos (Li y col., 2017). Por otro lado, el cuestionario empleado fue la Escala Multidimensional de Orientaciones a la Deportividad de Vallerand y col. (1997) cuyo uso está ampliamente cuestionado en la literatura científica debido a su baja consistencia interna y deficiente estructura factorial (Lamoneda y col., 2014). El uso de instrumentos que no sean completamente válidos y fiables y que, por lo tanto, no aseguren que el instrumento utilizado está midiendo el constructo deseado de forma correcta, podría dar lugar a conclusiones erróneas (Baumgartner y col., 2015).

Por otro lado, los estudios de Vidoni y Ward (2009), Perlman y Goc Karp (2010) y Wahl-Alexander y col. (2016), a diferencia del presente trabajo, se realizaron siguiendo una metodología cualitativa para evaluar la deportividad de los estudiantes durante las sesiones de EF. En primer lugar, Vidoni y Ward (2009) obtuvieron mediante la observación sistemática que el MED fue efectivo para disminuir los comportamientos violentos durante los partidos y mejorar la participación activa de los escolares. Sin embargo, el tamaño de muestra estudiado fue muy bajo (seis estudiantes) y la elección de los mismos estaba sesgada a únicamente aquellos estudiantes que, a juicio del profesor, presentaban baja incidencia de comportamientos de apoyo a los compañeros antes de la intervención. En segundo lugar, el estudio de Perlman y Goc Karp (2010) observaron a través de entrevistas que los estudiantes que seguían el MED actuaban con mayor deportividad (por ejemplo, no había protestas al árbitro y actuaban con calma ante situaciones que podían parecer injustas). Por último, Wahl-Alexander y col. (2016) observaron mediante entrevistas que los alumnos reportaban haber sido buenos compañeros y haber desarrollado una actitud de respeto a compañeros y rivales tras participar en varias temporadas del MED. No obstante, ambos estudios (Perlman y Goc Karp, 2010; Wahl-Alexander y col., 2016) utilizaron un diseño de grupo único, por lo que no existe un grupo control con el que controlar sesgos y poder afirmar que las mejoras son realmente debidas a la aplicación del MED. Además, los tres estudios descritos anteriormente también difieren en una mayor duración del programa, desde 18 a 103 sesiones, y en el contenido trabajado (por ejemplo, deportes alternativos como el rugby-tag, ultimate, speedball o lacrosse). El hecho de no haber registrado de forma cualitativa este tipo de conductas en el presente estudio junto con las diferencias descritas, hacen que los resultados no sean completamente comparables con el presente. Sin embargo, parece probable, que la modificación de conductas relacionadas con la deportividad sea más fácil de conseguir con una aplicación del MED durante un periodo de tiempo más prolongado.

Respecto a la influencia del MED sobre los niveles de AF de los estudiantes, estudios previos obtuvieron mejoras en los niveles de AF desarrollados durante las clases de EF del MED medidos con acelerometría (Perlman, 2012; Rocamora, González-Víllora, Fernández-Río y Arias-Palencia, 2019; Ward y col., 2017). Sin embargo, de lo que conocemos, no hay estudios previos que evalúen el efecto del MED sobre los niveles objetivos de AF habitual y CS de los escolares, la cual se considera una medida muy importante para comprobar el rol de la EF para la promoción de la práctica de AF de los escolares en contextos extraescolares. Desafortunadamente, el presente estudio no tuvo en efecto positivo en la AF habitual (objetiva ni percibida) practicada por los estudiantes del grupo ED frente a los estudiantes del grupo Tradicional. En línea con los resultados del presente estudio, Wallhead y col. (2014) tampoco obtuvieron mejoras en los niveles de AF habituales subjetivos, a pesar de tener una duración mucho mayor que el presente estudio (cuatro temporadas de 25 sesiones cada una). La ausencia de mejoras en los niveles de AF habitual podría ser debida a que los programas no estaban especialmente centrados en la transferencia de la práctica de AF de la clase de EF al contexto extraescolar y carecía de estrategias necesarias para ello (por ejemplo, información acerca de los beneficios de la práctica de AF o establecimiento de una meta diaria de AF). También puede deberse a que la intervención tiene una duración demasiado escasa para pretender el cambio de conducta en los jóvenes.

Por último, respecto a la intencionalidad de ser físicamente activo, en el presente estudio tampoco se obtuvieron diferencias significativas en los estudiantes del grupo ED en comparación con los estudiantes del grupo Tradicional. Estos resultados se encuentran en consonancia con los resultados obtenidos en otros estudios previos realizados con estudiantes de Educación Secundaria (Cuevas, García-López, y Serra-Olivares, 2016; Wallhead y col., 2014). Sin embargo, estudios previos con la aplicación de varias unidades didácticas consecutivas del MED en estudiantes de Educación Primaria ha encontrado altos cambios en la intencionalidad de práctica deportiva extraescolar (por ejemplo, Martínez de Ojeda y col., 2016). Parece probable que sea más sencillo conseguir el cambio de la intencionalidad de ser físicamente activo con una mayor duración del programa y con escolares más jóvenes.

Fortalezas y limitaciones

Respecto a las fortalezas, de lo que conocemos el presente estudio es el primero que examina el efecto del MED sobre los niveles de AF habituales de los adolescentes medidos objetivamente, lo que supone un paso adelante respecto a las investigaciones anteriores realizadas únicamente con medidas autorreportadas o evaluando los niveles de AF objetivos únicamente durante franjas concretas (por ejemplo, las clases de EF o los recreos escolares). Además, respecto a la medición de la deportividad, el presente estudio es el primero que demuestra un incremento en la participación con un programa de ED medido con el nuevo CMD de Iturbide-Luquin y Elosua-Oliden (2017). Por otro lado, el diseño del estudio (controlado y aleatorizado por grupos naturales) supone una ventaja frente a la mayoría realizados con un diseño de grupo único (por ejemplo, Perlman y Goc Karp, 2010). Por último, la evaluación del efecto del programa con un Modelo Lineal Multinivel con los participantes anidados dentro de las clases, supone un avance respecto a los análisis comúnmente aplicados (Li y col., 2017). En cuanto a las limitaciones, a pesar de que el propio creador del modelo estableciese como mínimo 12 sesiones de EF (Siedentop y col., 2011), es posible que la duración del programa (12 sesiones) haya sido un limitante para conseguir efectos en otras variables estudiadas, especialmente en los hábitos de AF. Respecto a los niveles objetivos de AF habitual, una limitación también podría haber sido el elevado número de participantes excluidos y, en consecuencia, el bajo tamaño de la muestra final (45 participantes). Además, el presente trabajo estudió el efecto de un programa de ED mediante el contenido de baloncesto y fútbol sala, y futuros estudios deberían examinar estos efectos con otros contenidos de enseñanza. Por último, cabe destacar que se trata de una muestra local perteneciente a un único centro educativo, por lo que los datos deben ser tomados con precaución y no generalizar.

Conclusiones

De lo que conocemos el presente estudio es el primero que examina el efecto de una unidad didáctica siguiendo el MED en EF sobre la deportividad y los niveles habituales objetivos de AF en estudiantes de Educación Secundaria con un estudio controlado por grupos y analizado mediante un Modelo Lineal Multinivel. Una unidad didáctica de ED de 12 sesiones aumentó el deseo y la voluntad del alumnado para participar en competiciones deportivas aplicando el máximo esfuerzo. Sin embargo, el programa no tiene un efecto mayor comparado con el modelo tradicional sobre el resto de dimensiones de la deportividad, ni sobre los niveles de AF habitual y la intención de ser físicamente activo. Por este motivo, el efecto del programa debería comprobarse en otros contextos (por ejemplo, con una mayor duración del programa y/o diferentes contenidos).

Agradecimientos

Agradecemos a todos los alumnos y las profesoras de Educación Física del centro escolar por su colaboración y participación en el estudio. Carolina Casado-Robles y Santiago Guijarro-Romero reciben una ayuda del Ministerio de Ciencia, Innovación y Universidades de España (FPU16/03314 y FPU15/02387, respectivamente). El estudio también está financiado por la "Beca de Iniciación a la Investigación para estudiantes de Máster" de la Universidad de Granada.

Referencias bibliográficas

Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes*, *50*(2), 179-211. https://doi.org/10.1016/0749-5978(91)90020-T

Baumgartner, T.; Jackson, A.; Mahar, M., y Rowe, D. (2015). *Measurement for evaluation in kinesiology* (9th ed.). Burlington, USA: Jones and Bartlett Learning.

Bernal, D., y Daniel-Huerta, M. (2016). Educación Física: una asignatura para mejorar el rendimiento académico, la cognición y los valores. *Revista Infancia, Educación y Aprendizaje*, *2*(1), 96-114. https://doi.org/10.22370/ieya.2016.2.1.586

Cain, K. L.; Sallis, J. F.; Conway, T. L.; Van Dyck, D., y Calhoon, L. (2013). Using accelerometers in youth physical activity studies: a review of methods. *Journal of Physical Activity and Health*, *10*(3), 437-450. https://doi.org/10.1123/jpah.10.3.437

Campbell, M. K.; Piaggio, G.; Elbourne, D. R., y Altman, D. G. (2012). Consort 2010 statement: extension to cluster randomised trials. *BMJ*, *345*, e5661. https://doi.org/10.1136/bmj.e5661

Carson, V.; Hunter, S.; Kuzik, N.; Gray, C. E.; Poitras, V. J.; Chaput, J.P.; … Tremblay, M. S. (2016). Systematic review of sedentary behaviour and health indicators in school-aged children and youth: An update. *Applied Physiology, Nutrition, and Metabolism,* *41*(6), 240–265. https://doi.org/10.1139/apnm-2015-0630

Chu, T. L., y Zhang, T. (2018). Motivational processes in Sport Education programs among high school students. *European Physical Education Review, 24*(3), 372–394. https://doi.org/10.1177/1356336X17751231

Corder, K.; Winpenny, E.; Love, R.; Brown, H. E.; White, M., y Sluijs, E. V. (2017). Change in physical activity from adolescence to early adulthood: a systematic review and meta-analysis of longitudinal cohort studies. *British Journal of Sports Medicine, 53*(8), 496-503. https://doi.org/10.1136/bjsports-2016-097330

Cuevas, R.; García-López, L. M., y Serra-Olivares, J. (2016). Sport Education Model and Self-determination Theory: An intervention in secondary school children. *Kinesiology, 48*(1), 30-38. https://doi.org/10.26582/k.48.1.15

Dössegger, A.; Ruch, N.; Jimmy, G.; Braun-Fahrländer, C.; Mäder, U.; Hänggi, J.; … Bringolf-Isler, B. (2014). Reactivity to accelerometer measurement of children and adolescents. *Medicine and Science in Sports and Exercise, 46*(6), 1140–1146. https://doi.org/10.1249/MSS.0000000000000215

Evenson, K. R.; Catellier, D. J.; Gill, K.; Ondrak, K. S., y McMurray, R. G. (2008). Calibration of two objective measures of physical activity for children. *Journal of Sports Sciences, 26*(14), 1557-1565. https://doi.org/10.1080/02640410802334196

Field, A. (2017). *Discovering statistics using IBM SPSS statistics*: Sage.

Iturbide-Luquin, L. M., y Elosua-Oliden, P. (2017). Los valores asociados al deporte: análisis y evaluación de la deportividad. *Revista de Psicodidáctica, 22*(1), 29-36. https://doi.org/10.1016/S1136-1034(17)30041-2

Ko, B.; Wallhead, T., y Ward, P. (2006). Professional development workshops – What do teachers learn and use? *Journal of Teaching in Physical Education, 25*(4), 397-412. https://doi.org/10.1123/jtpe.25.4.397

Lamoneda, J.; Huertas, F. J.; Córdoba, L. G., y García, A. V. (2014). Adaptación de la Escala Multidimensional de Orientaciones hacia la Deportividad al contexto del fútbol alevín. *Cuadernos de Psicología del Deporte, 14*(2), 71-80. https://doi.org/10.4321/S1578-84232014000200008

Li, W.; Xiang, P.; Chen, Y., y Xie, X. (2017). Unit of analysis: Impact of Silverman and Solmon's article on field-based intervention research in Physical Education in the U.S.A. *Journal of Teaching in Physical Education, 36*(2), 131–141. https://doi.org/10.1123/jtpe.2016-0169

Martínez de Ojeda Pérez, D.; Méndez-Giménez, A., y Pérez, J. V. (2016). Efectos del modelo Educación Deportiva en el clima social del aula, la competencia percibida y la intención de ser físicamente activo: un estudio prolongado en primaria. *Sport TK-Revista EuroAmericana de Ciencias del Deporte, 5*(2), 153-166.

Martínez-Gómez, D.; Martínez-De-Haro, V.; Del-Campo, J.; Zapatera, B.; Welk, G. J.; Villagra, A.; ... Veiga, Ó. L. (2009). Validez de cuatro cuestionarios para valorar la actividad física en adolescentes españoles. *Gaceta Sanitaria, 23*(6), 512-517. https://doi.org/10.1016/j.gaceta.2009.02.013

Méndez-Giménez, A.; Fernández-Río, J., y Méndez-Alonso, D. (2015). Modelo de educación deportiva versus modelo tradicional: efectos en la motivación y deportividad. *Revista internacional de medicina y ciencias de la actividad física y del deporte, 15*(59), 449–466. https://doi.org/10.15366/rimcafd2015.59.004

Menéndez-Ferreira, R.; Barquin, R.; Maldonado, A., y Camacho, D. (2018). Análisis y propuesta de una herramienta basada en gamificación para la educación en valores dentro del deporte. *En XVIII Conferencia de la Asociación española para la Inteligencia Artificial* (pp. 1039–1045).

Metzler, M. W. (2011). *Instructional models for physical education*. Scottsdale, Arizona: Holcomb Hathaway.

Migueles, J. H.; Cadenas-Sánchez, C.; Ekelund, U.; Nyström, C. D.; Mora-Gonzalez, J.; Löf, M., ... Ortega, F. B. (2017). Accelerometer data collection and processing criteria to assess physical activity and other outcomes: A systematic review and practical considerations. *Sports Medicine, 47*(9), 1821-1845. https://doi.org/10.1007/s40279-017-0716-0

Ministerio de Educación, Cultura y Deporte. (2015). Real Decreto 1105/2014, de 26 de diciembre, por el que se establece el currículo básico de la Educación Secundaria Obligatoria y del Bachillerato. *Boletín Oficial del Estado, 3,* 169–256.

Moreno, J. A.; Moreno, R., y Cervelló, E. (2007). El autoconcepto físico como predictor de la intención de ser físicamente activo. *Psicología y Salud, 17*(2), 261 – 267.

Nader, P. R.; Bradley, R. H.; Houts, R. M.; McRitchie, S. L., y O'Brien, M. (2008). Moderate-to-vigorous physical activity from ages 9 to 15 years. *Jama, 300*(3), 295-305. https://doi.org/10.1001/jama.300.3.295

Oliver, M.; Badland, H. M.; Schofield, G. M., y Shepherd, J. (2011). Identification of accelerometer nonwear time and sedentary behavior. *Research quarterly for* *exercise and sport, 82*(4), 779-783. https://doi.org/10.1080/02701367.2011.10599814

Orme, M.; Wijndaele, K.; Sharp, S. J.; Westgate, K.; Ekelund, U., y Brage, S. (2014). Combined influence of epoch length, cut-point and bout duration on accelerometry-derived physical activity. *International Journal of Behavioral Nutrition and Physical Activity, 11*(1), 34. https://doi.org/10.1186/1479-5868-11-34

Parra-Saldías, M.; Mayorga-Vega, D.; López-Fernández, I., & Viciana, J. (2018). How many daily steps are really enough for adolescents? A cross-validation study. *Retos, 33*, 241-246.

Pearson, N.; Haycraft, E.; Johnston, J. P., y Atkin, A. J. (2017). Sedentary behaviour across the primary-secondary school transition: A systematic review. *Preventive Medicine, 94*(1), 40–47. https://doi.org/10.1016/j.ypmed.2016.11.010

Perlman, D. (2012). The influence of the Sport Education Model on amotivated students' in-class physical activity. *European Physical Education Review, 18*(3), 335-345. https://doi.org/10.1177/1356336X12450795

Perlman, D., y Goc Karp, G. (2010). A Self-Determined Perspective of the Sport Education Model. *Physical Education and Sport Pedagogy, 15*(4), 401-418. https://doi.org/10.1080/17408980903535800

Poitras, V. J.; Gray, C. E.; Borghese, M.; Carson, V.; Chaput, J. P.; Janssen, I.; ... Sampson, M. (2016). Systematic review of the relationships between objectively measured physical activity and health indicators in school-aged children and youth. *Applied Physiology, Nutrition and Metabolism, 41*(6), 197-239. https://doi.org/10.1139/apnm-2015-0663

Rocamora, I.; González-Víllora, S.; Fernández-Río, J., y Arias-Palencia, N. M. (2019). Physical activity levels, game performance and friendship goals using two different pedagogical models: Sport Education and Direct Instruction. *Physical Education and Sport Pedagogy, 24*(1), 87-102. https://doi.org/10.1080/17408989.2018.1561839

Ruiz, J. V.; Ponce, A.; Sanz, E., y Valdemoros, M. A. (2015). La educación en valores desde el deporte: investigación sobre la aplicación de un programa integral en deportes de equipo. *Retos: nuevas tendencias en educación física, deporte y recreación, 28*, 270-276.

Siedentop, D.; Hastie, P. A., y Van der Mars, H. (2011). *Complete guide to Sport Education*. Champaign, Illinois: Human Kinetics

Stewart, A.; Marfell-Jones, M.; Olds, T., y De Ridder, J. (2011). *International standards for anthropometric assessment*. New Zealand: International Society for the Advancement of Kinanthropometry.

Tremblay, M. S.; Carson, V.; Chaput, J. P.; Connor-Gorber, S.; Dinh, T.; Duggan, M.; ... Zehr, L. (2016). Canadian 24-hour movement guidelines for children and youth: An integration of physical activity, sedentary behaviour, and sleep. *Applied Physiology, Nutrition, and Metabolism, 41*(6), S311-327. https://doi.org/10.1139/apnm-2016-0151

Trost S. G.; Loprinzi P. D.; Moore R., y Pfeiffer K. A. (2011). Comparison of accelerometer cut points for predicting activity intensity in youth. *Medicine Science and Sports Exercise, 43*(7), 1360-1368. https://doi.org/10.1249/MSS.0b013e318206476e

Vallerand, R.J.; Brière, N.M.; Blanchard, C., y Provencher, P. (1997). Development and validation of the multidimensional Sportspersonship orientations scale. *Journal of Sport and Exercise Psychology, 19,* 197-206. https://doi.org/10.1123/jsep.19.2.197

Viciana, J.; Mayorga-Vega, D.; Ruiz, J., y Blanco, H. (2016). La comunicación educativa de entrenadores de fútbol en competición. *Retos, 29*(1), 17-21.

Viciana, J.; Casado-Robles, C.; Pérez-Macías, L., & Mayorga-Vega, D. (2020). A Sport Education teaching unit as a citizenship education strategy in Physical Education. A group-randomized controlled trial. *Retos, 38*(38), 44-52.

Vidoni, C., y Ward, P. (2009). Effects of fair play instruction on student social skills during a middle school Sport Education unit. *Physical Education and Sport Pedagogy, 14*(3), 285–310. https://doi.org/10.1080/17408980802225818

Wahl-Alexander, Z.; Sinelnikov, O., y Curtner-Smith, M. (2016). A longitudinal analysis of students' autobiographical memories of participation in multiple

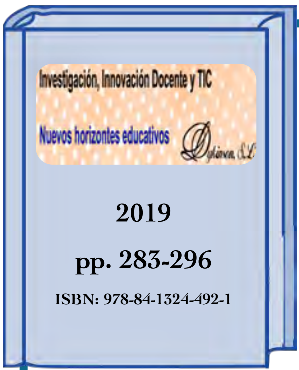
Sport Education seasons. *European Physical Education Review, 23*(1), 25-40. https://doi.org/10.1177/1356336X15624246

Wallhead, T. L.; Garn, A. C., y Vidoni, C. (2014). Effect of a Sport Education program on motivation for Physical Education and leisure-time physical activity. *Research Quarterly for Exercise and Sport, 85*(4), 478–487. https://doi.org/10.1080/02701367.2014.961051

Ward, J. K.; Hastie, P. A.; Wadsworth, D. D.; Foote, S.; Brock, S. J., y Hollett, N. (2017). A Sport education fitness season's impact on students' fitness levels, knowledge, and in-class physical activity. *Research quarterly for exercise and sport, 88*(3), 346-351. https://doi.org/10.1080/02701367.2017.1321100

World Health Organization. (2014). *Global Status Report on Noncommunicable Diseases 2014*. Geneva: WHO

Paper II



PLANIFICACIÓN EN EDUCACIÓN FÍSICA MEDIANTE UNIDADES DIDÁCTICAS INNOVADORAS PARA INCREMENTAR LOS NIVELES DE ACTIVIDAD FÍSICA HABITUAL DE LOS ESCOLARES

Casado-Robles, C., Guijarro-Romero, S., & Mayorga-Vega, D.

Investigación, Innovación docente y TIC. Nuevos Horizontes Educativos

2019. 283-296. ISBN: 978-84-1324-492-1

Paper II

ABSTRACT

Engaging in regular physical activity and reducing time spend in sedentary behavior have numerous benefits on students' health and well-being. Unfortunately, worldwide 81% of adolescents do not achieve the physical activity recommendations and also spend more than two-thirds of the day involved in sedentary behavior. Consequently, the increase of habitual physical activity levels and the reduction of time spent in sedentary behavior among young people is an important educational challenge. However, in the task of Physical Education planning, teachers deal with some difficulties that make more difficult the achievement of this objective. To overcome these limitations, four applications of the innovative teaching units (intermittent, alternated, reinforced and irregular) (Viciana & Mayorga-Vega, 2016) are proposed to help Physical Education teachers to design effective and viable programs to increase habitual physical activity levels among students.

KEYWORDS

School programs. Children. Adolescents. Educational curriculum. Educational innovation.

La práctica de actividad física en los escolares

En la actualidad existe una fuerte evidencia científica de que la práctica regular de actividad física (AF), así como la reducción de conducta sedentaria (CS), suponen numerosos beneficios sobre la salud y bienestar de las personas (Reiner, Niermann, Jekauc y Woll, 2013), especialmente sobre los niños y adolescentes (Carson et al., 2016; Poitras et al., 2016). Por ejemplo, en los jóvenes la práctica regular de AF no solo tiene beneficios físicos (por ejemplo, disminuyendo el riesgo de mortalidad o el desarrollo de un sistema cardiovascular sano), sino también efectos psicológicos (por ejemplo, control de la ansiedad y la depresión) y sociales (por ejemplo, integración o la autoconfianza) (Organización Mundial de la Salud, 2014). Para alcanzar los beneficios derivados de la práctica de AF, la Organización Mundial de la Salud (2010) ha establecido unas recomendaciones mundiales sobre la AF para la salud, indicando que los escolares desde los 5 a los 17 años de edad deben realizar al menos 60 minutos diarios de AF moderada-vigorosa (AFMV). Específicamente, la AF diaria debería componerse principalmente de ejercicio aeróbico e incluir actividades vigorosas que refuercen los músculos y huesos (Organización Mundial de la Salud, 2010).

Por otro lado, una gran cantidad de CS se encuentra relacionada con marcadores negativos de salud cardiovascular en niños y adolescentes (Carson et al., 2016). Por ejemplo, se relaciona con mayor riesgo de poseer una composición corporal desfavorable, peores niveles de estado físico o menor autoconfianza (Carson et al., 2016). Por lo tanto, resulta de vital importancia que independientemente de los niveles de AF, los escolares no empleen grandes cantidades de tiempo diario en CS para evitar posibles problemas de salud (Santos et al., 2014). De acuerdo con las Recomendaciones Canadienses del Movimiento, se debería limitar el tiempo de uso recreacional de pantalla a no más de dos horas diarias y evitar estar sentado periodos prolongados de tiempo (Tremblay et al., 2016). Además, dichas recomendaciones también indican que reemplazar el tiempo en CS por más AF, en especial AFMV, puede proporcionar mayores beneficios en la salud de los escolares (Tremblay et al., 2016).

Desafortunadamente, los niveles de práctica de AF disminuyen drásticamente en la transición de la etapa infantil a la adolescencia (Corder, Winpenny, Love, Brown, White y Sluijs, 2017) a favor de un mayor tiempo empleado en CS (Pearson, Haycraft, Johnston y Atkin, 2017). Estudios previos constatan el incumplimiento de las recomendaciones mundiales de AFMV en la población en general (De Moraes, Guerra y Menezes, 2013). Concretamente, el 81% de los adolescentes a nivel mundial no cumplen con la recomendación de los 60 minutos diarios de AFMV, siendo la tasa de incumplimiento entre los adolescentes españoles del 84% (Organización Mundial de la Salud, 2018). Además, los escolares emplean más de dos tercios del tiempo de vigilia diario involucrados en CS (Ruiz et al., 2015). Estos resultados son alarmantes teniendo en cuenta que la inactividad física se considera el cuarto factor de riesgo de mortalidad a nivel mundial (International Sport and Culture Association, 2015).

La promoción de la práctica de actividad física en el contexto escolar

El incremento de la AFMV y la reducción de la CS entre los jóvenes es una prioridad de salud pública (International Sport and Culture Association, 2015) y un importante reto educativo (Ministerio de Educación, Cultura y Deporte, 2015). Los organismos responsables de la sanidad pública, así como diferentes autores en la literatura científica, han promovido diversos objetivos y estrategias de promoción de AF en los jóvenes (Ministerio de Educación, Cultura y Deporte, 2013; Viciana, Lozano, Cocca y Mayorga-Vega, 2012).

La escuela es considerada un contexto ideal para la promoción de AF (Organización Mundial de la Salud, 2018; Viciana, Martínez-Baena y Mayorga-Vega, 2015) y según el informe Eurydice de la Comisión Europea, el 80% de los niños en edad escolar solo practican AF en la escuela (Comisión Europea, 2013). Dado que la escuela tiene carácter obligatorio y la mayoría del tiempo diario los niños y adolescentes están en el centro escolar, algunos autores recomiendan que durante este periodo de tiempo, los estudiantes se involucren en al menos 30 minutos de AFMV (Kriemler, Meyer, Martin, van Sluijs, Andersen y Martin, 2011). En especial, la asignatura de Educación Física juega un papel clave, debido a su carácter obligatorio en el currículum educativo y a que está dirigida por profesionales en el ámbito de la AF y la salud (Viciana, Martínez-Baena y Mayorga-Vega, 2015). La Educación Física también ha demostrado incrementar los niveles diarios de AF de los escolares frente a los índices registrados los días en que no hay Educación Física

(Mayorga-Vega, Martínez-Baena y Viciana, 2018; Viciana, Mayorga-Vega y Parra-Saldías, 2019), jugando un rol esencial en el cumplimiento de las recomendaciones diarias de AFMV. Además, desde esta asignatura los profesores de Educación Física tienen la oportunidad de sensibilizar a los alumnos de la importancia de mantener un estilo de vida activo y los beneficios que conlleva su práctica fuera del contexto escolar. Dentro del marco curricular de la asignatura de Educación Física se destaca como uno de sus objetivos principales el desarrollo y consolidación de hábitos saludables de AF regular en la etapa de Educación Secundaria Obligatoria y la ocupación activa del ocio y tiempo libre de los escolares (Ministerio de Educación, Cultura y Deporte, 2013, 2015; SHAPE-America, 2014).

Principales limitaciones relacionadas con la planificación en educación física y las unidades didácticas innovadoras

Sin embargo, diferentes limitaciones en el ámbito educativo impiden a los profesores de Educación Física cumplir con la promoción de la AF habitual entre los escolares de manera eficaz. En primer lugar, la asignatura de Educación Física tiene una presencia muy reducida dentro del sistema educativo. Según el informe de la UNESCO a nivel mundial en las escuelas se dedica un tiempo muy limitado a la Educación Física (Hardman, Murphy, Routen y Tones, 2014). Concretamente en los países europeos, de media se dedican aproximadamente 110 minutos a la semana en Educación Primaria, lo que representaría tan solo dos clases de 55 minutos semanales, y esta duración es aún menor en la etapa de Educación Secundaria Obligatoria donde se reduce en beneficio de otras áreas curriculares (Comisión Europea, 2013; Hardman et al., 2014). De forma reiterada, los profesores de Educación Física han denunciado esta baja carga horaria asignada a la asignatura y puesto de manifiesto los consecuentes perjuicios para el desarrollo de todos los objetivos presentes en el currículo educativo de forma adecuada (Albarracín, Moreno y Beltrán, 2014). En segundo lugar, el currículo educativo indica que los profesores de Educación Física deben desarrollar un gran número de contenidos y objetivos educativos a lo largo del curso escolar. Además del fomento de la práctica de AF, se deben trabajar otros contenidos tales como el desarrollo de la condición física y motriz o la expresión corporal (Ministerio de Educación, Cultura y Deporte, 2015; SHAPE-America, 2014).

Por otro lado, la distribución de las clases de Educación Física no se basa en criterios relacionados con la AF, sino que simplemente está regido por ajustar correctamente el horario del profesorado del centro educativo (Viciana, Mayorga-Vega y Merino-Marban, 2014). Por lo tanto, en algunos centros las clases de Educación Física pueden ser impartidas dos días consecutivos sin dejar el tiempo necesario de descanso, o también, pueden impartirse a primera hora de la mañana cuando puede hacer mucho frío durante el invierno en ciertas zonas del país, o por el contrario a última hora cuando puede hacer mucho calor en primavera en las zonas más cálidas. Además, otra limitación es el hecho de que las clases de Educación Física pueden verse interrumpidas constantemente a lo largo del curso escolar por diversos factores (Viciana, Mayorga-Vega y Merino-Marban, 2014). Por ejemplo, los aspectos meteorológicos (es decir, lluvia o calor/ frío extremo) pueden provocar que se vea suspendida la práctica de Educación Física si no se disponen de espacios cubiertos, o si, aunque los hubiera se produce coincidencia horaria con otro profesor de Educación Física y ya están ocupados. Otro ejemplo de interrupciones son los periodos vacacionales (por ejemplo, Navidad, Semana Santa o días festivos a lo largo del curso escolar) que dificultan el proceso de planificación del periodo de enseñanza.

Esto provoca que se apliquen programas muy cortos e insuficientes para la promoción de la AF y que la asignatura por sí sola no pueda satisfacer las necesidades de AF de los escolares y conseguir resultados significativos. Las unidades didácticas propuestas tradicionalmente se basan en trabajar un número concreto de contenidos de forma aislada, en lugar de estar realmente centradas en alcanzar objetivos significativos para la asignatura. Comúnmente, los profesores de Educación Física consideran la planificación anual como un grupo de segmentos de tiempo vacíos que hay que rellenar con contenidos, basando su elección en criterios personales o en los recursos disponibles, pero que no están correctamente diseñados para alcanzar los objetivos propuestos. Además, las sesiones que componen una unidad didáctica tradicional se desarrollan todas seguidas durante un tiempo determinado y dedicando la sesión completa al mismo objetivo. En consecuencia, no promueven la conexión de aprendizajes entre diferentes objetivos de la asignatura de Educación Física debido a la falta de coherencia de la planificación anual (Viciana y Mayorga-Vega, 2016).

Lamentablemente, las unidades didácticas tradicionales enfocadas a la promoción de hábitos de AF presentan una eficacia escasa en la actualidad. Por lo tanto, es necesario plantear y desarrollar nuevas estrategias que sean efectivas para el fomento de los hábitos de AF saludable en los escolares (Mayorga-Vega, Viciana y Cocca, 2013). De este modo, surgen los modelos de unidades didácticas innovadoras (Viciana y Mayorga-Vega, 2016), con el objetivo de resolver diferentes problemas en la planificación de la asignatura de Educación Física y que pueden ser aplicados para la promoción de la AF y reducción de CS. Estas unidades didácticas están basadas en los principios originales de la planificación educativa: la flexibilidad, la eficacia y basada en objetivos (Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, UNESCO, 1962). Las unidades didácticas innovadoras pretenden profundizar en los conceptos de dinamismo y flexibilidad de la planificación educativa, planteando una "distribución innovadora del tiempo" en función del objetivo educativo perseguido (Viciana y Mayorga-Vega, 2016). Aprovechando estos modelos alternativos a los empleados hasta ahora para llevar a cabo la planificación de las clases de Educación Física, se pueden aportar soluciones a los profesores y solventar los problemas relacionados con la promoción de hábitos de AF y reducción de CS. A continuación, se hace una propuesta para promover los hábitos de AF saludables en escolares mediante la aplicación de las unidades didácticas intermitentes, alternadas, irregulares y reforzadas (Viciana y Mayorga-Vega, 2016).

Propuestas de intervención en educación física

Unidad Didáctica Intermitente: Las pulseras de fitness aplicadas a una Unidad Didáctica Intermitente en Educación Física sobre la actividad física habitual y la conducta sedentaria

La propuesta de Unidad Didáctica Intermitente se basa en dedicar únicamente una franja de tiempo de la sesión de Educación Física al objetivo curricular propuesto (por ejemplo, de 5 a 15 minutos), durante un periodo de tiempo específico de cada sesión. Esta unidad didáctica permite conectar el aprendizaje de los estudiantes entre los contenidos desarrollados en las dos o más franjas de la sesión. De este modo, se podría proponer una Unidad Didáctica Intermitente basada en la promoción de los hábitos de AF y reducción de CS durante la última franja de la sesión de Educación Física (es decir, la vuelta a la calma) con la ayuda de pulseras de fitness, dedicando el resto de la sesión a otro contenido curricular diferente.

En la Figura 1 se puede observar el esquema de la Unidad Didáctica Intermitente. En concreto, se propone que los escolares lleven pulseras de fitness durante un trimestre académico completo (aproximadamente 9 semanas) como estrategia para aumentar la AF habitual y reducir la CS. Durante ese tiempo se aplica simultáneamente una Unidad Didáctica Intermitente, únicamente durante la vuelta a la calma de las sesiones regulares de Educación Física (es decir, los 10 últimos minutos de la sesión) para analizar el *feedback* registrado por la pulsera de fitness. El programa intermitente engloba diferentes estrategias: (a) Sesiones educativas sobre los niveles de AF recomendados, beneficios de la práctica de AF, posibilidades de práctica saludables para su tiempo de ocio, transporte activo al colegio o actividades sedentarias a evitar; (b) Establecimiento de metas o retos diarios de AF (por ejemplo, 10.000 pasos diarios) (Parra-Saldías, Mayorga-Vega, López-Fernández y Viciana, 2018) y su registro diario en la aplicación móvil propia de la pulsera de fitness; (c) Asesoramiento sobre el uso de las pulseras de actividad para que reporten *feedback* útil a tiempo real y continuo durante todo el programa (por ejemplo, avisos de CS o consecución de la meta propuesta); y (d) seguimiento diario de la AF de los estudiantes y establecimiento de retos competitivos entre clases mediante un grupo privado de Facebook, donde anotará cualquier cuestión anecdótica diaria de importancia para el participante, como estrategia de apoyo social del programa.

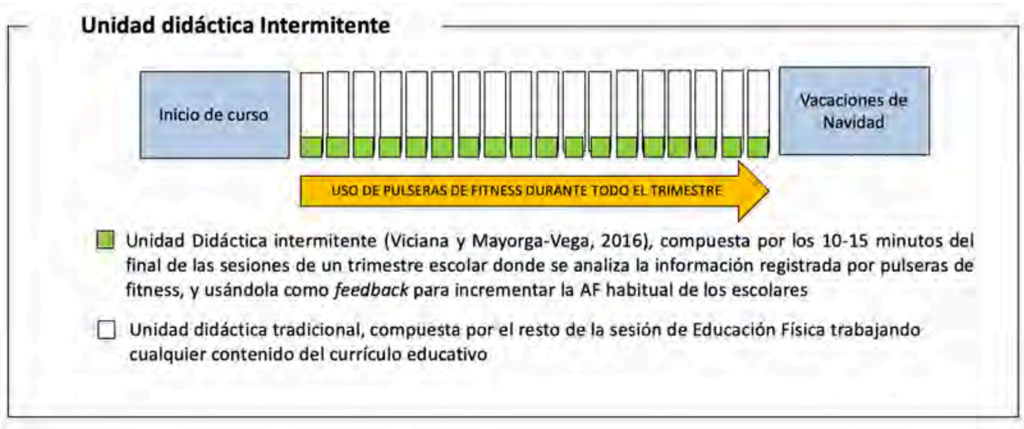


Figura 1. Esquema de la propuesta de Unidad Didáctica Intermitente

Unidad Didáctica Reforzada: Las pulseras de fitness aplicadas durante las vacaciones de Navidad para el mantenimiento de la actividad física habitual

Las Unidades Didácticas Reforzadas tienen en cuenta que, para alcanzar una mejora en un determinado aprendizaje, es necesario el desarrollo del mismo durante un determinado periodo de tiempo, pero también se precisa de un posterior mantenimiento a lo largo de todo el curso escolar (Mayorga-Vega, Montoro-Escaño, Merino-Marban y Viciana, 2015). El refuerzo podría mantener las mejoras alcanzadas durante la unidad didáctica de desarrollo y evitar las posibles pérdidas de aprendizaje o hábitos alcanzados después de un tiempo. Además, permitiría el "sobreaprendizaje", al extender la práctica exitosa del alumno para motivar y consolidar el aprendizaje obtenido previamente. Por ejemplo, los periodos vacacionales existentes a lo largo del curso escolar (es decir, Navidad y Semana Santa) son periodos donde se pueden producir pérdidas del aprendizaje obtenido y puede ser un momento ideal para aplicar una unidad didáctica de refuerzo y producir la "reactivación del aprendizaje" antes de retornar a valores basales.

En la Figura 2 se observa el esquema de la propuesta Unidad Didáctica Reforzada. Siguiendo con la aplicación de estos modelos para la promoción de hábitos de AF, la propuesta de Unidad Didáctica Reforzada está encaminada a fomentar el mantenimiento, durante las vacaciones de Navidad, de los niveles de AF alcanzados durante la Unidad Didáctica Intermitente propuesta anteriormente (entendida como unidad didáctica de desarrollo). Para aplicar la Unidad Didáctica Reforzada, antes de las vacaciones de Navidad se proponen y entregan programas individualizados de práctica de AF adaptados al nivel de práctica de los estudiantes. El objetivo es que, tras haber incrementado los niveles de AF y aprendido a utilizar las pulseras de fitness, las utilicen como una herramienta de apoyo para seguir el programa individualizado de forma autónoma durante las vacaciones de Navidad (aproximadamente 3 semanas). En dicho programa individualizado se indicará el reto de pasos diarios a cumplir y las actividades a desarrollar destacando el tiempo y la intensidad de trabajo. Además, de forma simultánea se llevará un seguimiento de la actividad realizada por los alumnos a través de una red social (por ejemplo, grupo privado de Facebook), para animar a que continúen con el programa y solventar posibles dudas.

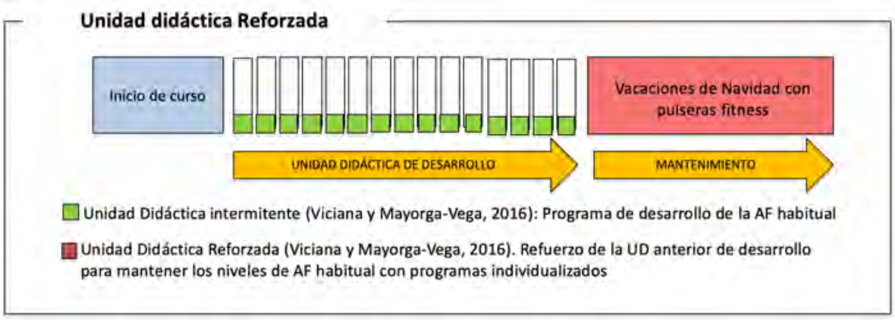


Figura 2. Esquema de la propuesta de Unidad Didáctica Reforzada

Unidades Didácticas Alternadas: Unidades Didácticas Alternadas *indoor y outdoor* en Educación Física para fomentar el conocimiento del entorno y la práctica de actividad física en el tiempo de ocio

La propuesta de Unidades Didácticas Alternadas surge con el propósito de conectar el aprendizaje desarrollado entre dos unidades didácticas desarrolladas durante el mismo periodo de tiempo, para permitir a los estudiantes entender las conexiones fundamentales entre ellas y construir un aprendizaje significativo. Una de las principales ventajas del empleo de estas unidades didácticas es transferir el aprendizaje desarrollado en el aula de Educación Física a la vida diaria del alumnado. Aunque dicha transferencia y la relación entre diferentes contenidos se podría aplicar entre dos o más tareas dentro de la misma sesión de Educación Física, es una propuesta interesante si se quiere desarrollar en diferentes contextos o con diferentes materiales (por ejemplo, AF en espacios interiores y AF en el entorno próximo). Además, permite solucionar el problema de coincidencia de horario con otros profesores de Educación Física para utilizar el gimnasio escolar a la misma hora, ya que un día el contenido puede ser enseñado dentro del gimnasio (es decir, con los materiales y espacios convencionales del mismo) y el otro día de la semana el contenido puede ser enseñado fuera (es decir, utilizando los espacios y recursos del entorno), facilitando la distribución de instalaciones entre el profesorado. De este modo, se podrían proponer Unidades Didácticas Alternadas basadas en el conocimiento del entorno próximo al centro educativo para la práctica de AF durante el tiempo de ocio de los escolares.

En la Figura 3 se observa el esquema de la propuesta de Unidades Didácticas Alternadas, que consiste en la realización de dos unidades didácticas de cuatro sesiones de Educación Física cada una (ocho sesiones en total), siguiendo la dinámica de impartir una sesión dentro del centro escolar (unidad didáctica *indoor*) seguida de otra utilizando el entorno próximo al centro (unidad didáctica *outdoor*) durante todo el programa. El objetivo es establecer una conexión en el aprendizaje de tareas para practicar AF en ambos espacios. Los contenidos a trabajar durante ambas unidades didácticas son la condición física y deportes, pero la principal diferencia reside en el material y espacio utilizado en cada unidad didáctica. La unidad didáctica *indoor* se desarrolla utilizando los materiales tradicionales del aula de Educación Física (por ejemplo, bancos suecos, balones medicinales, espalderas o colchonetas) y los espacios habituales dentro del centro educativo (gimnasio cubierto y pista polideportiva). Por su parte, la unidad didáctica *outdoor* se desarrolla aprovechando los espacios y recursos materiales que ofrecía el entorno próximo al centro educativo (por ejemplo, parques o polideportivos municipales). Además, mientras se desarrolla la sesión *indoor* el profesor de Educación Física puede hacer referencia a cómo realizar estos ejercicios de forma similar *outdoor* para establecer la conexión entre ambos espacios, y viceversa.

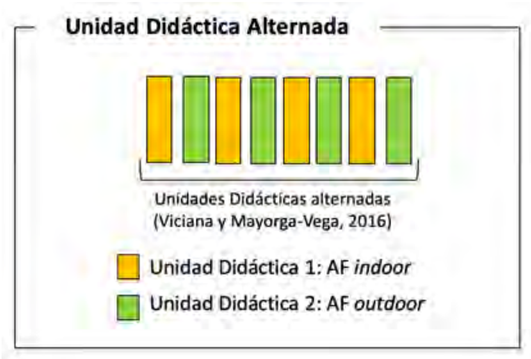


Figura 3. Esquema de la propuesta de Unidades Didácticas Alternadas

Unidad Didáctica Irregular: Programa de Educación Deportiva en Educación Física para el incremento de la actividad física desarrollada en los recreos escolares

Las unidades didácticas irregulares surgen del carácter flexible y dinámico de la planificación educativa en su máximo nivel, con el objetivo de alcanzar objetivos y resultados auténticos. La aplicación de estas unidades didácticas permite distribuir las sesiones de Educación Física de todo el curso escolar con un amplio abanico de posibilidades de organización, en función del centro de interés perseguido por el profesor. Una de las ventajas principales de estas unidades didácticas se basa en incrementar la cantidad de oportunidades de aprendizaje del alumnado utilizando otros contextos extracurriculares fuera de las clases de Educación Física (por ejemplo, recreos escolares, tiempo después del colegio o fines de semana), en función de las características y madurez del alumnado. Por ejemplo, la aplicación de una Unidad Didáctica Irregular que combine el contexto formal de Educación Física con la franja extracurricular del recreo escolar, puede permitir el incremento de los niveles de AF del alumnado durante esa franja y su consecuente contribución las recomendaciones diarias de AFMV.

En la Figura 4 se observa el esquema de la propuesta de Unidad Didáctica Irregular. La propuesta de Unidad Didáctica Irregular es la aplicación del Modelo de Educación Deportiva (Siedentop, 1994) durante 12 sesiones regulares de Educación Física, combinado con el uso de la franja extracurricular del recreo escolar (ocho recreos) para desarrollar la fase de competiciones del Modelo de Educación Deportiva. Por lo tanto, la estructura de la unidad didáctica quedaría dividida en una sesión de Educación Física introductoria, dos sesiones dirigidas, ocho sesiones de práctica autónoma, ocho recreos de competición formal y una sesión de evento final. De este modo, se utiliza una franja extracurricular como un tiempo extra de aprendizaje y además incrementar los niveles de AF durante el recreo y la reducción de CS.

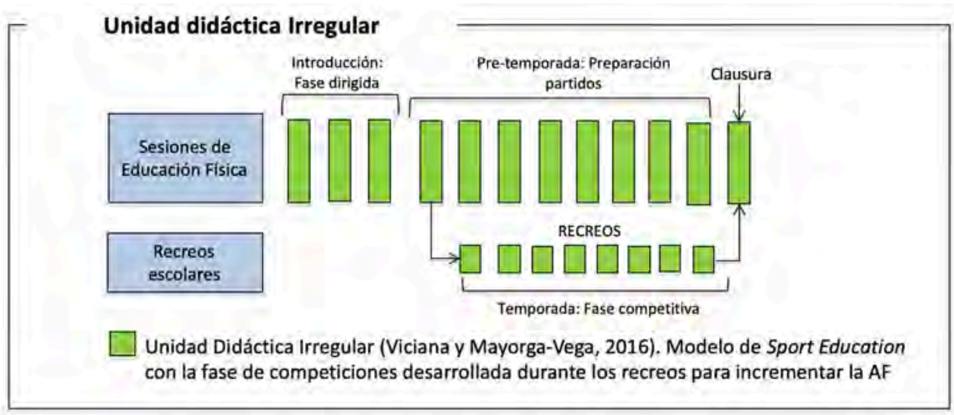


Figura 4. Esquema de la propuesta de Unidad Didáctica Irregular

Conclusiones

La eficacia de los programas tradicionales diseñados para promover los hábitos de AF y la reducción de CS es reducida en la actualidad. Por lo tanto, parece lógico plantearse la importancia y necesidad de buscar estrategias efectivas para fomentar los hábitos saludables de AF en la edad escolar. Basándonos en las características de la flexibilidad y dinamismo de la planificación educativa, los modelos de unidades didácticas innovadoras propuestos parecen ser una guía excelente para ayudar a los profesores de Educación Física a diseñar programas efectivos y viables con el objetivo de incrementar los niveles habituales de AF y reducir el tiempo en CS de los escolares. Además, el empleo de estas unidades didácticas innovadoras permite solventar algunas limitaciones comunes relacionadas con la planificación de la asignatura de Educación Física (por ejemplo, el gran volumen de contenidos y objetivos a desarrollar durante el curso escolar sumado a un escaso tiempo curricular dedicado a la asignatura, se podría solventar utilizando periodos extracurriculares para incrementar el tiempo de práctica o el desarrollo de varios objetivos curriculares al mismo tiempo con una unidad didáctica intermitente). Por lo tanto, se recomiendan estudios futuros que pongan en práctica las unidades didácticas innovadoras propuestas para comprobar su efectividad en el incremento de la AF habitual y reducción de CS en estudiantes de Educación Secundaria Obligatoria.

Referencias bibliográficas

Albarracín, A., Moreno, J.A. y Beltrán, V. (2014). La situación actual de la Educación Física según su profesorado: Un estudio cualitativo con profesorado de la Región de Murcia. *Cultura Ciencia y Deporte, 27*, 225-234.

Carson, V., Hunter, S., Kuzik, N., Gray, C. E., Poitras, V. J., Chaput, J. P., ... y Kho, M. E. (2016). Systematic review of sedentary behaviour and health indicators in school-aged children and youth: an update. *Applied Physiology, Nutrition, and Metabolism*, *41*(6), 240-265.

Comisión Europea/EACEA/Eurydice. (2013). *Physical education and sport at school in Europe Eurydice Report*. Luxemburgo: Oficina de Publicaciones de la Unión Europea.

Corder, K., Winpenny, E., Love, R., Brown, H. E., White, M. y Sluijs, E.V. (2017). Change in physical activity from adolescence to early adulthood: a systematic review and meta-analysis of longitudinal cohort studies. *British Journal of Sports Medicine, 53*(8), 496-503.

De Moraes, A.C., Guerra, P.H. y Menezes, P.R. (2013). The worldwide prevalence of insufficient physical activity in adolescents; a systematic review. *Nutrición hospitalaria*, *28*(3), 575-584.

Hardman, K., Murphy, C., Routen, A. y Tones, S. (2014). *UNESCO-NWCPEA: World-wide survey of school Physical Education*. París: United Nations Educational, Scientific and Cultural Organization.

International Sport and Culture Association (2015). *The economic cost of Physical inactivity in Europe. An ISCA/ Cebr report.* UK: ISCA/ Cebr Office.

Kriemler, S., Meyer, U., Martin, E., van Sluijs, E. M., Andersen, L. B. y Martin, B. W. (2011). Effect of school-based interventions on physical activity and fitness in children and adolescents: a review of reviews and systematic update. *British Journal of Sports Medicine, 45*, 923-930.

Mayorga-Vega, D., Martínez-Baena, A. y Viciana, J. (2018). Does school physical education really contribute to accelerometer-measured daily physical activity and non sedentary behaviour in high school students? *Journal of sports sciences, 36*(17), 1913-1922.

Mayorga-Vega, D., Montoro-Escaño, J., Merino-Marban, R. y Viciana, J. (2015). Effects of a physical education-based programme on health-related physical fitness and its maintenance in high school students: A cluster-randomized controlled trial. *European Physical Education Review, 22*(2), 243-259.

Mayorga-Vega, D., Viciana, J. y Cocca, A. (2013). Effects of a circuit training program on muscular and cardiovascular endurance and their maintenance in schoolchildren. *Journal of Human Kinetics, 37*(1), 153-160.

Ministerio de Educación, Cultura y Deporte. (2013). Ley Orgánica 2/2006, de 3 de mayo, de Educación. *Boletín Oficial del Estado, 106*, 17158-17207.

Ministerio de Educación, Cultura y Deporte. (2015). Real Decreto 1105/2014, de 26 de diciembre, por el que se establece el currículo básico de la Educación Secundaria Obligatoria y del Bachillerato. *Boletín Oficial del Estado, 3*, 169–256.

Organización Mundial de la Salud. (2010). *Global recommendations on physical activity for health.* Geneva: WHO

Organización Mundial de la Salud. (2014). *Global Status Report on Noncommunicable Diseases 2014*. Geneva: WHO

Organización Mundial de la Salud. (2018). *Promoting physical activity in the education sector.* Copenhagen: World Health Organization.

Parra Saldías, M., Mayorga-Vega, D., López-Fernández, I. y Viciana, J. (2018). How many daily steps are really enough for adolescents? A cross-validation study. *Retos, 33*, 241–246.

Pearson, N., Haycraft, E., Johnston, J. P. y Atkin, A. J. (2017). Sedentary behaviour across the primary-secondary school transition: A systematic review. *Preventive Medicine, 94*, 40–47.

Poitras, V. J., Gray, C. E., Borghese, M. M., Carson, V., Chaput, J. P., Janssen, I., ... y Sampson, M. (2016). Systematic review of the relationships between objectively measured physical activity and health indicators in school-aged children and youth. *Applied Physiology, Nutrition and Metabolism, 41*(6), 197-239.

Reiner, M., Niermann, C., Jekauc, D. y Woll, A. (2013). Long-term health benefits of physical activity: A systematic review of longitudinal studies. *BMC Public Health, 13*(1), 813.

Ruiz, J. R., Huybrechts, I., Cuenca-García, M., Artero, E. G., Labayen, I., Meirhaeghe, A., ... y Marcos, A. (2015). Cardiorespiratory fitness and ideal cardiovascular health in European adolescents. *Heart*, *101*(10), 766-773.

Santos, R., Mota, J., Okely, A., Pratt, M., Moreira, C., Coelho-e-Silva, M., Vale, S. y Sardinha, L. (2014) The independent associations of sedentary behaviour and physical activity on cardiorespiratory fitness. *British Journal of Sports Medicine 48*, 1508-1512.

Shape America (2014). *National standards & grade-level outcomes for K-12 physical education.* Reston, VA: Human Kinetics.

Siedentop, D. (1994). *The Sport Education Model. Sport education: Quality PE through positive sport experiences*. Champaign, Illinois: Human Kinetics

Tremblay, M. S., Carson, V., Chaput, J. P., Connor Gorber, S., Dinh, T., Duggan, M., ... y Janssen, I. (2016). Canadian 24-hour movement guidelines for children and youth: an integration of physical activity, sedentary behaviour, and sleep. *Applied Physiology, Nutrition, and Metabolism*, *41*(6), S311-S327.

UNESCO/ED/190. (1962). *New methods and techniques in education. Report of a meeting of experts.* Paris, Francia. Recuperado de https://unesdoc.unesco.org/ark:/48223/pf0000126329

Viciana, J. y Mayorga-Vega, D. (2016). Innovative teaching units applied to physical education–changing the curriculum management for authentic outcomes. *Kinesiology, 48*(1), 142-152.

Viciana, J., Lozano, L., Cocca, A. y Mayorga-Vega, D. (2012). Influence of the organizational system on motor engagement time in physical education on high school students. *Procedia-Social and Behavioral Sciences*, *69*, 1160-1167.

Viciana, J., Martínez-Baena, A. y Mayorga-Vega, D. (2015). Contribución de la educación física a las recomendaciones diarias de actividad física en adolescentes según el género: un estudio con acelerometría. *Nutrición Hospitalaria*, *32*(3), 1246-1251.

Viciana, J., Mayorga-Vega, D. y Merino-Marban, R. (2014). Physical education-based planning for developing and maintaining students' health-related physical fitness levels. In R. Todaro (Ed.), *Handbook of physical education research: Role of school programs, children's attitudes and health implications* (pp. 237–252). New York: Nova Science Publishers.

Viciana, J., Mayorga-Vega, D. y Parra-Saldías, M. (2019). Adolescents' physical activity levels on physical education and non-physical education days according to gender, age, and weight status. *European Physical Education Review*, *25*(1), 143-155.

Paper III

SPORT EDUCATION-BASED IRREGULAR TEACHING UNIT AND STUDENTS' PHYSICAL ACTIVITY DURING SCHOOL RECESS

Casado-Robles, C., Mayorga-Vega, D., Guijarro-Romero, S., & Viciana, J.

The Journal of Educational Research

2020, *113*(4), 262-274. https://doi.org/10.1080/00220671.2020.1806014

Paper III

ABSTRACT

Objective: The aim of the study was to compare the effect of a Sport Education-based irregular teaching unit on the high-school students' physical activity (PA) during school recess.

Methods: A sample of 165 high-school students (88 boys and 77 girls) aged 13-16 years old participated in the study. The six classes were cluster-randomly assigned to the Irregular Sport Education group (Sport Education methodology with competitions in recess), Traditional Sport Education group (Sport Education methodology with competitions in Physical Education lessons), or to the control group (non-intervention). The students' PA during school recess periods were objectively assessed by GT3X accelerometers.

Results: The results of the repeated-measures nested analysis of variance showed that the Sport Education-based irregular teaching unit significantly improved the students' PA during recesses ($p < 0.001$).

Conclusions: The competitions phase of the Sport Education model performed at school recesses seems to be a valuable strategy for increasing students' PA contributing to the achievement of the daily recommendations.

KEYWORDS

School-based program. Sport Education Model. Accelerometry. Adolescents. Secondary school. Playtime.

Introduction

The World Health Organization (WHO, 2010) recommends that adolescents should achieve at least 60 minutes daily of moderate-to-vigorous physical activity (MVPA), and being that young people spend the majority of their waking hours at school, which is compulsory for all adolescents, Kriemler et al. (2011) recommended that at least 30 minutes should occur during school time. School, and especially the Physical Education (PE) subject, has been considered by the WHO (2018) as a key setting for the promotion of adolescents' physical activity (PA), playing an essential role in achieving daily PA recommendations. In recent years, there has been a rise in the number of school-based interventions focused on PA promotion (e.g., Dobbins et al., 2013). Although, how Dobbins et al. (2013) outlined in their systematic review, it is not clear what the most effective strategies are to promote healthy behaviors, Carson et al. (2014) suggested that multi-component school-based interventions should be considered in order to increase the quantity and quality of PA opportunities throughout the school day. Within these multi-component programs, a variety of factors have been considered to increase PA levels: active commuting (i.e., before/after school), PE lessons, school recess (SR), and the involvement of teachers (Carson et al., 2014; WHO, 2018). All of these factors could be controlled by the PE teacher during the school day, except active commuting as it happens outside the school. Therefore, focusing only on those which could be supervised by the teacher is a more straightforward strategy to control students' behavior in the school center and to promote healthy PA habits at school.

Therefore, SR (as a non-curricular time during school) can be an ideal context in providing supplemental PA opportunities for adolescents (Parrish et al., 2013). Some organizations such as the Center for Disease Control and Prevention and the Society of Health and Physical Educators (CDC & SHAPE-America, 2017) and the Institute of Medicine (IOM, 2013) highlight the advantage that the SR offers daily practice opportunities outside of PE, which usually has a frequency of only two days a week. Previous descriptive studies have observed that approximately 40% of SR time should be engaged in MVPA (Bailey et al., 2012) and that SR can contribute to around 33% of adolescents recommended daily MVPA levels (Viciana et al., 2016). In addition to providing the opportunity to engage in PA, SR also helps students to practice social skills, obtain academic benefits and enhance cognitive performance

(IOM, 2013). Therefore, schools have the potential to organize and implement school policies during SR to provide daily and voluntary opportunities for students to experience PA at the same time that they address these benefits (CDC & Shape-America, 2017).

A variety of strategies focusing on intervention trials carried out with the aim of enhancing adolescents' PA levels during SR have been used, including: (a) the addition of adequate sports equipment and facilitating its access to students (Méndez-Giménez et al., 2017); (b) modifications to the facilities, markings or paintings on the playground surface (Huberty et al., 2011); (c) structured activities or competitions organized by the teacher (Howe et al., 2012; Larson et al., 2014), and (d) active video game play or exergaming (Erwin et al., 2013). Recently, in recognition of the great potential of SR, the CDC and SHAPE-America (2017) proposed strategies to provide PA practice to students during SR. Among these strategies, providing planned sports activities is emphasized for increasing students' participation in PA (CDC & Shape-America, 2017).

In addition to these aforementioned strategies, the Sport Education (SE) model proposed by Siedentop et al. (2011), offers one potential strategy for promoting a physically active lifestyle. The SE model emerged to provide positive, quality sporting experiences within PE lessons, centered on taking the competition to the educational context and taking advantage of all the educational values of sport such as respect, sportsmanship or teamwork (Siedentop et al., 2011). To achieve these goals, this pedagogical approach is based on six main features: (a) Students developing a "sense of affiliation" because they become members of a team throughout the season and assume different roles (e.g., coach or referee); (b) the sport "season" (i.e., teaching unit) being longer than a traditional one and including practices and games; (c) the "culminating event" with a tournament and award ceremony to celebrate successes; (d) students developing a "formal competition" phase interspersed with autonomous practice accordingly with the schedule of competitions; (e) a "record keeping" of behaviors and results takes place throughout the competitions to give feedback and establish goals; and finally, (f) the "festive" nature of the sport being present throughout the season, creating distinctive team elements, publishing photographs or celebrating achievements (Siedentop et al., 2011).

The SE model has provided several positive effects such as improvements in students' motivation toward PA practice (Chu & Zang, 2018), intention to be physically active (Martínez de Ojeda, 2015), the increase of in-class MVPA levels (Perlman, 2012), and the improvement of leisure-time MVPA levels (Wallhead et al., 2014). Concerning the SR, Coolkens et al. (2018), combined a parkour-based SE season in PE with SR where access to the gym and parkour equipment was offered. Moreover, Wallhead et al. (2010) found that this model has the capacity to promote students' voluntary participation in a sport-based SR to explain the transfer of learning and motivation from PE to other contexts. Based on these previous outcomes, and since participation in PA during SR is completely voluntary, the use of the SE model, which has provided great scientific evidence to increase students' motivation towards PA, could increase the likelihood of students participating in PA in extra-curricular contexts.

Together with the SE model, the use of the novel distribution of time planning in PE termed "innovative teaching units", designed by Viciana and Mayorga-Vega (2016), could give PE teachers the opportunity to better organize learning/teaching periods. Specifically, the irregular teaching units were designed to increase students' total amount of learning using other extracurricular contexts beyond PE lessons, with the aim of achieving authentic outcomes. For this reason, a SE-based irregular teaching unit, using compulsory PE and extracurricular SR period to implement the SE model, would be an opportunity to examine the use of both contexts in the improvement of students' MVPA levels during SR. Previous studies (e.g., Coolkens et al., 2018; Perlman, 2012; Wallhead et al., 2010, 2014) have investigated the effect of the SE model on students' PA. Nevertheless, in these studies, SR were not used to carry out the formal competition phase of the SE. Additionally, while in the previous studies (e.g., Coolkens et al., 2018; Wallhead et al., 2010), the students' PA levels were not measured by instruments which quantify PA intensity objectively; thus, there is no strong evidence for how these interventions contributed objective data to the daily recommended MVPA. Kelly et al. (2016) pointed out that the use of valid and reliable measures of PA are needed to understand the impact of PA lifestyle interventions. Unlike other commonly used instruments, such as pedometers or self-reported methods (Sternfeld & Goldman-Rosas, 2012), accelerometers have shown to be highly reliable and valid monitors for assessing PA and SB among school-aged children (Trost et al., 2011). Additionally, accelerometers have the advantage of assessing different dimensions of PA accurately (e.g., frequency, intensity or time) (O'Neill et al., 2017; Strath et al., 2013) with a reasonable compromise between ease of administration and cost (Troiano, 2007). In this context, according to Bornstein et al. (2011), an accelerometer-based measure of PA would contribute to the advancement of scientific knowledge regarding previous studies, since it is considered the reference method for measuring children's PA in free-living conditions.

As a result of the findings from these previous studies, although the SE model has been extensively researched in scientific literature, this study proposes a new format of SE including PE lessons and the extracurricular time of SR. Since the PE subject has a very limited time allocation in the educational system (Casado-Robles, Guijarro-Romero, & Mayorga-Vega, 2020), this study expects to transfer in-class activities in PE to other contexts of daily life (e.g., SR) in order to increase the quantity of PA opportunities, with the ultimate goal of promoting adherence to PA practice. Specifically, this study was designed to address how the engagement in a SE season during PE lessons could promote adolescents' voluntary participation in organized competitions during SR and its consequent influence in students' overall daily MVPA levels. Therefore, the main aim of the present study was to compare the effect of a SE-based irregular teaching unit on the PA levels of high-school students during SR. Secondarily, the present study was aimed at: (a) evaluating the effect of a SE-based irregular teaching unit in PE on the retention of students' PA levels in the SR following the intervention, and (b) examining the effect of the program on students' intention to take part in sporting competitions during SR organized by the PE teacher and by the students themselves.

Method

Design

As the present study was performed in a school setting, the intervention was focused on natural high-school classes (i.e., instead of isolated partipants), using a cluster-randomized controlled trial (Dreyhaupt et al., 2017). However, the students had been previously assigned randomly into classes by the school center. Six established classes, balanced by grade, were randomly assigned to one of the following groups: Control group (non-intervention), Traditional SE group (SE with competitions in PE), or the Irregular SE group (SE with competitions in SR). The

Ethical Committee of the University of Granada approved the protocol of the present study.

Participants and setting

A sample of 165 high-school students (88 boys and 77 girls) aged 13-16 years old (average age = 14.0 ± 1.1 years old) from six different eighth/tenth-grade PE classes of a state high-school center participated in the present study. The inclusion criteria for students were: (a) being enrolled in the eighth or tenth grades of the secondary educational level (classes where the school approvals were obtained); (b) being free from any health disorder which would make them unable to undergo PE; (c) presenting written signed consent of their parents or legal guardians, and (d) presenting written signed assent of the students. The exclusion criteria were: (a) not attending at least 10 PE classes during the intervention period, and (b) not having at least two SR with the required wear time recorded in each measured moment (i.e., pre-intervention, competitions and follow-up). Finally, although all of the 165 invited students agreed to participate and met the inclusion criteria, only 91 (46 boys and 45 girls) passed the exclusion criteria. The loss of 74 students resulted from: six students by the exclusion criterion (a), 61 students by the exclusion criterion (b), and seven students by both.

Table 1. General characteristics of the included participants[a]

	CG (*n* = 24)	TG (*n* = 22)	IG (*n* = 45)	*p*[b]
Age (13-14/15-16)	54.2/45.8	63.6/36.4	53.3/46.7	0.711
Grade (2nd/4th)	54.2/45.8	63.6/36.6	53.3/46.7	0.711
Gender (boys/girls)	62.5/37.5	50.0/50.0	44.4/55.6	0.360
Body mass (kg)	55.0 (8.1)	56.8 (13.7)	57.1 (8.2)	0.668
Body height (cm)	162.8 (8.3)	162.1 (6.0)	163.8 (7.4)	0.659
Body mass index (kg/m^2)	20.8 (3.0)	21.5 (4.4)	21.3 (2.6)	0.716
Physically active (days/week)	3.1 (1.5)	2.6 (1.2)	2.9 (1.6)	0.529
Extracurricular sport activities (hours/week)	2.6 (2.1)	3.1 (2.0)	2.5 (2.0)	0.578

Note. CG = Control Group; TG = Traditional Sport Education Group; IG = Irregular Sport Education Group; PA = Physical activity

[a]Data are reported as mean (standard deviation) or percentage. [b]Significance level from the chi-squared test for the categorical variables and from the one-way analysis of variance for continous variables.

The school center was situated in the urban area of the city of Granada and all the students came from middle-to-high income households. According to the school schedule, all students received two 60-minute PE lessons per week and 30-minute SR every school day. Table 1 shows the general characteristics of the included participants. The three groups were balanced in terms of age, grade and gender ($p > 0.05$); and there were no differences in terms of body mass, body height, body mass index, and PA baseline values between the groups ($p > 0.05$). The PE teachers were two men and one woman who had taught at least two seasons of the SE model (Siedentop et al., 2011) before this study.

Intervention programs

The intervention programs carried out by each group is shown in Figure 1 and Table 2. The SE seasons (i.e., Irregular SE and Traditional SE groups) were designed considering Hastie & Casey (2014) recommendations and according to the five phases proposed by Siedentop et al. (2011): (a) Introductory phase, where the SE model and its main characteristics were explained (e.g., phases or students' roles) and homogeneous teams were created according to their initial level. Each class was divided into four-student teams; (b) Teacher-directed phase, where the teacher proposed technical-tactical skills games and tasks for the whole class. Moreover, the students' roles were assigned and the observational items were elaborated; (c) Pre-season phase, where the students assumed their roles (i.e. coach, physical trainer, material assistant or captain). The instructions were given through the student-coach to his teammates after planning and consensus with the teacher. The teacher role was to assist the coaches during the lessons; (d) Season phase, where matches were held according to the competition schedule and one duty-team assumed referee and chronicler functions; and (e) final event, where the season-closing was held with the participation in friendly-matches and the awards ceremony prepared by the students in a festive atmosphere. The control intervention (i.e., control group) was based on direct instruction methodology (Metzler, 2017) and small-sided games. The main characteristics of the control intervention were: (a) teacher as protagonist of the teaching-learning process by issuing a model of proven performance and well-defined; (b) the student's role is to receive and reproduce that model as accurately as possible; (c) all the decisions are taken by the teacher (objectives, contents, activities and evaluation); (d) the organization is formal and not completely individualized.

Table 2. Detailed schematic of the intervention programs

Lesson	Irregular Sport Education	Traditional Sport Education
1	- Theoretical explanation of the Sport Education Model, its methodology and students' roles. - Introduction to the sport (Soccer/Basketball) - Circuit level test by stations, to assess initial technical-tactical skills. Station 1: Control and manage the ball 1 vs. 1; Station Free kicks/shots; Station 3: Small-sided games 3 vs. 3; Station 4: Pig-in-the-middle game.	
2	- Announcement of teams chosen according to students' initial level. - Assignation, by student preferences, their roles and responsibilities (i.e., coach, physical trainer, material assistant and capt - Whole class skill practice: Techniques and tactical skills tasks directed by the teacher about ball dribbling and passing gam Including explanations about technical skills and rules.	
3	- Elaboration of the observational items. - Distinctive team elements selection (i.e., name selection, presentation of logos and t-shirts colors). - Whole class skill practice: Techniques and tactical skills tasks directed by the teacher about dynamic passing/reception and stationary/dynamic shooting exercises. Including explanations about technical skills and rules. - Introduce game play 3 vs. 3	
4	***Physical education** - Team practice (First contact according to their roles). - Ball control and passing/reception tasks. - Intra-team games (2 vs. 1). Introduce duty-team (Role of the referee). ***School recess** - Regular season (League: 3 vs. 3). 2 rounds of matches. - Duty-team (one referee and one scorekeeper to each match)	- Team practice (First contact according t their roles). - Ball control and passing/reception tasks. - Intra-team games (2 vs. 1).
5	***Physical education** - Team practice. - Marking/Unmarking tasks. - Intra-team games (2 vs. 1 / 1 vs. 1). Introduce duty-team (Role of the referee). ***School recess** - Regular season (League: 3 vs. 3). 2 rounds of matches. - Duty-team (one referee and one scorekeeper to each match)	- Team practice. - Marking/Unmarking tasks. - Intra-team games (2 vs. 1 / 1 vs. 1).
6	***Physical education** - Chronicle and observations analysis about previous competitions (*Fair-play*). - Team practice. - Dribbling and shooting tasks. - Intra-team games (1 vs. 1 / 2 vs. 2). ***School recess** - Regular season (League: 3 vs. 3). 2 rounds of matches. - Duty-team (one referee and one scorekeeper to each match)	- Team practice. - Dribbling and shooting tasks. - Intra-team games (1 vs. 1 / 2 vs. 2). Introduce duty-team (Role of the referee)
7	***Physical education** - Chronicle and observations analysis about previous competitions (*Fair-play*). - Team practice. - Offence/Defense tactics. - Intra-team games (2 vs. 2).	- Team practice. - Offence/Defense tactics. - Intra-team games (2 vs. 2). Introduce du team (Role of the referee).

	***School recess** - Regular season (League: 3 vs. 3). 2 rounds of matches. - Duty-team (one referee and one scorekeeper to each match)	
8	***Physical education** - Chronicle and observations analysis about previous competitions (*Fair-play*). - Team practice. - Training for the championship. Tasks chosen by the coach and fitness trainer. - Intra-team games (2 vs. 2). ***School recess** - Regular season (League: 3 vs. 3). 2 rounds of matches. - Duty-team (one referee and one scorekeeper to each match)	- Regular season (League: 3 vs. 3). 4 roun of matches. - Duty-team (two referees, one scorekeep and one reporter to each match)
9	***Physical education** - Chronicle and observations analysis about previous competitions (*Fair-play*). - Team practice. - Training for the championship. Tasks chosen by the coach and fitness trainer. - Friendly inter-team games (3 vs. 3) ***School recess** - Regular season (Semi-final: 3 vs. 3). 2 rounds of matches. - Duty-team (one referee and one scorekeeper to each match)	- Regular season (League: 3 vs. 3). 4 roun of matches. - Duty-team (two referees, one scorekeep and one reporter to each match)
10	***Physical education** - Chronicle and observations analysis about previous competitions (*Fair-play*). - Team practice. - Training for the championship. Tasks chosen by the coach and fitness trainer. - Intra-team games (2 vs. 2). ***School recess** - Regular season (Semi-final: 3 vs. 3). 2 rounds of matches. - Duty-team (one referee and one scorekeeper to each match)	- Regular season (Semi-final: 3 vs. 3). 4 rounds of matches. - Duty-team (two referees, one scorekeep and one reporter to each match)
11	***Physical education** - Chronicle and observations analysis about previous competitions (*Fair-play*). - Team practice. - Training for the championship. Tasks chosen by the coach and fitness trainer. - Friendly inter-team games (3 vs. 3) ***School recess** - Regular season (Final: 3 vs. 3). - Duty-team (one referee and one scorekeeper to each match)	- Regular season (Final: 3 vs. 3). - Duty-team (two referees, one scorekeep and one reporter to each match)
12	- Friendly championship games (3 vs. 3 / 5 vs. 5) - Awards ceremony prepared by the students in a festive atmosphere	

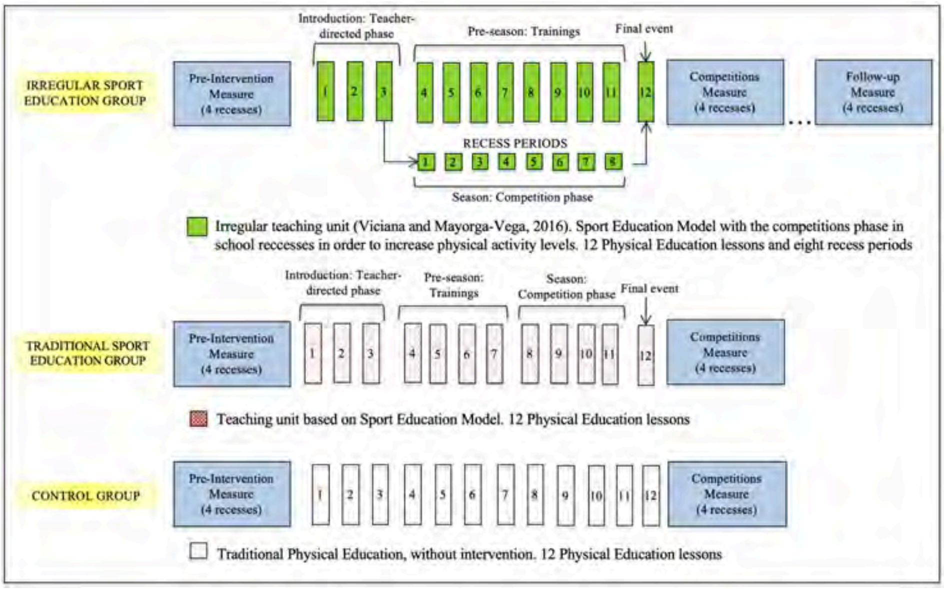


Figure 1. Intervention programs. Sport Education-based irregular teaching unit versus Traditional Sport Education-based teaching unit and Control teaching unit (adapted from Viciana & Mayorga-Vega, 2016)

Irregular Sport Education season

An irregular teaching unit (Viciana & Mayorga-Vega, 2016) based on SE methodology was applied to the Irregular SE group during 12 PE classes (two sessions a week, lasting 60 minutes each) to carry out the whole SE model except the formal competition phase which was carried out during eight SR (two SR a week, lasting 30 minutes each). The scheme of this Irregular SE season with the key features of each phase are detailed in Table 2. The teaching unit structure was distributed into one introductory PE class, two teacher-directed PE classes, eight pre-season PE classes, eight formal competition recesses, and one final event PE class. This irregular distribution has three differentiating characteristics with the traditional SE model: (a) During the pre-season phase, the first 15 minutes of the sessions were dedicated to analyze the chronicles and observations registered by the duty-team in the previous competition; leaving the rest of the session to train for the next match by the team themselves; (b) all competitions were held in SR, where teams played two rounds of matches each day supported by a rotating duty team. Before students went out to the playground, the teacher put out the calendar of matches, assigning the duty-team functions and play zones. Subsequently, the first round of 10-minute matches was performed, the duty-team observation sheets were collected, and the next round of matches was started (including the duty-team rotation). In this way, all participants were involved in PA during SR; and (c) the voluntary duty team distributed its components to perform support for two different matches (i.e., one referee and one scorekeeper to each match and one student observing and supporting both matches).

Traditional Sport Education season

The teaching unit applied to the Traditional SE group was based on the SE model during 12 PE classes (two 60-minute sessions a week). In contrast with the Irregular SE group, the methodology was applied according to the original model, that is, all phases developed in the PE lessons, without the implication of SR during the formal competition phase. The scheme of this Traditional SE season with the key features of each phase are also detailed in Table 2. The teaching unit structure was distributed into one introductory lesson, two teacher-directed lessons, four pre-season lessons, four formal competition lessons, and one final event lesson. Moreover, during SR students had access to all playground facilities but without any additional play material or any organized activities.

Control intervention

The Control group followed a traditional intervention of sport in PE (12 PE classes, two 60-minute sessions a week), mixing tasks with small-sided games and direct instruction methodologies (Metzler, 2017). All sessions were oriented to the technical-tactical learning of sports with this structure: a warm-up of 5-10 minutes with aerobic exercises and games; a 40-minutes main part consisting of three-to-five technique-centered tasks and concluding with small-sided games (3 vs. 3); and a cool-down of 5 minutes with common flexibility tasks. As well as the Traditional SE group, students were not provided with any additional play material or any organized activities during SR.

Intervention program fidelity

Sport Education fidelity

Hastie and Casey (2014) guidelines were followed to ensure the fidelity of the intervention based on the SE model in the Irregular SE and Traditional SE groups. The main limitation arises with the absence of students' previous experience with the model. However, all teachers had previous experience in the SE model and they received three training sessions focused on the characteristics of the programs and the week-by-week plan for the season was designed. Regarding the curricular description of the intervention, attending to the proposed recommendations: (a) the model was implemented over a prolonged period of time (i.e., 12 PE lessons and eight SR in the Irregular SE group); (b) students maintained their four-student teams throughout the whole season; (c) competitions were based on small-sided games adapted to the educational level (i.e., 3 vs. 3); (d) students adopted different roles and their corresponding responsibilities (i.e., coach, physical trainer, material assistant and captain), and (e) the whole season was celebrated with a festive atmosphere. Moreover, an experienced teacher oversaw the application of the programs based on the SE model following the 19-item benchmarks proposed by Ko et al. (2006).

Control intervention fidelity

Metzler's (2017) guidelines were followed to ensure the fidelity of the intervention programs based on direct instruction in the control group intervention. PE teachers attended a workshop where the intervention program and the direct instruction's benchmarks were fully explained and the week-by-week plan was designed. Additionally, an experienced teacher oversaw the application of the program using the seven-category checklist proposed by Metzler (2017).

Data collection

Descriptive characteristics

At the beginning of the study, the anthropometric measures (i.e., participants' body mass, body height and body mass index) were registered following the international standard for anthropometric assessment (Stewart et al., 2011). Moreover, adolescents' gender, grade, and age information were obtained from the school reports. Finally, about PA habits, the number of days per week that students were involved in PA for at least 60 minutes was measured by the PACE questionnaire for adolescents (Prochaska et al., 2001). While the question of PA from the enKID study (Serra-Majen & Rodríguez-Santos, 2003) was used to assess the weekly hours that students spend in extracurricular sports activities.

Objectively-measured physical activity

GT3X accelerometers (ActiGraph, LLC, Pensacola, FL, USA) were used to assess students' objective PA levels. Using the same instruments and protocols, evaluation was carried out in two different moments to examine differences between the three groups: (a) Pre-intervention, during four SRs the week before beginning the intervention; and (b) Competitions, during the eight SRs corresponding to the formal competition phase of the Irregular SE group. Moreover, students of the Irregular SE group had an extra evaluation (follow-up), during four SRs two weeks after the end of the intervention program to examine possible retention of PA levels at SR. The pre-intervention and follow-up measures consists of four SRs for practical reasons, since accelerometers were fitted on Monday and collected on Friday in the SR with four-valid SRs (i.e., From Monday to Thursday). The accelerometers were fitted on the participants' right hip using an elastic waistband, and all data were downloaded and analyzed using the ActiLife Lifestyle Monitoring System Software version 6.13.3, following the protocols that have shown greater evidence of validity for adolescents (Migueles et al., 2017). Since short bursts of rapidly changing activity characterize adolescents' behavior, a one-second *epoch* was used (Cain et al., 2013). According to the school schedule, SR activity was recorded from 11:15 to 11:45 (i.e., 30 minutes was the length of SR in this school, as in most Spanish school centers) in all measurement periods (i.e., pre-intervention, competitions and follow-up). Evenson's cut-off points (Evenson et al., 2008) were used to determine sedentary (≤100 *counts*/min), light (101-2295 *counts*/min), moderate (2296-4011 *counts*/min), and vigorous PA (≥4012 *counts*/min) during the 30 minutes of SR. According to the cross-validation study performed by Trost et al. (2011), these cut-off points are the most valid for estimating PA intensity with one-second *epochs* among adolescents. To avoid potential bias due to differences in the length of total time of valid wear time, time-based standardized scores were used (i.e., all the variables were expressed in terms relative to the valid wear time recorded) (Mayorga-Vega et al., 2018).

Intention to continue participating in sports competitions at school recess

After the intervention program, intention to continue to take part in following sporting competitions organized in SR was evaluated by self-report with two *ad hoc* items: "Would you participate in sports competitions at SR organized by the PE teacher?" and "Would you participate in sports competitions at SR organized by the students themselves" These items were preceded by the sentence: "Regarding your intention to participate in sports competitions at SR…". A Likert-type scale from 1 to 10 (from "strongly disagree" to "strongly agree") was used.

Data analysis

Means and standard deviation/standard error or percentage for general characteristics of the participants and dependent variables were calculated. Statistical tests assumptions were checked by common procedures (e.g., histograms and Q-Q plots for normality). First, as exploratory analyses, the chi-squared analyses for categorical variables (i.e., age, grade and gender), and the one-way analyses of variance (ANOVA) for continuous variables (i.e., body mass, body height, body mass index, and pre-intervention PA levels) were conducted to examine potential differences between the three groups. Afterward, the effect of a SE-based irregular teaching unit on students' PA levels during SR was examined. Because the unit of randomization and intervention was the class, a Mixed Multilevel Linear Model with participants nested within classes was selected (i.e., repeated-measures nested ANOVA) (Li et al., 2017). Additionally, the intra-class correlation coefficients (ICC, also known as intra-cluster correlation coefficients) of some variables showed empirical evidence of dependency of the observations (variables in Table 3: ICC = 0.34-0.43, Wald Z, $p < 0.05$; variables in Table 4: ICC = 0.00-0.00, Wald Z, $p > 0.05$). According to Field's (2017) recommendation, the approach started from "basic" models in which all the parameters were fixed and then progressively random coefficients and exploring confounding variables were followed. The – 2 log-likelihood (- 2LL, i.e., comparing the change in the chi-square test) was used to compare the model´s fit. From all the potential confounding variables explored (i.e., age, grade, gender, body mass, body height, body mass index, and intervention attendance), gender was used for the analyses reported in Table 3 (no covariables were used for analyses reported in Table 4). The maximum likelihood estimation method was used. After the Mixed Multilevel Linear Model, for the *post-hoc* analyses, the within-group pairwise comparisons (for each group independently) with the Bonferroni adjustment was used. Effect sizes were estimated using the Cohen's ∂ for pairwise comparisons. Finally, the Kruskal-Wallis test was used to assess the effect of the intervention program on students' intention to take part in sport competitions during SR. All statistical analyses were performed using the SPSS version 25.0 for Windows (IBM® SPSS® Statistics). The statistical significance level was set at $p < 0.05$.

Results

Table 3 shows the effect of the intervention program on objective PA levels during SR. The results of the two-way repeated-measures nested ANOVA showed a statistically significant interaction effect in all measured variables ($p < 0.001$). Then, the within-group pairwise comparisons showed that the Irregular SE group students statistically significantly improved objective PA levels during SR ($p < 0.001$). However, for both the Control group and Traditional SE group no statistically significant differences were found during SR ($p > 0.05$), except for the Control group in the MVPA variable that decreased statistically ($p < 0.05$).

Table 4 shows the effect of the intervention program and retention of PA during SR on Irregular SE group students' objective PA levels during SR. The results of the one-way repeated-measures nested ANOVA on the objective PA levels showed a statistically significant effect in all measured variables ($p < 0.001$). Later, the within-group pairwise comparisons showed that students had a statistically significant improvement in objective PA levels ($p < 0.001$). However, the follow up measure showed statistically decreases in objective PA levels during SR ($p < 0.001$) and statistical significant differences were not found with the pre-intervention value ($p > 0.05$).

With respect to the effect of the intervention program on students' intention to take part in sport competitions during SR, the results of the Kruskal-Wallis test did not show statistically significant differences between the three groups for both those organized by the PE teacher [median (interquartile range); Control group = 8.0 (3.0), Traditional SE group = 6.5 (5.8), Irregular SE group = 6.0 (5.0)] nor those organized by students themselves [median (interquartile range); Control group = 7.0 (5.0), Traditional SE group = 5.5 (6.3), Irregular SE group = 5.0 (6.5)] ($p > 0.05$). The students' participation rate in sports competitions organized in SR, which were completely voluntary, was 100% for the Irregular SE group students.

Table 3. Effect of the intervention program on students' objective physical activity levels during school recess

	Pre-intervention	Competitions	Mixed Multilevel Lineal Model			Effect sizes	
	Mean (SE)	Mean (SE)	*-2LL*	*F*	*p*	Comp.	*d*[a]
Sedentary (%)							
CG (*n* = 24)	62.8 (3.0)	69.2 (3.0)	1492.53	30.65	< 0.001	CG-TG	-0.23
TG (*n* = 22)	71.1 (3.1)	73.2 (3.1)				CG-IG	-1.81
IG (*n* = 45)	64.5 (2.2)	37.2 (2.2)***				TG-IG	-1.58
Light (%)							
CG (*n* = 24)	24.2 (2.2)	22.0 (2.2)	1368.32	15.20	< 0.001	CG-TG	0.16
TG (*n* = 22)	19.8 (2.2)	19.5 (2.2)				CG-IG	1.36
IG (*n* = 45)	22.2 (1.6)	36.2 (1.6)***				TG-IG	1.20
Moderate-to-vigorous (%)							
CG (*n* = 24)	13.0 (1.6)	8.9 (1.6)*	1254.80	31.15	< 0.001	CG-TG	0.26
TG (*n* = 22)	9.2 (1.6)	7.3 (1.6)				CG-IG	1.97
IG (*n* = 45)	13.7 (1.2)	26.6 (1.2)***				TG-IG	1.71
Light-to-vigorous (%)							
CG (*n* = 24)	37.2 (3.0)	30.8 (3.0)	1492.53	30.65	< 0.001	CG-TG	0.23
TG (*n* = 22)	29.0 (3.1)	26.8 (3.1)				CG-IG	1.81
IG (*n* = 45)	36.0 (2.2)	62.8 (2.2)***				TG-IG	1.58
Vertical axis (counts/min)							
CG (*n* = 24)	820.8 (97.8)	651.1 (97.8)	2762.59	51.76	< 0.001	CG-TG	0.34
TG (*n* = 22)	620.1 (99.9)	626.3 (99.9)				CG-IG	2.89
IG (*n* = 45)	790.1 (71.8)	2097.3 (71.8)**				TG-IG	2.54

Note. SE = Standard error; - 2LL = -2 log-likelihood; CG = Control Group; TG = Traditional Sport Education Group; IG = Irregular Sport Education Group.

[a] Cohen's *d* effect size.

* $p < 0.05$, *** $p < 0.001$ *post-hoc* pairwise comparisons with Bonferroni adjustment for within-groups analyses.

Table 4. Effect of the intervention program and follow-up on objective physical activity levels during school recess
Education group students (n = 31)[a]

	Pre-intervention (0)	Competitions (1)	*Follow up* (2)	Multilevel lineal mod		
	Mean (SE)	Mean (SE)	Mean (SE)	*- 2LL*	*F*	
Sedentary (%)	65.3 (3.1)	35.6 (1.8)*	70.4 (3.0)*	753.584	93.542	<
Light (%)	21.7 (1.8)	36.7 (1.3)*	21.0 (2.3)*	688.366	41.054	<
Moderate-to-vigorous (%)	13.0 (1.7)	27.7 (1.8)*	8.6 (1.1)*	654.464	52.998	<
Light-to-vigorous (%)	34.7 (3.1)	64.4 (1.8)*	29.6 (3.0)*	753.584	93.531	<
Vertical axis (*counts*/min)	777.5 (96.4)	2226.7 (127.3)*	642.6 (75.2)*	1433.894	68.848	<

Note. SD = Standard error.

[a]Due to the loss of valid data as follow-up, the number of participants is lower. [b]Cohen's ∂ effect size.

* $p < 0.001$ *Post-hoc* within-group pairwise comparisons with the Bonferroni adjustment (Column Competitions: Pre
Column *Follow up*: Competitions – *Follow up*).

Discussion

The main purpose of the present study was to examine the effect of a SE-based irregular teaching unit on the PA levels developed by high-school students in SR. The findings of the present study have shown that the students who performed the irregular teaching unit reached higher PA levels in SR during the intervention. The Irregular SE group participants showed an increase of 12.9% in MVPA compared with the baseline, obtaining a total of 26.6% of the SR involved in MVPA. If the light-to-vigorous PA is considered, this increase was even greater (26.8%), obtaining a total of 62.8% of the SR time (≈18 min) engaged in total PA.

Numerous studies focused on evaluating PA levels during SR among the school-aged population show an average percentage of 40% of SR time spent in MVPA (Parrish et al., 2013; Ridgers et al., 2012). However, almost all of them were carried out in elementary school children and it cannot be compared with adolescents due to the declines in PA levels, and specifically during SR (Ridgers et al., 2012), in the transition from elementary to high-school. There is very little evidence about SR MVPA levels in high-school students and even less if only those studies with objective evaluation instruments are considered (Reilly et al., 2016). From the eight similar studies found (i.e., in high-school adolescents and evaluated with objective instruments), the average of SR time where high-school students were involved in MVPA was 13.6% (from 4.4% to 52.0%) (e.g., Andersen et al., 2015; Viciana et al., 2016). It should be noted that, although these studies have been carried out in different countries, and SR length or culture were different, results were very similar with the baseline levels in the present study (i.e., 13.7%). Therefore, these results are far below the recommendations proposed by Bailey et al. (2012) and it is necessary to implement school-based policies on SR to increase MVPA levels and to contribute to achieving the adolescents' daily PA recommendations (CDC & Shape-America, 2017; WHO, 2018).

Although previous studies have carried out intervention programs designed to increase those MVPA levels during SR, there is limited scientific knowledge about interventions based on semi-structured sports activities in high-school students and objectively-measured as the present study. Most of them were focused on improving sports equipment or modifying school playgrounds with elementary students (Parrish et al., 2013), and these kinds of interventions may not be the most appropriate in adolescents. Moreover, the program proposed in the present study is simpler and easier to implement in any school center, without the need to invest in infrastructure/equipment or modifications at the playground surface as in previous studies (e.g., Huberty et al., 2011; Méndez-Giménez et al., 2017). Given the limited studies carried out with adolescents, it is difficult to highlight an ideal recess-based intervention to increase MVPA levels in this population. However, some similar intervention programs (i.e., games and competitions) obtained significant effects in elementary students. For instance, Erwin et al. (2013) proposed a program based on dance videos during five indoor SR, increasing MVPA levels up to 22.2% of SR time. Also, Larson et al. (2014) performed an intervention with games and semi-structured sports during four SR, increasing MVPA levels in 5.7% (∂ = 0.34) during structured SR. The present study obtained better improvements (i.e., 12.9%; ∂ = 1.97) than the previous one. These differences may be because of the students' baseline MVPA levels during unstructured SR. Students in Larson's et al. (2014) intervention, spent an average of 34% of unstructured SR time in MVPA while students in the present study spent just a 13.7%. Therefore, it is easier to improve MVPA levels in students with low baseline PA levels, than in those with relatively good baseline levels. Moreover, differences could also be due to a longer length of the intervention in the present study (i.e., eight vs. four SR).

Moreover, longer intervention programs seem to have a more significant effect on elementary school students' PA levels. In this line, Huberty et al. (2011) and Howe et al. (2012) applied an intervention based on structured sports activities and competitions organized by the teacher during SR during seven months and nine weeks, respectively. Authors obtained great improvements from 25.1% to 48% (increasing a 22.9%) and from 23% to 49.7% (increasing a 26,7%, ∂ = 3.02) of SR time involved in MVPA, respectively. These interventions overcome the Bailey et al. (2012) recommendations of 40% SR time in MVPA, but it should be noted that baseline PA levels were far higher than present baseline results (i.e., 13.8%), probably because of participants' age (i.e., elementary school). In addition to the program length and age, the present study also differs in the contents and structure of the program. For example, Huberty et al. (2011) proposed an intervention not only based on sports activities but also includes a more complex restructuring of the playground with activity zone maps, activity cards and extra sport equipment.

Moreover, in these above-mentioned studies, students were free to choose from a wide range of activities (e.g., individual/group activities, or playful games/competitions), while in the present study the sports competitions were linked to the content developed during PE lessons. Thus, it was based on traditional and widely practiced sports (i.e., soccer and basketball) and it was the same for all students. Maybe if the program is applied with alternative and less common sports, it could be more interesting for all adolescents and better results could be obtained. Furthermore, all activities offered to the students were led and supervised by the teacher. Meanwhile, the present intervention promoted PA allowing for participation in sports competitions voluntarily as part of the SE Model (Siedentop et al., 2011). It is important to highlight that students' autonomy as a key principle of the present study, since ordinary SR is non-curricular and free time for students to play or interact with friends without the intervention of teachers. Therefore, PE teachers should encourage students to perform PA autonomously, which hopefully they will then apply to being physically active during their leisure time without the help of teachers. However, it is possible that no more pretentious results could have been obtained due to the total autonomy of the students participating in sports competitions and the lack of students' choice of activities, in combination with the short length of the program. Regardless, these results are consistent with other studies indicating that offering multi-component structured programs at SR seems to be desirable, taking advantage of this extracurricular time as a useful complement to adolescents' habitual PA levels (Carson et al., 2014).

Regarding SE-based programs, a previous study with similar students' characteristics and length revealed increases in adolescents' MVPA and total PA levels during the formal competition phase (Perlman, 2012). Although this aforementioned study was developed in PE lessons, the formal competition phase seems to be an excellent strategy for increasing students' PA levels. However, only two previous studies were found that regard the effects of the SE model on PA levels developed during SR (Coolkens et al., 2018; Wallhead et al., 2010). Both studies observed that students spent a higher time of SR involved in PA while they performed a SE based program in PE. However, the approach was different because they offered the SR as extra opportunities for students to play the sport they were learning in PE, but they did not provide continuity of any phase of the model between both contexts as has been done in the present study with the competition phase. The use of an irregular teaching unit, which linked both settings, could be a good option for transferring the students' motivation in a SE season from PE lessons to an extracurricular-time context, such as SR. Moreover, the present study provides new evidence for the positive effect of a SE season on adolescents' participation in sport-based PA during SR measured with accelerometry instead of systematic observation (i.e., Coolkens et al., 2018) or pedometers (i.e., Wallhead et al., 2010).

A secondary purpose of the present study was to examine the effect of a SE-based irregular teaching unit in PE on the retention of students' PA levels in the following SR after the intervention. The results of the present study showed that the Irregular SE group adolescents' PA levels in SR decreased significantly, returning to the baseline levels (pre-intervention) after finishing the program. Therefore, it seems that if adolescents revert to depending on their habit to engage in MVPA during SR, the effects of a short-term SE program do not have a lasting influence. These findings are consistent with previous studies which determined that it is necessary to propose an appropriately designed and structured PA environment to increase adolescents' PA levels (in periods of time where doing PA is voluntary), as it was performed during the intervention program (Dobbins et al., 2013; Wallhead et al., 2014). Additionally, results are in line with Baquet et al. (2018) who determined the intervention had only an acute effect on the adolescents' PA levels, but once the "novelty effect" is lost and structured programs are no longer offered, adolescents spend most of their time in sedentary activities again. This theory should be tested with further voluntarily structured programs developed throughout the school year to check whether the increase happens only in the first weeks or if it remains for a desirably prolonged period with the aim of sustainable change in adolescents' PA levels during SR.

Another secondary purpose of this study was to examine the effect of the program on students' intention to take part in sporting competitions during SR organized by PE teachers and by the students themselves. It should be noted that 100% of the Irregular SE group students participated in the formal competition phase completely voluntarily. This high proportion of participation may be due to this direct connection between PE lessons and SR through the SE model. As part of the SE model, students feel like members of a team and they develop a feeling of affiliation (Siedentop et al., 2011), which could make them more interested in participating in sports competitions with their peers putting into practice everything learned in PE lessons and being encouraged by other schoolmates during SR. These results also complement those found by previous studies. For instance, Coolkens et al. (2018) also reported a high rate of participation (i.e., almost 80%) in structured parkour SR sessions. Moreover, Martínez de Ojeda (2015) obtained a participation rate of 100% of students in organized training sessions at SR while they were performing a SE program, compared to 0% while following a traditional methodology. Nevertheless, this difference is reduced if the sports competition phase is considered, where 100% of the students participated at SR while they followed the SE program, and 91.1% of them while they developed a traditional methodology. It can be understood as a positive attitude towards sports competitions in SR (slightly higher while students experience a SE session).

Furthermore, results showed that after the intervention, there were no differences in adolescents' intention to participate in sports competitions at SR between the three groups. The lack of differences could be explained due to the participants of Control group and Traditional SE group experienced the sport competitions celebrated by the Irregular SE group at SR (the three groups belonged to the same school and shared the playground). Therefore, these students followed the matches and encouraged other classes classmates, sometimes even asking whether they could participate in competitions that could be organized later. These findings are very positive because they mean that despite being an extracurricular period, most students (independently of the intervention group) were interested in participating in sport-based activities during SR, showing that these programs are feasible and in accord with students' interests. For this reason, the design of practical and viable sports competitions during recess periods could be an excellent strategy to promote PA levels.

Considering the global results obtained in the present study and their relation with the objectives (i.e., the success of this program incrementing the students' PA levels and the time involved in PA during SR, but exclusively for the program duration, as well as the students' desire of participating in this type of SE competition activities connected with the PE program, even when they belonged to the control group), and to avoid losing the gain obtained, it is desirable to convert these short-term SR programs into whole year programs. This could even allow connecting the PE subject with other curricular subjects such as Maths (statistics from the SE competitions), Ethics and Citizenship education (analyzing the referee protests or conflicts among students and respect for game rules), getting the whole school to participate in the SR program. This is related to the effectiveness of previous programs with similar characteristics, although not developed in SR, but in any moment of the school time (e.g., "The Daily Mile" program, which consists of 15 minutes doing PA at any intensity lead by a teacher, was established in Scotland in one primary school in 2012, and now is being followed by over 1000 schools worldwide, Marchant et al., 2020). It is also important to denote that the increment of students' health is the final goal of the kind of programs developed during SR periods. Therefore, connecting the SR periods with organized sports and other physical activities, lead by teachers, and connected with other subjects could be a solution for the success of these kinds of programs in order to be developed for the whole year and implicate the maximum number of participants (teachers and students).

Regarding the limitations of the present study, the use of a convenience sample could limit the generalization of the findings, although this type of intervention must be complemented with others to conclude major generalizations. Moreover, only the effect of an Irregular SE program based on soccer and basketball was studied, and future research should study these effects with other teaching contents. The program length (12 PE lessons) could also be a limitation to achieving more outstanding results. Although shorter scientific experiences have been found in previous scientific literature (Chu & Zang, 2018), the present program was designed according to the creator of the model recommendations, who established a minimum length of 12 sessions (Siedentop et al., 2011). Furthermore, some PA domains, such as the context or students' intentionality, cannot be measured by accelerometry. However, the accelerometer-based measures are considered the most appropriate reference method for measuring PA levels in free-living conditions (Bornstein et al., 2011) with great validity and reliability (Troiano, 2007). Besides that, because our measurement periods were SR, all MVPA presumably was composed of games for developing motor skills or sports. Regarding the strengths of the present study, to our knowledge, this is the first study that applies the irregular teaching unit model proposed by Viciana and Mayorga-Vega (2016) concerning increasing students' MVPA levels during SR. Thus, the findings of the present study give empirical support for the effectiveness of this innovative teaching unit. Furthermore, the accelerometer-measured PA reveals objective data about the quantity and intensity of PA, taking a step ahead in regard to studies carried out with self-reported measures or pedometers. The use of a research design that compares three groups (i.e., control, traditional SE, and irregular SE groups) also adds quality to the study. It allows us to control that results are due to the new methodology used and not to possible external factors.

Future research studies are suggested for overcoming these limitations and verifying the present findings in several other contexts (i.e., sports modalities, older and earlier ages of participants and/or a more extended intervention). These future programs should include explicit strategies of PA promotion (e.g., advice about PA benefits and problems related with sedentary behavior, goal setting based on international guidelines, strategies to reduce sedentary behavior and the increase of PA or resolution of barriers to PA practice) which make students feel that they are making an informed decision about their health (Dobbins et al., 2013; Jago et al., 2015). Because of this, students may be more likely to change their behavior in order to pursue long-term improvements in PA levels and not only during the intervention. For example, these orientations could be included through the season phase, sharing the time spent in fair-play analysis during PE lessons with this counselling time. Additionally, future research studies focused on monitoring habitual PA during a full week could be interesting for checking the contribution of the program on adolescents' habitual PA levels and not only during the SR, in order to transfer the improvements from the school context to students' daily life (Viciana & Mayorga-Vega, 2018).

In conclusion, the application of an irregular teaching unit in PE following the SE is an appropriate intervention in order to offer the opportunity for high-school students to participate in sports competitions during SR, and to increase their PA levels in this time period. This knowledge could help PE teachers to design effective and feasible programs that allow for the increase of adolescents' PA levels during SR, contributing to the daily adolescents' PA recommendation and their consequent benefits on adolescents' health.

Acknowledgements

Authors want to thank you to all the students and the Physical Education teachers who made this research possible. We also thank Aliisa Hatten for the English revision of the manuscript.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

Carolina Casado-Robles and Santiago Guijarro-Romero are supported by a research grant from the Spanish Ministry of Universities [grant number: FPU16/03314 and FPU15/02387, respectively]. Carolina Casado-Robles is also supported by a "Research Initiation grant" from the University of Granada.

References

Andersen, H. B., Klinker, C. D., Toftager, M., Pawlowski, C. S., & Schipperijn, J. (2015). Objectively measured differences in physical activity in five types of schoolyard area. *Landscape and Urban Planning*, *134*, 83-92. https://doi.org/10.1016/j.landurbplan.2014.10.005

Bailey, D. P., Fairclough, S. J., Savory, L. A., Denton, S. J., Pang, D., Deane, C. S., & Kerr, C. J. (2012). Accelerometry-assessed sedentary behaviour and physical activity levels during the segmented school day in 10-14-year-old children: The HAPPY study. *European Journal of Pediatrics, 171*(12), 1805-1813. https://doi.org/10.1007/s00431-012-1827-0

Baquet, G., Aucouturier, J., Gamelin, F. X., & Berthoin, S. (2018). Longitudinal follow-up of physical activity during school recess: impact of playground markings. *Frontiers in Public Health*, *6*, 283. https://doi.org/10.3389/fpubh.2018.00283

Bornstein, D. B., Beets, M. W., Byun, W., & McIver, K. (2011). Accelerometer-derived physical activity levels of preschoolers: a meta-analysis. *Journal of Science and Medicine in Sport*, *14*(6), 504-511. https://doi.org/10.1016/j.jsams.2011.05.007

Cain, K. L., Sallis, J. F., Conway, T. L., Van Dyck, D., & Calhoon, L. (2013). Using accelerometers in youth physical activity studies: a review of methods. *Journal of Physical Activity and Health, 10,* 437-450. https://doi.org/10.1123/jpah.10.3.437

Carson, R. L., Castelli, D. M., Beighle, A., & Erwin, H. (2014). School-based physical activity promotion: A conceptual framework for research and practice. *Childhood Obesity, 10*(2), 100-106. https://doi.org/10.1089/chi.2013.0134

Casado-Robles, C., Guijarro-Romero, S., & Mayorga-Vega, D. (2020). Planificación en Educación Física mediante unidades didácticas innovadoras para incrementar los niveles de actividad física habitual de los escolares. En Alonso, G., Romero, J.M., Rodríguez-Jiménez, C., & Sola, J.M. (Ed.), Investigación, Innovación docente y TIC. Nuevos Horizontes Educativos (pp. 283-296). Madrid, España: Dykinson.

Centers for Disease Control and Prevention & SHAPE America. (2017). S*trategies for recess in schools*. Atlanta, GA: Centers for Disease Control and Prevention.

Chu, T. L., & Zhang, T. (2018). Motivational processes in Sport Education programs among high school students: A systematic review. *European Physical Education Review, 24,* 372-394. https://doi.org/10.1177/1356336X17751231

Coolkens, R., Ward, P., Seghers, J., & Iserbyt, P. (2018). Effects of generalization of engagement in Parkour from Physical Education to recess on physical activity. *Research Quarterly for Exercise and Sport, 89,* 429-439. https://doi.org/10.1080/02701367.2018.1521912

Dobbins, M., Husson, H., DeCorby, K., & LaRocca, R. L. (2013). School-based physical activity programs for promoting physical activity and fitness in children and adolescents aged 6 to 18. *Cochrane Database of Systematic Reviews, 2*. https://doi.org/10.1002/14651858.CD007651.pub2

Dreyhaupt, J., Mayer, B., Keis, O., Öchsner, W., & Muche, R. (2017). Cluster-randomized Studies in Educational Research: Principles and Methodological Aspects. *GMS Journal for Medical Education, 34*(2), 1-25.

Erwin, H. E., Koufoudakis, R., & Beighle, A. (2013). Children's physical activity levels during indoor recess dance videos. *Journal of School Health, 83*, 322-327. https://doi.org/10.1111/josh.12034

Evenson, K. R., Catellier, D. J., Gill, K., Ondrak, K. S., & McMurray, R. G. (2008). Calibration of two objective measures of physical activity for children. *Journal of Sports Sciences, 26*, 1557-1565. https://doi.org/10.1080/02640410802334196

Field, A. (2017). *Discovering statistics using IBM SPSS Statistics* (5th ed.). London: SAGE Publications.

Hastie, P. A., & Casey, A. (2014). Fidelity in models-based practice research in sport pedagogy: A guide for future investigations. *Journal of Teaching in Physical Education*, *33*, 422-431. https://doi.org/10.1123/jtpe.2013-0141

Howe, C. A., Freedson, P. S., Alhassan, S., Feldman, H. A., & Osganian, S. K. (2012). A recess intervention to promote moderate-to-vigorous physical activity. *Pediatric Obesity, 7*, 82-88. https://doi.org/10.1111/j.2047-6310.2011.00007.x

Huberty, J. L., Siahpush, M., Beighle, A., Fuhrmeister, E., Silva, P., & Welk, G. (2011). Ready for recess: A pilot study to increase physical activity in elementary school children. *Journal of School Health, 81*, 251-257. https://doi.org/10.1111/j.1746-1561.2011.00591.x

Institute of Medicine. (2013). *Educating the student body: Taking physical activity and physical education to school*. National Academies Press.

Jago, R., Rawlins, E., Kipping, R. R., Wells, S., Chittleborough, C., Peters, T. J., ... & Campbell, R. (2015). Lessons learned from the AFLY5 RCT process evaluation: implications for the design of physical activity and nutrition interventions in schools. *BMC public health*, *15*(1), 946. https://doi.org/10.1186/s12889-015-2293-1

Kelly, P., Fitzsimons, C., & Baker, G. (2016). Should we reframe how we think about physical activity and sedentary behaviour measurement? Validity and reliability reconsidered. *International Journal of Behavioral Nutrition and Physical Activity, 13*(1), 32. https://doi.org/10.1186/s12966-016-0351-4

Ko, B., Wallhead, T., & Ward, P. (2006). Professional development workshops—What do teachers learn and use? *Journal of Teaching in Physical Education, 25*, 397-412. https://doi.org/10.1123/jtpe.25.4.397

Kriemler, S., Meyer, U., Martin, E., van Sluijs, E. M., Andersen, L. B., & Martin, B. W. (2011). Effect of school-based interventions on physical activity and fitness in children and adolescents: a review of reviews and systematic update. *British Journal of Sports Medicine, 45*, 923-930. https://doi.org/10.1136/bjsports-2011-090186

Larson, J. N., Brusseau, T. A., Chase, B., Heinemann, A., & Hannon, J. C. (2014). Youth physical activity and enjoyment during semi-structured versus unstructured school recess. *Open Journal of Preventive Medicine, 4*, 631-639. https://doi.org/10.4236/ojpm.2014.48072

Li, W., Xiang, P., Chen, Y., & Xie, X. (2017). Unit of analysis: Impact of Silverman and Solmon's article on field-based intervention research in Physical Education in the U.S.A. *Journal of Teaching in Physical Education*, *36*, 131–141. https://doi.org/10.1123/jtpe.2016-0169

Marchant, E., Todd, C., Stratton, G., Brophy, S. (2020). The Daily Mile: Whole-school recommendations for implementation and sustainability. A mixed-methods study. *PLoS ONE 15*(2), e0228149. https://doi.org/10.1371/journal.pone.0228149

Martínez de Ojeda D. (2015). Student participation in physical activity in leisure time depending on the methodology used in the Physical Education. *Actividad Física y Deporte: Ciencia y Profesión, 22*, 13-25.

Mayorga-Vega, D., Martínez-Baena, A., & Viciana, J. (2018). Does school physical education really contribute to accelerometer-measured daily physical activity and non sedentary behaviour in high school students? *Journal of Sports Sciences, 36*(17), 1913–1922. https://doi.org/10.1080/02640414.2018.1425967

Méndez-Giménez, A., Cecchini, J. A., & Fernéndez-Rio, J. (2017). The effect of a self-constructed material on children's physical activity during recess. *Saúde Pública, 51*, 58.

Metzler, M. (2017). *Instructional models for Physical Education*. Scottsdale, AZ: Routledge.

Migueles, J. H., Cadenas-Sanchez, C., Ekelund, U., Delisle Nystr€om, C., Mora-Gonzalez, J., Löf, M., Labayen, I., Ruiz, J. R., & Ortega, F. B. (2017). Accelerometer data collection and processing criteria to assess physical activity and other outcomes: a systematic review and practical considerations. *Sports medicine, 47*(9), 1821-1845. https://doi.org/10.1007/s40279-017-0716-0

O'Neill, B., McDonough, S. M., Wilson, J. J., Bradbury, I., Hayes, K., Kirk, A., … & Tully, M. A. (2017). Comparing accelerometer, pedometer and a questionnaire for measuring physical activity in bronchiectasis: a validity and feasibility study? *Respiratory research*, *18*(1), 16.

Parrish, A. M., Okely, A. D., Stanley, R. M., & Ridgers, N. D. (2013). The effect of school recess interventions on physical activity: A systematic review. *Sports Medicine, 43*, 287-299. https://doi.org/10.1007/s40279-013-0024-2

Perlman, D. (2012). The influence of the Sport Education Model on amotivated students' in-class physical activity. *European Physical Education Review, 18*, 335-345. https://doi.org/10.1177/1356336X12450795

Prochaska, J. J., Sallis, J. F., & Long, B. (2001). A physical activity screening measure for use with adolescents in primary care. *Archives of pediatrics & adolescent medicine*, *155*(5), 554-559.

Reilly, J. J., Johnston, G., McIntosh, S., & Martin, A. (2016). Contribution of school recess to daily physical activity: systematic review and evidence appraisal. *Health Behavior and Policy Review*, *3*, 581-589. https://doi.org/10.14485/HBPR.3.6.7

Ridgers, N. D., Salmon, J., Parrish, A. M., Stanley, R. M., & Okely, A. D. (2012). Physical activity during school recess: A systematic review. *American Journal of Preventive Medicine, 43*, 320-328. https://doi.org/10.1016/j.amepre.2012.05.019

Serra-Majem, L., & Rodríguez-Santos, F. (2003). Crecimiento y desarrollo. *Estudio enKid. Barcelona: Editorial Masson*,

Siedentop, D., Hastie, P. A., & van der Mars, H. (2011). *Complete guide to Sport Education*. (2nd Ed.) Champaign, IL: Human Kinetics.

Sternfeld, B., & Goldman-Rosas, L. (2012). A systematic approach to selecting an appropriate measure of self-reported physical activity or sedentary behavior. *Journal of Physical Activity & Health, 9*(S1), S19–28.

Stewart A, Marfell-Jones M, Olds T, De Ridder J. (2011). *International Standards for Anthropometric Assessment*. New Zealand: International Society for the Advancement of Kinanthropometry.

Strath, S. J., Kaminsky, L. A., Ainsworth, B. E., Ekelund, U., Freedson, P. S., Gary, R. A., Richardson, C. R., Smith, D. T., & Swartz, A. M, American Heart Association Physical Activity Committee of the Council on Lifestyle and Cardiometabolic Health and Cardiovascular, Exercise, Cardiac Rehabilitation and Prevention Committee of the Council on Clinical Cardiology, and Council. (2013). Guide to the assessment of physical activity: clinical and research applications: a scientific statement from the American Heart Association. *Circulation*, *128*(20), 2259-2279. https://doi.org/10.1161/01.cir.0000435708.67487.da

Troiano, R. P. (2007). Large-scale applications of accelerometers: new frontiers and new questions. *Medicine & Science in Sports & Exercise*, *39*(9), 1501. https://doi.org/10.1097/mss.0b013e318150d42e

Trost, S. G., Loprinzi, P. D., Moore, R., & Pfeiffer, K. A. (2011). Comparison of accelerometer cut points for predicting activity intensity in youth. *Medicine & Science in Sports & Exercise, 43*, 1360-1368. https://doi.org/10.1249/MSS.0b013e318206476e

Viciana, J., & Mayorga-Vega, D. (2016). Innovative teaching units applied to Physical Education–changing the curriculum management for authentic outcomes. *Kinesiology, 48*(1), 142–152. https://doi.org/10.26582/k.48.1.1

Viciana, J., Mayorga-Vega, D., & Martínez-Baena, A. (2016). Moderate-to-vigorous physical activity levels in physical education, school recess, and after-school time: influence of gender, age, and weight status. *Journal of physical activity and health, 13*(10), 1117–1123. https://doi.org/10.1123/jpah.2015-0537

Viciana, J., & Mayorga-Vega, D. (2018). The three-axes model of planning in physical education. *Retos, 33*, 313–319. https://doi.org/10.47197/retos.v0i33.54533

Wallhead, T. L., Garn, A. C., & Vidoni, C. (2014). Effect of a Sport Education program on motivation for Physical Education and leisure-time physical activity. *Research Quarterly for Exercise and Sport, 85,* 478-487. https://doi.org/10.1080/02701367.2014.961051

Wallhead, T. L., Hagger, M., & Smith, D. T. (2010). Sport Education and extracurricular sport participation: An examination using the Trans-Contextual Model of motivation. *Research Quarterly for Exercise and Sport, 81*, 442-455.

World Health Organization. (2010). *Global recommendations on physical activity for health*. Geneva: World Health Organization.

World Health Organization. (2018). *Promoting physical activity in the education sector*. Copenhagen: World Health Organization.

Paper IV

IS IT MORE FUN TO DO PHYSICAL EDUCATION OUTSIDE THE SCHOOL CENTER?

Casado-Robles, C., Mayorga-Vega, D., Guijarro-Romero, S., & Viciana, J.

The Journal of Educational Research

Under review

Paper IV

ABSTRACT

Objective: To examine the effect of a Physical Education-based inside-outside alternated teaching unit in pupils' enjoyment levels in Physical Education.

Methods: One hundred seventy-nine pupils (94 females) aged 13-15 years old were cluster-randomly assigned to the Alternated group (four-week alternated teaching unit, with one Physical Education lesson inside and one outside the school each week) or Inside group (four-week inside teaching unit, only with Physical Education lessons inside the school). Pupils' enjoyment towards Physical Education subject was assessed before and after the intervention, as well as at the end of each Physical Education lesson of the teaching unit.

Results: Results showed that the alternated teaching unit improved pupils' enjoyment in Physical Education subject. Moreover, pupils in the Alternated group reported higher enjoyment values during the outside Physical Education sessions compared to their inside sessions.

Conclusions: These findings may guide the design of Physical Education-based programs enjoyable for pupils to promote their physical activity levels.

KEYWORDS

Intervention program. Enjoyment. Innovative teaching units. Secondary Education. Adolescents. Sport Satisfaction Instrument.

Introduction

Despite the regular practice of physical activity (PA) is considered a powerful health indicator in school-aged children (World Health Organization, WHO, 2020), and its well-recognized benefits (e.g., reduce the likelihood of obesity and diabetes, or improvements in academic achievement) (US Department of Health and Human Services, 2018; WHO, 2020), worldwide, about 80% of school-aged children do not meet the daily recommendation of 60 minutes of moderate-to-vigorous PA (Guthold et al., 2020). Moreover, previous studies have demonstrated a decline in PA levels as age progresses (Corder et al., 2017), in favor of more time spent in sedentary behavior (Pearson et al., 2017). Therefore, nowadays the promotion of PA practice among school-aged children is a health priority (World Health Organization, 2018). For this reason, Physical Education subject has been recognized as an ideal setting to encourage pupils to be physically active in both in-school and out-of-school contexts, promoting the acquisition of lifelong healthy PA habits (World Health Organization, 2018). Furthermore, this acquisition is also included as a primary aim to achieve throughout the Secondary Education stage in the Physical Education subject in most countries (e.g., American Association of Health, Physical Education, Recreation and Dance, 2014; Spanish Ministry of Education, Culture and Sport, 2015).

Besides, it should be considered that Ryan & Deci (2020) highlighted pupils' enjoyment as a relevant factor influencing their intrinsic motivation, which is of great importance since without motivation learning is not possible in the educational context (Rehman & Haider, 2013). Therefore, pupils' enjoyment in carrying out educational tasks is key in leading them to learning achievement. The term enjoyment is widely understood as a positive physiological state, which involves feelings of pleasure and fun related to Physical Education practice (Scanlan & Simons, 1992). Moreover, to achieve an effective healthy habits promotion from the Physical Education setting, the enjoyment of Physical Education is a primary determinant that influences the school-aged children's willingness to participate in PA (Bai et al., 2018). Specifically, the pupils' enjoyment during Physical Education lessons would affect their PA levels in out-of-school contexts, tending to look for more opportunities to be physically active in out-of-school contexts those with higher levels of enjoyment during Physical Education lessons (Cox et al., 2008). These correlations suggest that Physical Education teachers should try to achieve high levels of enjoyment in school-aged children during Physical Education lessons in order to increase their PA levels during leisure time (Yli-Piipari et al., 2013). Unfortunately, as well as the decline observed in PA levels, previous longitudinal studies observed a progressive decrease in pupils' enjoyment levels during Physical Education lessons as they grow up (e.g., Hashim, 2007; Barkoukis et al., 2010; Yli-Piipari et al., 2013). Therefore, the design and implementation of enriching and motivating experiences in Physical Education that address the improvement of pupils' enjoyment levels are needed to achieve satisfactory learning and, ultimately, to increase their PA practice during leisure time effectively (Baena-Extremera, Granero-Gallegos, Bracho-Amador et al., 2012; Cox et al., 2008; Dishman et al., 2005).

In this line, some previous intervention studies have been carried out to improve school-aged children's enjoyment in different Physical Education-based teaching units (TU). On the one hand, some studies were focused on assessing the pupils' enjoyment according to the teaching content. For example, Carriedo (2019) found schoolchildren's high levels of enjoyment during Physical Education lessons when implementing a TU based on innovative content (i.e., combat sports). However, these findings do not agree with those found by Hortiguela et al. (2017), who did not find differences in pupils' enjoyment comparing a combat sport TU and two teaching units of popular sports such as football and basketball. Therefore, the relationship between the content and pupils' enjoyment is not completely clear. On the other hand, previous findings suggest that beyond the content developed, it seems necessary to include specific strategies in those TU, such as the design of activities and objectives adapted to pupils' level or creating opportunities for decision-making, to affect positively the pupils' enjoyment during Physical Education (e.g., Sevil et al., 2016). Besides that, previous studies included another interesting option to favorably improve school-aged children's enjoyment during Physical Education lessons based on developing all the lessons of the TU in the close environment (i.e., outside of the school center) (Baena-Extremera & Granero-Gallegos, 2015; Baena-Extremera, Granero-Gallegos & Ortíz-Camacho, 2012)

Regarding the inclusion of these outside Physical Education lessons, it could also provide pupils with authentic performances that they can apply to their leisure time, achieving the principle of transferability of learning from Physical Education classes to pupils' daily life competencies (Viciana & Mayorga-Vega, 2018), and leading ultimately to lifelong outside PA participation (Institute for Outdoor Learning, 2020). Therefore, in order to design effective and authentic interventions in Physical Education context, Viciana & Mayorga-Vega (2016) proposed the "Innovative TU" to better organize learning/teaching periods. In particular, the alternated TU were designed to link the learning developed by two complementary contents, thus allowing students to understand the fundamentals that connect them, and reaching authentic and meaningful outcomes. Specifically, the implementation of a Physical Education-based alternated TU (Viciana & Mayorga-Vega, 2016) which combines possibilities of PA practice in the school center (inside) and out of the school center (outside), could allow Physical Education teachers to connect the learning developed in the classroom to the pupils' life, while also seeking to increase the pupils' levels of enjoyment.

Following the evidence showing that school-aged children's enjoyment of Physical Education and PA levels decline during the Secondary Education stage (Corder et al., 2017; Yli-Piipari et al., 2013), the secondary education stage seems to be an ideal target context for Physical Education TU to prevent these declines. Unfortunately, to our knowledge, no previous studies are examining the effect of a Physical Education-based alternated TU on the possibilities of practicing PA inside-outside on the pupils' enjoyment. Consequently, the present study pursued two main aims: (a) to examine the effect of an alternated TU doing lessons inside-outside the school center on pupils' enjoyment levels during Physical Education lessons; and (b) to compare the effect of the inside and outside lessons on pupils' enjoyment during today's Physical Education lesson.

Method

Study design

The present study is reported according to the CONSORT for cluster randomized trials guidelines (Campbell et al., 2012). The Ethical Committee for human studies of the University of Granada (452/CEIH/2017) approved the present study protocol. The protocol conforms to the Declaration of Helsinki statements (64th WMA, Brazil, October 2013). A cluster-randomized controlled trial design was used for practical reasons and due to the nature of the present study (i.e., natural groups from an educational setting) (Dreyhaupt et al., 2017). Recruitment was carried out in December 2018, and data collection was done from January to May 2019.

Participants

The principal and the Physical Education teachers of a public high school center (chosen by convenience) of the province of Granada (Spain) were consulted for the implementation of the study. They were informed about the project, and permission to conduct the study was requested. After school approvals were obtained, all the 179 pupils aged 13-15 years old (94 females and 85 males), from the second to the third grade of Secondary Education in the Spanish education system (i.e., eighth to ninth-grade in the United States education system), were invited to participate in the present study. Pupils and their legal tutors were completely informed about the features of the study. Pupils' signed written informed assent and their legal guardians' signed written informed consent were obtained before taking part in the study. According to the center's reports, all the pupils' families had a middle socioeconomic level.

The inclusion criteria were: a) being enrolled in the second to third grade at the Secondary Education level (grades in which approval of the school was obtained); b) participating in the normal Physical Education classes; c) being free of any health disorder that would make them unable to engage in PA normally; d) presenting the corresponding signed consent by their legal guardians, and e) presenting the corresponding signed assent by the pupils. The exclusion criterion was not having performed the dependent variables correctly following their rules of administration.

Sample size

A priori sample size calculation was estimated with the Optimal Design Plus Empirical Evidence Software Version 3.01 for Windows. Parameters were set as follows: significance level $\alpha = 0.05$, number of participants per cluster $n = 25$, effect size $\delta = 0.50$, intra-cluster correlation coefficient $\rho = 0.01$, and statistical power $(1 - \beta) = 0.80$. A total number of six clusters (150 participants) was estimated.

Randomization

Randomization was conducted at the class level, using a computerized random number generator. Before the pre-intervention evaluation was administered, six natural classes of the selected school center (four second and two third-grade classes), balanced by grade (i.e., two second-grade and one third-grade classes per group), were randomly assigned by an independent researcher, blinded to the study objectives to form one of the following study groups: Inside group (n = 88) or Alternated group (n = 91). Moreover, according to the education rules, the pupils had been previously assigned randomly into classes by the school center, following the criterion that the classes should be balanced between males and females, at the beginning of the school year.

Intervention

Pupils in the Alternated group carried out a TU of eight Physical Education lessons (i.e., a four-week TU, two lessons per week) based on the alternated TU proposed by Viciana & Mayorga-Vega (2016) for the practice of PA inside and outside the school center. Specifically, it consists of four Physical Education lessons for PA practice inside and four Physical Education lessons for PA practice outside, delivering one inside lesson (i.e., inside the school center) followed by one outside lesson (i.e., in the immediate environment near the school center), which allow establishing a transference of learning from the conventional Physical Education classroom inside the school center to outside contexts. Table 1 shows the Alternated teaching unit scheme.

Table 1. Scheme of the Alternated teaching unit.

Lesson	Inside/Outside	Physical fitness contents	Sport and games contents
1	Inside	Total body circuit-training workout in pairs (three levels of difficulty)	Baseball or kickball games (pupils chose one of them).
2	Outside		
3	Inside	CrossFit teens competitions in teams of 3-4 players.	Indoor football or handball games (pupils chose one of them).
4	Outside		
5	Inside	High interval intensity training through relay games in teams of 3-4 players.	Ultimate or American handball games (pupils chose one of them)
6	Outside		
7	Inside	8-stations gymkhana about physical fitness tasks and sports developed in the teaching unit, in teams of 3-4 players, using an augmented reality app (HP-Reveal app).	
8	Outside		

The useful time for each lesson was 55 minutes divided into a 10-minute warm-up, 40-minute main part, and 5-minute cool-down. Throughout all the lessons, physical fitness tasks (e.g., high-intensity interval training methods) and sports (e.g., volleyball) were developed in the main part of the lessons. Those contents were selected since they are the most broad-spread globally in Physical Education (Hardman et al., 2014) and could be the most applicable to students' free time (individually or in small groups with friends). The main difference between inside and outside lessons was the material and the space used in each environment. In this way, the inside lessons were developed using the traditional Physical Education materials (e.g., gymnastic benches, medicine balls, vaulting boxes, wall bars, or mats) and conventional spaces within the educational center (e.g., indoor gym and sports court). On the other hand, the outside lessons were developed taking advantage of the spaces and material resources offered by the immediate environment. Specifically, one natural park and another urban park with their resources (e.g., trees, swings, benches, and other urban furniture), and the municipal sports center which had several sports courts (e.g., a football field or paddle courts) were used.

The teaching methodology included strategies to increase pupils' enjoyment during Physical Education classes trying to maximize their involvement, stimulation, and excitement based on Hashim's (2007) recommendations. Specifically, these strategies were to: (a) offer a wide range of involvement choices with the same goal in order to individualize according to pupils' preferences (e.g., tactical attack principles choosing between indoor football or handball) (b) match school-aged children's needs and skills, individualizing the tasks according to their level and guiding the task as challenging (e.g., three different levels of difficulty in physical fitness tasks); (c) propose small-group activities in which the opportunity to choose activity partners during lessons were given to pupils (e.g., the election of three or four peers for CrossFit competitions); and (d) include variety and innovative contents and resources which lead to perceive the task as exciting (e.g., HP-Reveal mobile app).

On the other hand, pupils in the Inside group also carried out a TU of eight Physical Education lessons (i.e., a four-week TU, two lessons per week). The useful class time, the structure of the lesson divide into three parts, and tasks were the same as the Alternated group. However, in contrast to the Alternated group, pupils only received the lessons inside the school center. Therefore, the eight Physical Education lessons were developed using conventional Physical Education materials and school facilities within the educational center. Unlike the Alternated group, no transference of learning from the Physical Education context to out-of-school contexts nor PA practice in the immediate environment was promoted in the Inside group.

Measures

All evaluations were carried out during the Physical Education class time by the same researchers, instruments, and protocols. The pupils filled out the test and questionnaires in an ordinary classroom in quiet and comfortable conditions. Pupils were asked for their maximum sincerity, and they were guaranteed the confidentiality of the obtained data. Although instructions on how to correctly respond to the questionnaires were explained before starting, the main researcher was present during the whole evaluation session in order to solve any question that might emerge. The measurement protocol followed with each variable is detailed below.

Demographic characteristics

Participants' age, grade, and gender data were registered at the beginning of the study.

Anthropometric measurement

Pupils' body mass and body height were assessed at the beginning of the study (twice each one and the mean was retained) following the International Standards for Anthropometric Assessment (Stewart et al., 2011). Then, body mass index was calculated as body mass divided by the square of body height (kg/m^2). Finally, participants' body weight status was categorized by the body mass index thresholds as overweight/obesity (i.e., gender- and age-adjusted cut-point values equal to or higher than the equivalent value of 25 kg/m^2 at the age of 18 years) or non-overweight/obesity (i.e., lower than the equivalent cut-point values of 25 kg/m^2 at the age of 18 years) (Cole et al., 2000).

Extracurricular physical activity

The weekly hours that pupils spend practicing PA after-school were assessed with the enKid questionnaire at the beginning of the study (Martínez-Gómez et al., 2009). It consists of a single question (i.e., How many hours do you spend weekly in extracurricular sports activities?). A 6-point Likert-type scale, ranging from "0 hours/week" to "More than 5 hours/week" was used. This questionnaire showed an adequate convergent validity comparing with accelerometer-measured total PA and moderate-to-vigorous PA in Spanish students (r = 0.43 and 0.46, respectively) (Martínez-Gómez et al., 2009).

Enjoyment in Physical Education subject

The pupils' intrinsic enjoyment towards Physical Education subject was measured during a Physical Education class time at the beginning and another one at the end of the intervention (i.e., pre-intervention and post-intervention, respectively), using a short version of the Sport Satisfaction Instrument (SSI-Physical Education, Baena-Extremera, Granero-Gallegos, Bracho-Amador et al., 2012) proposed by Casado-Robles et al. (in press). This questionnaire measures the pupils' enjoyment dimension using items like "I usually have fun in Physical Education lessons". The scale was preceded by the sentence "Indicate your degree of disagreement or agreement with the following statements, referring to your Physical Education lessons…". According to previous studies (e.g., Casado-Robles et al., 2020), in order to facilitate the evaluation of the items by the students, a 10-point Likert-type scale, ranging from 1 ("Strongly disagree") to 10 ("Strongly agree") was used since is the scale that Spanish school-aged children receive in their scholar marks. The Spanish version of the adapted SSI-Physical Education has shown adequate psychometric properties among school-aged children (GFI = 1.00; CFI = 1.00; RMSEA = 0.00; Ω = 0.92; Cronbach's alpha = 0.80) (Casado-Robles et al., in press).

Enjoyment in today's Physical Education lesson

The pupils' intrinsic enjoyment towards today's Physical Education lessons was measured at the end of each Physical Education lesson of the intervention (i.e., during the five-minute cool-down). The above-mentioned Spanish version of the SSI-Physical Education (Baena-Extremera, Granero-Gallegos, Bracho-Amador et al., 2012) was adapted by Casado-Robles et al. (in press). The original questionnaire was referred to Physical Education subject in general, then, in this version, the items were slightly modified to refer to the lesson that had just finished. This questionnaire measures the pupils' enjoyment dimension using items like "In today's Physical Education lesson I had fun" instead of the original "I usually have fun in Physical Education lessons". Moreover, the sentence which preceded the scale was also modified by "Indicate your degree of disagreement or agreement with the following statements, referring to today's Physical Education lesson…". A 10-point Likert-type scale, ranging from 1 ("Strongly disagree") to 10 ("Strongly agree") was used. Afterward, the pupils' enjoyment in today's Physical Education lesson was calculated as the mean value of at least two valid lessons in each condition (i.e., inside and outside lessons). The Spanish version of SSI for today's Physical Education lesson has shown adequate psychometric properties among school-aged children (GFI = 1.00; CFI = 1.00; RMSEA = 0.00; Ω = 0.92; Cronbach's alpha = 0.88) (Casado-Robles et al., in press).

Statistical analysis

Means and standard deviation/standard error or percentage for general characteristics of the participants and dependent variables were calculated. Statistical tests assumptions were checked by common procedures (e.g., histograms and Q-Q plots for normality). The internal consistency of the dependent variables measured by the questionnaires was examined with Cronbach's alpha. Alpha values < 0.40 were considered poor, 0.40–0.59 fair, 0.60–0.74 good, and 0.75–1.00 excellent (Hernaez, 2015). First, as exploratory analyses, the chi-squared analyses for categorical variables (i.e., gender and grade), and the one-way analyses of variance (ANOVA) for continuous variables (i.e., age, body mass, body height, body mass index, extra-curricular PA, and attendance rate) were conducted to examine potential differences between the two groups. Afterward, the effect of a Physical Education-based alternated TU on pupils' enjoyment in Physical Education subject was examined. An intention-to-treat approach was followed, that's means, all the participants were included in the statistical analyses regardless of adherence to the protocol. Because the unit of randomization and intervention was the class, a Multilevel Linear Model with participants nested within classes and measures within participants was selected (i.e., two-way mixed nested ANOVA) (Li et al., 2017). The maximum likelihood estimation method was used. According to Field's (2017) recommendation, the approach started from "basic" models in which all the parameters were fixed and then progressively random coefficients and exploring confounding variables were followed. The – 2 log-likelihood (- 2LL, i.e., comparing the change in the chi-square test) was used to compare the model's fit. From all the potential confounding variables explored (i.e., gender, age, grade, body mass, body height, body mass index, extra-curricular PA, and intervention attendance), no covariables were used. Finally, for the *post-hoc* analyses, the within-group pairwise comparisons with the Bonferroni adjustment were used. Similarly, regarding the effect of a Physical Education-based alternated TU on pupils' enjoyment during today's Physical Education lesson, a Multilevel Linear Model with participants nested within classes was selected (i.e., one-way nested ANOVA). From all the potential confounding variables explored, no covariables were used (except age for the inside-outside comparison). Effect sizes were estimated using Cohen's d for pairwise comparisons. Finally, although in the present study an intention-to-treat approach was followed, as sensitivity analyses, all the above-mentioned analyses were also carried out with a per-protocol approach (i.e., including only the participants with an adherence to the protocol ≥ 87.5%, that is, seven lessons). All statistical analyses were performed using the SPSS version 25.0 for Windows (IBM® SPSS® Statistics). The statistical significance level was set at $p < 0.05$.

Results

Final sample and general characteristics

Figure 1 shows the flow chart corresponding to the participants included in the present study. Although all the invited 179 pupils (94 females and 85 males) agreed to participate and met the inclusion criteria, the number of students that satisfactorily passed the exclusion criterion to become the final sample group was different between pupils' intrinsic enjoyment in Physical Education subject (n = 169) and today's Physical Education lesson (n = 174). All the participants that did not follow the protocol were because they did not have an attendance rate equal to the 100%. No participants were lost because of the rejection to continue in the study or change of the school. Supplementary Figure 1 shows the flow chart for the sensitivity analysis (i.e., per-protocol approach).

Table 2 shows the general characteristics of the participants. The results of the one-way ANOVA and the chi-square test did not show statistically significant differences in terms of general characteristics and attendance rate between the two groups ($p > 0.05$). In the sample of the present study, the Cronbach's alpha values for enjoyment during Physical Education subject was 0.86, and enjoyment during today's Physical Education lesson was 0.89.

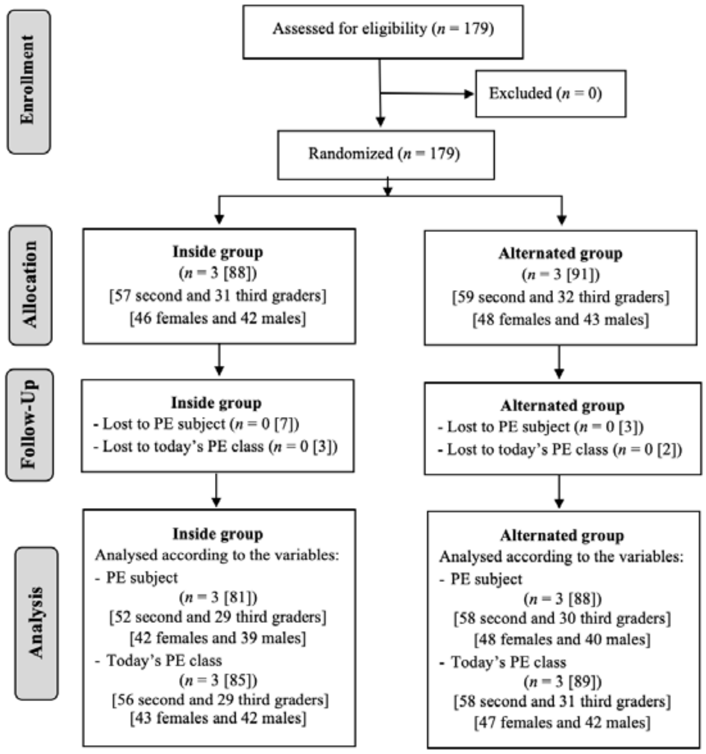


Figure 1. Flow chart of the school classes and students of the present study. All numbers are reported as school classes [students]. PE = Physical Education.

Table 2. General characteristics of the participants and differences between the two groups.

	Total	Alternated	Inside	ANOVA/Chi-squared test	
	(*N* = 179)	(*n* = 91)	(*n* = 88)	F/χ^2	*p*
Age (years)	14.08 (0.95)	14.03 (0.81)	14.13 (1.07)	0.532	0.467
Gender (females/males)	52.50/47.50	52.70/47.30	52.30/47.70	0.004	0.949
Grade (2nd/3rd)	64.80/35.20	64.80/35.20	64.80/35.20	0.000	0.993
Body mass (kg)	59.76 (13.37)	58.39 (11.73)	61.19 (14.82)	1.966	0.163
Body height (cm)	164.26 (8.18)	164.08 (7.98)	164.45 (8.43)	0.091	0.736
Body mass index (kg/m^2)	22.05 (4.13)	21.60 (3.60)	22.53 (4.60)	2.241	0.136
Overweight-obese (no/yes)	67.40/32.60	69.20/30.80	65.50/34.50	0.279	0.597
Extra-curricular PA (hours/week)	2.30 (2.15)	2.41 (1.97)	2.18 (2.33)	0.495	0.483
Intervention rate	90.43 (14.99)	91.48 (13.93)	89.35 (16.03)	0.908	0.342

Note. Continous variables (i.e., age, body mass, body height, body mass index, and extracurricular PA) are reported as mean (standard deviation) and categorical variables (i.e., gender, grade, overweight-obese and intervention rate) as percentage. PA = Physical activity.

[a] One-way analysis of variance for continous variables and the chi squared test for categorical variables.

Enjoyment in Physical Education subject

Table 3 shows the effect of the inside-outside alternated TU on enjoyment in Physical Education subject. The MLM results on enjoyment showed a statistically significant interaction effect ($p < 0.001$). Subsequently, the within-group pairwise comparisons showed that the Alternated group participants improved statistically significant their enjoyment from the pre- to the post-intervention measure ($p < 0.01$; $\partial = 0.65$). However, for the Inside group statistically significant differences were not found ($p > 0.05$). The sensitivity analysis (i.e., per-protocol approach) found the same outcomes as the main analysis (i.e., intention-to-treat approach) (Supplementary Table 1).

Table 3. Effect of the inside-outside alternated teaching unit on enjoyment in Physical Education subject.

Group	Pre-intervention	Post-intervention	Multilevel Lineal Model			ES
	M (SE)	M (SE)	- 2LL	*F*	*p*	∂
Alternated	8.1 (0.2)	8.8 (0.2)*	1317.08	11.510	0.001	0.65
Inside	8.8 (0.2)	8.4 (0.2)				

ES = Effect size; M = Mean; SE = Standard error; - 2LL = -2 log-likelihood; ∂ = Cohen's ∂ effect size; Alternated, n = 88, Inside, n = 81.

* $p < 0.01$, *post-hoc* pairwise comparisons with Bonferroni adjustment for within-groups analyses.

Enjoyment in today's Physical Education lesson

Table 4 shows the effect of the inside-outside alternated TU on enjoyment during today's Physical Education lesson. The MLM results showed that the Alternated group participants reported statistically significant higher enjoyment values during the outside sessions than during their inside sessions ($\partial = 0.46$), as well as compared to the Inside group participants ($\partial = 0.52$) ($p < 0.05$). However, for the comparison between Alternated inside sessions and the Inside group, statistically significant effects were not found ($p > 0.05$). The sensitivity analysis (i.e., per-protocol approach) found the same outcomes as the main analysis (i.e., intention-to-treat approach), except that no statistically significant differences in enjoyment values between the Alternated group participants during their outside sessions and the Inside group participants were found, although there is a tendency towards statistical significance being $p < 0.10$ ($p = 0.062$ vs. $p = 0.039$) (Supplementary Table 2).

Table 4. Effect of the inside-outside alternated teaching unit on enjoyment during today's Physical Education lesson.

Group	Inside	Outside	Multilevel Lineal Model				ES
	M (SE)	M (SE)	Comparison	- 2LL	*F*	*p*	*∂*
Alternated	8.5 (0.1)	9.0 (0.1)	AG-IG Inside	599.19	0.231	0.632	0.07
Inside	8.4 (0.1)	-	AG-IG Outside	564.89	6.985	0.039	0.52
			AG In-outside	515.85	31.272	< 0.001	0.46

ES = Effect size; M = Mean; SE = Standard error; - 2LL = -2 log-likelihood; *∂* = Cohen's *∂* effect size; Alternated, *n* = 89, Inside, *n* = 85. AG = Alternated group; IG = Inside group.

Discussion

The purposes of the present study were: (a) to examine the effect of an alternated TU doing lessons inside-outside the school center on pupils' enjoyment levels during Physical Education lessons; and (b) to compare the effect of the inside and outside lessons on pupils' enjoyment during today's Physical Education lesson. The findings of the present study showed that the alternated TU improved pupils' enjoyment during Physical Education lessons compared with the inside TU. Moreover, regarding the comparison of the pupils' enjoyment during each Physical Education lesson, pupils in the Alternated group reported higher enjoyment values during their outside sessions than during their inside sessions across the TU, as well as higher enjoyment values than the Inside group participants during the inside sessions.

The success of the present TU to increase school-aged children's enjoyment may be possibly due to the teaching methodology and evidence-based strategies known to maximize students' involvement, stimulation, and excitement (Briggs, 1994; Hashim, 2007). Among others strategies: (1) the opportunity for pupils to engage in a wide variety of activities (Michael et al., 2016); (2) the provision of choice between different activities with the same objective considering the current pupils' interests (Granero-Gallegos et al., 2014); (3) the inclusion of interactions and cooperation with peers (i.e., small-group activities) (Dishman et al., 2005); or (4) the individualization of tasks to maximize feelings of pupils' ability or competence according to the degree of challenge (Lubans et al., 2017) may foster their enjoyment in Physical Education lessons and facilitate lifelong PA habits. These results agree with those found by Sevil et al. (2016) and Abós et al. (2017), which also increase pupils' enjoyment levels after the implementation of a Physical Education-based intervention based on specific motivational strategies in comparison with the control group, although with all their lessons inside the school center. Furthermore, boredom in the Physical Education classroom tends to appear when the lessons are monotonous and repetitive (White et al., 2018), so the novelty of the TU may be another influential factor in improving enjoyment since pupils are tired of always doing the same activities (Baena-Extremera & Granero-Gallegos, 2013). This novelty is reflected in the innovative contents and resources included (e.g., CrossFit-teens or mobile apps) (Cummiskey, 2011) and the use of outside spaces for some Physical Education lessons, which are very different from the traditional Physical Education to which all students are accustomed.

Moreover, the outside public spaces are also classified as the most favorite places for PA practice by school-aged children (Rehrer et al., 2011), so conducting Physical Education lessons in these spaces could be conducive to the enjoyment of its practice more than inside the school center. In this line, these improvements in students' enjoyment during outside lessons support previous findings studies (e.g., Mackenzie et al., 2018; Baena-Extremera & Granero-Gallegos, 2013; Baena-Extremera & Granero-Gallegos, 2015) which indicate that the realization of experiences outside the conventional Physical Education classroom supposes an increase in the student's interest and enjoyment. However, the present study represents an advance over previous knowledge because it obtains improvements in pupils' enjoyment during outside Physical Education lessons, but also an effect of the whole alternated TU on pupils' enjoyment using only half of the outside sessions than the previous ones which developed all their lessons outside. In other words, a combination of inside-outside sessions such as the one proposed also improves pupils' enjoyment and could reduce some disadvantages of outside TU such as, the teachers fear and concerns about students' safety outside the school grounds, the limited and strict institution timetables or adverse weather conditions that disadvantage to undertake all the lessons outside (Dyment, 2005; Ernst, 2014). Furthermore, previous studies differ from the present one in the contents developed during the TU. The previous outdoor studies were based on adventure activities (e.g., climb, rappel, zip line, or skiing), which are one of the less common activities included in Physical Education (Hardman et al., 2014), while the present study proposes the development of those contents most broad-spread globally in Physical Education (i.e., physical fitness tasks and individual/team sports) (Hardman et al., 2014).

Regarding this transference from Physical Education lessons to their daily life, following the Granero-Gallegos et al. (2012) suggestion, the environmental context of the school (e.g., extracurricular activities, green zones, or municipal sports centers) should be analyzed previous the study. In this way, the content covered during Physical Education lessons could be continued by school-aged children in their free time, as many hours as possible. Moreover, the contents used in the present study could be the most applicable to students' free time (individually or in small groups with friends), without the need for special equipment or a specialist's supervision like previous studies (e.g., adventure activities). Therefore, the present alternated TU provides students authentic and situational PA practices using the environment that surrounds the school center and their particular community, making pupils capable to practice PA autonomously in the out-of-school context (European Commission/EACEA/Eurydice, 2013; SHAPE America, 2013). Another advantage of connecting both contexts (i.e., inside the school center and the immediate environment) could be a better understanding of the relationship between what is developed in Physical Education lessons and how it can be applied to pupils' daily life (Viciana & Mayorga-Vega, 2018). It could make them realize that the way of working in both contexts is very similar and therefore, in the future, pupils could pay more attention to inside Physical Education lessons acquiring the necessary knowledge and feeling competent to apply it autonomously to their free time (Bandura, 2004).

Besides this transference of learning, it could also be facilitated that if pupils have higher feelings of enjoyment during Physical Education lessons outside the school center, this enjoyment could also be transferred to PA practice in their free time using the same environment and ultimately promoting the acquisition of healthy PA habits. In this line, several studies have suggested that if school-based programs are enjoyable for students, they may contribute to becoming more active during their daily life (Gråstén & Yli-Piipari, 2019), improving positively variables like the intention to be physically active in the future (e.g., Baños et al., 2019) or PA and sedentary habits (e.g., Bai et al., 2018; Yli-Piipari et al., 2013), so this TU may be suitable for promoting pupils' participation in PA. Besides that, significant international agencies, like the General Conference of UNESCO (2015) or the National Association for Sport & Physical Education (2004), also highlights the need to provide quality and motivating Physical Education lessons which promote fun and enjoyment for lifelong participation in PA. Therefore, the present alternated TU is very interesting considering that the promotion of PA practice among pupils is a global health priority (World Health Organization, 2018).

Strengths and limitations

The main strength of the present study was that, to the best of our knowledge, this is the first study that analyzes the effect of an alternated inside-outside TU of physical fitness and sports on pupils' enjoyment during Physical Education lessons. Moreover, there are no previous studies that have evaluated enjoyment day by day during all the lessons of the TU, which allows us to evaluate short-term fluctuations in mood using a minimal amount of time, and to check if there is any difference between inside and outside lessons. Furthermore, the use of a cluster-randomized controlled trial design was the most appropriate considering the nature of the context (i.e., natural school classes) (Campbell et al., 2012; Dreyhaupt et al., 2017). Finally, due to this cluster design, the evaluation of the effect of the TU with a Mixed Multilevel Linear Model with participants nested within classes, represents an advance with respect to the commonly applied analyses (Li et al., 2017). However, the present study also has some limitations that should be considered. Firstly, a limitation was that the study was carried out in a specific region of one country and using a convenience sample, so the findings could not be generalized. However, due to human and material resource restrictions, a probabilistic and larger sample could not be examined. Furthermore, the comparison with a third group (i.e., Control group) which worked another completely different content than Inside and Alternated group, inside the school center, could allow us to check if the teaching contents may affect the outcomes and it could add quality to this study. Nonetheless, the comparison with the Inside group that also worked physical fitness and sports, but only inside the school, allows us to check that the improvements are due to the outside lessons.

Conclusions

In conclusion, the results of the present study showed that the alternated inside-outside TU is effective for improving pupils' enjoyment during Physical Education lessons compared with the inside TU. Moreover, outside Physical Education lessons have increased pupils' enjoyment in comparison with inside lessons. Overall, the study's findings have several important theoretical and practical implications for policymakers as well as for Physical Education teachers when designing and implementing programs, considering it as a meaningful TU to enhance pupils' enjoyment of Physical Education. However, future research studies should examine if these improvements obtained in enjoyment are actually translated into improvements in school-aged children out-of-school PA habits as highlight in previous theories.

Acknowledgments

The authors want to thank you to all the students and the Physical Education teachers who made this research possible. We also thank Aliisa Hatten for the English revision of the manuscript. Carolina Casado-Robles is supported by a research grant from the Spanish Ministry of Universities [grant number: FPU16/03314].

Disclosure statement

No potential conflict of interest was reported by the authors.

Supplementary Files

Scan or click here to see Supplementary Files [Escanea el código o pincha aquí para ver los Archivos Suplementarios]

References

Abós, Á., Sevil, J., Julián, J. A., Abarca-Sos, A., & García-González, L. (2017). Improving students' predisposition towards physical education by optimizing their motivational processes in an acrosport unit. *European Physical Education Review, 23*(4), 444-460. https://doi.org/10.1177/1356336X16654390

American Association of Health, Physical Education, Recreation and Dance (2014). *National standards & grade-level outcomes for K12 physical education.* Human Kinetics

Baena-Extremera, A. & Granero-Gallegos, A. (2015). Efectos de las actividades en la naturaleza en la predicción de la satisfacción de la Educación Física. *Retos, 28*, 9-14. https://doi.org/10.47197/retos.v0i28.34816

Baena-Extremera, A., & Granero-Gallegos, A. (2013). Efecto de un programa de Educación de Aventura en la orientación al aprendizaje, satisfacción y autoconcepto en secundaria. *Revista Iberoamericana de Diagnostico y Evaluacion Psicologica, 2*(36), 163–182.

Baena-Extremera, A., Granero-Gallegos, A., & Ortíz-Camacho, M. M. (2012). Quasi-experimental study of the effect of an Adventure Education program on classroom satisfaction, physical self-concept and social goals in Physical Education. *Psychologica Belgica, 52*(4), 386-396. http://doi.org/10.5334/pb-52-4-369

Baena-Extremera, A., Granero-Gallegos, A., Bracho-Amador, C., & Pérez-Quero, F.J. (2012). Spanish Version of "Sport Satisfaction Instrument (SSI)" adapted to physical education. *Revista de Psicodidáctica, 17*(2), 375-395.

Bai, Y., Allums-Featherston, K., Saint-Maurice, P. F., Welk, G. J., & Candelaria, N. (2018). Evaluation of youth enjoyment toward physical activity and sedentary behavior. *Pediatric Exercise Science, 30*(2), 273-280. https://doi.org/10.1123/pes.2017-0101

Bandura, A. (2004). Health promotion by social cognitive means. *Health Education and Behavior, 31*(2), 143–164. https://doi.org/10.1177/1090198104263660

Baños, R., Marentes, M., Zamarripa, J., Baena-Extremera, A., Ortiz-Camacho, M., & Duarte-Félix, H. (2019). Influencia de la satisfacción, aburrimiento e importancia de la educación física extraescolar en adolescentes mexicanos. *Cuadernos de Psicología del Deporte,19*(3), 205-215. https://doi.org/10.6018/cpd.358461

Barkoukis, V., Ntoumanis, N., & Thøgersen-Ntoumani, C. (2010). Developmental changes in achievement motivation and affect in physical education: Growth trajectories and demographic differences. *Psychology of sport and exercise*, *11*(2), 83-90. https://doi.org/10.1016/j.psychsport.2009.04.008

Briggs, J. D. (1994). An investigation of participant enjoyment in the physical activity instructional setting. *The Physical Educator*, *51*(4), 213-224.

Campbell, M. K., Piaggio, G., Elbourne, D. R., & Altman, D. G. (2012). Consort 2010 statement: Extension to cluster randomised trials. *BMJ, 345,* e5661. https://doi.org/10.1136/bmj.e5661

Carriedo, A. (2019). Satisfaction, learning, and perceptions of elementary students about grappling combat sports. *Sportis, Scientific Technical Journal of School Sport, Physical Education and Psychomotricity,* *5*(1), 133-150. https://doi.org/10.17979/sportis.2017.3.1.1726

Casado-Robles, C., Mayorga-Vega, D., Guijarro-Romero, S., & Viciana, J. (2020). Sport education-based irregular teaching unit and students' physical activity during school recess. *The Journal of Educational Research, 113*(4), 262-274. https://doi.org/10.1080/00220671.2020.1806014

Casado-Robles, C., Guijarro-Romero, S., Viciana, J., & Mayorga-Vega, D. (In press). Adaptación y validación del cuestionario de satisfacción hacia la Educación Física en escolares españoles: Evaluación general de la asignatura y de la clase de hoy. *Proceedings of I Congreso de Estudiantes de Doctorado (CAED)*. Editorial Electrónica

Cole, T. J., Bellizzi, M. C., Flegal, K. M., & Dietz, W. H. (2000). Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ, 320*, 1240-1243. https://doi.org/10.1136/bmj.320.7244.1240

Corder, K., Winpenny, E., Love, R., Brown, H. E., White, M., & Sluijs, E. V. (2017). Change in physical activity from adolescence to early adulthood: a systematic review and meta-analysis of longitudinal cohort studies. *British Journal of Sports Medicine*, 1–9. https://doi.org/10.1136/bjsports-2016-097330

Cox, A., Smith, A., & Williams, L. (2008). Change in physical Education motivation and physical activity behavior during middle school. *The Journal of Adolescent Health 43*(5), 506-513. https://doi.org/10.1016/j.jadohealth.2008.04.020

Cummiskey, M. (2011). There's an App for That Smartphone Use in Health and Physical Education. *Journal of Physical Education, Recreation & Dance, 82*(8), 24-30. https://doi.org/10.1080/07303084.2011.10598672

Dishman, R., Motl, R., Saunders, R., Felton, G., Ward, D., & Pate, R. (2005). Enjoyment mediates the effects of a school-based physical activity intervention among adolescent girls. *Medicine and Science in Sports and Exercise 37*(3), 478-487. https://doi.org/10.1249/01.mss.0000155391.62733.a7

Dreyhaupt, J., Mayer, B., Keis, O., Öchsner, W., & Muche, R. (2017). Cluster-randomized Studies in Educational Research: Principles and Methodological Aspects. *GMS Journal for Medical Education, 34*(2), 1-25. https://doi.org/10.3205/zma001103

Dyment, J. E. (2005). Green School Grounds as Sites for Outdoor Learning: Barriers and Opportunities. *International Research in Geographical and Environmental Education, 14*(1), 28-45. https://doi.org/10.1080/09500790508668328

Ernst, J. (2014). Early childhood educators' use of natural outdoor settings as learning environments: an exploratory study of beliefs, practices, and barriers. *Environmental Education Research, 20*(6), 735-752. https://doi.org/10.1080/13504622.2013.833596

European Commission/EACEA/Eurydice. (2013). *Physical Education and sport at school in Europe Eurydice Report*. Publications Office of the European Union.

Field, A. (2017). *Discovering statistics using IBM SPSS Statistics (5th ed.)*. SAGE Publications.

Granero-Gallegos, A., Baena-Extremera, A., Pérez-Quero, F. J., Ortiz-Camacho, M. M., & Bracho-Amador, C. (2012). Analysis of motivational profiles of satisfaction and importance of physical education in high school adolescents. *Journal of Sports Science & Medicine*, *11*(4), 614.

Granero-Gallegos, A., Baena-Extremera, A., Sánchez-Fuentes, J. A., & Martínez-Molina, M. (2014). Perfiles motivacionales de apoyo a la autonomía, autodeterminación, satisfacción, importancia de la educación física e intención de práctica física en tiempo libre. *Cuadernos de Psicología del Deporte, 14*(2), 59-70.

Gråstén, A., & Yli-Piipari, S. (2019). The Patterns of Moderate to Vigorous Physical Activity and Physical Education Enjoyment Through a 2-Year School-Based Program. *Journal of School Health*, *89*(2), 88-98. https://doi.org/10.1111/josh.12717

Guthold, R., Stevens, G. A., Riley, L. M., & Bull, F. C. (2020). Global trends in insufficient physical activity among adolescents: a pooled analysis of 298 population-based surveys with 1.6 million participants. *The Lancet Child & Adolescent Health, 4*(1), 23-35. https://doi.org/10.1016/S2352-4642(19)30323-2

Hardman, K., Murphy, C., Routen, A., & Tones, S. (2014). *UNESCO-NWCPEA: World-wide survey of school Physical Education.* United Nations Educational, Scientific and Cultural Organization

Hashim, H. (2007). *Components of enjoyment in physical education*. University of Western Australia.

Hernaez, R. (2015). Reliability and agreement studies: A guide for clinical investigators. *Gut*, *64*(7), 1018-1027. https://doi.org/10.1136/gutjnl-2014-308619

Hortiguela, D., Gutierrez-Garcia, C., & Hernando-Garijo, A. (2017). Combat versus team sports: the effects of gender in a climate of peer-motivation, and levels of fun and violence in physical education students. *Ido Movement for Culture. Journal of Martial Arts Anthropology,* *17*(3), 11-20. https://doi.org/10.14589/ido.17.3.2

Institute for Outdoor Learning. (2020). *Outdoor learning research.* Retrieved from: https://www.outdoor-learning.org/Good-Practice/Research-Resources/About-Outdoor-Learning

Li, W., Xiang, P., Chen, Y. J., Xie, X., & Li, Y. (2017). Unit of analysis: Impact of Silverman and Solmon's article on field-based intervention research in physical education in the USA. *Journal of Teaching in Physical Education, 36*(2), 131-141. https://doi.org/10.1123/jtpe.2016-0169

Lubans, D., Lonsdale, C., Cohen, K., Eather, N., Beauchamp, M., Morgan, P., Sylvester, B., & Smith, J.J. (2017). Framework for the design and delivery of organized physical activity sessions for children and adolescents: rationale and description of the 'SAAFE' teaching principles. *International journal of behavioral nutrition and physical activity, 14*(1), 1-11. https://doi.org/10.1186/s12966-017-0479-x

Mackenzie, S. H., Son, J. S., & Eitel, K. (2018). Using outdoor adventure to enhance intrinsic motivation and engagement in science and Physical activity: an exploratory study. *Journal of outdoor recreation and tourism, 21,* 76-86. https://doi.org/10.1016/j.jort.2018.01.008

Martínez-Gómez, D., Martínez-De-Haro, V., Del-Campo, J., Zapatera, B., Welk, G. J., Villagra, A., Marcos, A., & Veiga, Ó. L. (2009). Validez de cuatro cuestionarios para valorar la actividad física en adolescentes españoles. *Gaceta Sanitaria*, *23*(6), 512-517.

Michael, S. L., Coffield, E., Lee, S. M., & Fulton, J. E. (2016). Variety, enjoyment, and physical activity participation among high school students. *Journal of physical activity and health*, *13*(2), 223-230. https://doi.org/10.1123/jpah.2014-0551

National Association for Sport and Physical Education. (2004). *Moving into the future: National standards for physical education (2nd ed.)*. National Association for Sport and Physical Education.

Pearson, N., Haycraft, E., Johnston, J. P., & Atkin, A. J. (2017). Sedentary behaviour across the primary-secondary school transition: A systematic review. *Preventive Medicine, 94,* 40–47. https://doi.org/10.1016/j.ypmed.2016.11.010

Rehman, A., & Haider, K. (2013). The impact of motivation on learning of secondary school students in Karachi: An analytical study. *Educational Research International*, *2*(2), 139-147.

Rehrer, N. J., Freeman, C., Cassidy, T., Waters, D. L., Barclay, G. E., & Wilson, N. (2011). Through the eyes of young people: Favourite places for physical activity. *Scandinavian journal of public health, 39*(5), 492-500. https://doi.org/10.1177/1403494811401478

Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology, 61*, 101860. https://doi.org/10.1016/j.cedpsych.2020.101860

Scanlan, T. K., & Simons, J. P. (1992). The construct of sport enjoyment. In: G. Roberts (Ed.), *Motivation in sport and exercise* (pp. 199–215). Human Kinetics.

Sevil, J., Abós, Á., Aibar, A., Julián, J. A., & García-González, L. (2016). Gender and corporal expression activity in physical education: Effect of an intervention on students' motivational processes. *European Physical Education Review, 22*(3), 372-389. https://doi.org/10.1177/1356336X15613463

SHAPE America. (2013). *Grade-level outcomes for K-12 Physical Education*. SHAPE America

Spanish Ministry of Education, Culture, and Sport. (2015). Royal Decree 1105/2014, laying down the curriculum for Secondary Education and Baccalaureate. *Goverment Gazette, 3*, 480–486.

Stewart, A., Marfell-Jones, M., Olds, T., & De Ridder, J. (2011). *International standards for anthropometric assessment*. International Society for the Advancement of Kinanthropometry.

United Nations Educational, Scientific and Cultural Organisation (UNESCO). (2015). *International charter of physical education, physical activity and sport.* Retrieved from: https://en.unesco.org/themes/sport-and-anti-doping/sport-charter

US Department of Health and Human Services. (2018). *Physical Activity Guidelines Advisory Committee Scientific Report.* US Department of Health and Human Services.

Viciana, J., & Mayorga-Vega, D. (2016). Innovative teaching units applied to Physical Education – changing the curriculum management for authentic outcomes. *Kinesiology 48*(1), 142–152.

Viciana, J., & Mayorga-Vega, D. (2018). The three-axes model of planning in physical education. *Retos, 33*, 313–319. https://doi.org/10.47197/retos.v0i33.54533

White, M. L., Renfrow, M. S., Farley, R. S., Fuller, D. K., Eveland-Sayers, B. M., & Caputo, J. L. (2018). A Cross-Training Program Does Not Alter Self-Reported Physical Activity Levels in Elementary School Children. *International Journal of Exercise Science, 11*(5), 308-318.

World Health Organization. (2018). *Global action plan on physical activity 2018-2030: More active people for a healthier world*. World Health Organization.

World Health Organization. (2020). *WHO guidelines on physical activity and sedentary behaviour*. World Health Organization

Yli-Piipari, S., Barkoukis, V., Jaakkola, T., & Liukkonen, J. (2013). The effect of physical education goal orientations and enjoyment in adolescent physical activity: A parallel process latent growth analysis. *Sport, Exercise, and Performance Psychology, 2*(1), 15.

Paper V

Journal of Teaching in Physical Education
2021
Advance online publication

EFFECT OF AN INSIDE-OUTSIDE SCHOOL ALTERNATED TEACHING UNIT OF KNOWLEDGE OF THE ENVIRONMENT FOR PRACTICING PHYSICAL ACTIVITY: A CLUSTER RANDOMIZED CONTROL TRIAL

Casado-Robles, C., Viciana, J., Guijarro-Romero, S., & Mayorga-Vega, D.

Journal of Teaching in Physical Education

2021, *Advance online publication*. https://doi.org/10.1123/jtpe.2020-0132

Paper V

ABSTRACT

Objective: To examine the effect of two Physical Education-based alternated teaching units on students' environmental knowledge for practicing out-of-school physical activity (PA), perceived autonomy support, self-determined and controlled motivation towards PA, intention to be physically-active, self-reported and objective PA levels, and sedentary behavior.

Methods: A sample of 179 students (94 females), aged 13-15 years old, were cluster-randomly assigned to the Innovative group (two alternated teaching units for practicing PA, with one lesson inside and one outside the school grounds) or the Traditional group (a teaching unit for practicing PA, solely inside the school center).

Results: The alternated teaching units improved students' knowledge of their environment for practicing PA, perceived autonomy, autonomous motivation, intention to be physically active, and self-reported PA during the whole week ($p < 0.05$).

Conclusions: The innovative program improved students' knowledge about their environment for practicing PA and self-reported PA, but did not improve objectively-measured PA levels.

KEYWORDS

Autonomy support. Secondary school. Self-determined motivation.
Accelerometry. Trans-contextual model

Introduction

Currently, there is strong evidence that regular physical activity (PA) is associated with many health benefits in school-aged children (Poitras et al., 2016). Unfortunately, worldwide only one-fifth of high-school students achieve the recommended 60 minutes of moderate-to-vigorous PA (MVPA) (Guthold, Stevens, Riley, & Bull, 2020). Therefore, the promotion of adequate PA levels among high-school students is an important public health challenge, which can take place in various settings of daily life (World Health Organization, WHO, 2018). Among others, school context and specifically the Physical Education (PE) subject play an essential role in students' PA promotion (WHO, 2018).

Although many educational objectives have to be developed throughout the annual PE plan, promoting students' lifelong PA is one of the fundamental aims in most countries (e.g., Spanish Ministry of Education, Culture and Sport, 2015). In order to achieve this purpose, PE teachers need to provide students with tools for practicing PA autonomously, making them competent and transferring learning from the classroom to students' daily life (Viciana & Mayorga-Vega, 2018). A strategy for promoting this transferability of learning may be the education outside the classroom, a teaching method in which teachers relocate teaching activities from the classroom to the environment that surrounds the educational center to provide students authentic and situational PA practices (Becker, Lauterbach, Spengler, Dettweiler, & Mess, 2017; Viciana & Mayorga-Vega, 2018). Previous studies regarding education outside the classroom embedded within the school curricula, have shown to be effective in stimulating students' skills learning, as well as improving key psychological variables, such as students' autonomous motivation (Becker et al., 2017; Bølling, Otte, Elsborg, Nielsen, & Bentsen, 2018). Although education outside the classroom has been used recently to promote students' PA (e.g., Schneller et al., 2017), the systematic review by Becker et al. (2017) highlighted that more research is needed to further investigate these findings.

Furthermore, this transference of learning could be facilitated through the application of the alternated teaching units proposed by Viciana and Mayorga-Vega (2016), which are based on delivering PE lessons belonging to two different teaching units alternatively while putting them into practice. The implementation of two teaching units regarding complementary contents makes students aware that both contents are based on the same learning principles and avoids unconnected learning perceived by students in traditional and isolated units (Viciana & Mayorga-Vega, 2016). Consequently, the use of these alternated teaching units in combination with the education outside the classroom strategy could be an excellent structure of planning to increase students' out-of-school PA. Specifically, it could allow PE teachers to connect in-school PA (inside the school) with the immediate environment practice (outside the school), promoting the transference of significant learning in the PE context to students' daily life.

Additionally, understanding the determinants and factors influencing PA behavior is crucial to implement effective PA interventions for students in the PE setting (Sheeran, Klein, & Rothman, 2017). The Social Cognitive Theory (Bandura, 2004) includes health-related knowledge as a core determinant on the design of PA promotion interventions, which can be conveyed readily to students and supposes a first step toward generating a behavior change. Previous studies have shown that the students' knowledge about PA practice leads to a change in PA behavior and voluntary participation in PA (Wang & Chen, 2019). Moreover, the Self-Determination Theory (SDT, Ryan & Deci, 2020) is a theoretical framework widely used to explain the antecedents and consequences of students' motivation to engage in PA. The SDT considers motivation as a multidimensional term with different levels along a continuum according to the degree of autonomy, ranging from more self-determined (i.e., autonomous) to less self-determined (i.e., controlled) forms of behavioral regulations (Ryan & Deci, 2020). Furthermore, the SDT highlights the concept of three basic psychological needs and their satisfaction as a fundamental factor for understanding students' motivation: autonomy (referring to the need for initiative and ownership in one's own behavior), competence (referring to the need to feel capable of carrying out a certain behavior effectively), and relatedness (referring to the need of feeling connected and supported by significant others).

According to the SDT, autonomy support is an important social context for encouraging higher levels of students' more self-determined regulations from the SDT perspective (Deci & Ryan, 2008). Specifically, teachers who support students' autonomy are favoring the satisfaction of the three psychological needs (i.e., autonomy, relatedness, and competence), and, therefore, promoting the students' most self-determined forms of motivation, while in contexts where autonomy support is not provided, these needs are thwarted (Ryan & Deci, 2020). This autonomy support in educational contexts is characterized by an environment that supports students' choice, initiation, and understanding, where students feel that they can participate in their own learning (Deci & Ryan, 2008). According to this framework, the use of an autonomy-supportive teaching style and the improvement of students' autonomous motivation are considered key determining factors related to the acquisition and maintenance of students' PA (Teixeira, Carraça, Markland, Silva, & Ryan, 2012). Consequently, school-based interventions with the aim of promoting students' lifelong PA should be based on psychological theories that have motivation and autonomy as a core point of its tenets.

For this purpose, the Trans-Contextual Model (TCM, Hagger & Chatzisarantis, 2016) could be an ideal framework on which to base PE interventions focused on promoting motivation toward PA and PA participation out-of-school. Hagger & Chatzisarantis (2016) proposed three empirically-testable propositions to explain how teachers can encourage students' PA participation during out-of-school contexts. The first proposition hypothesizes that students' perceived autonomy support from teachers can develop students' self-determined motivation toward in-class activities, and it will predict self-determined motivation toward similar activities within the educational context. The second hypothesizes that self-determined motivation within the educational context will predict self-determined motivation toward similar activities in out-of-school contexts. Finally, the third proposition hypothesizes that self-determined motivation toward activities in an out-of-school context will predict future intention to engage in similar activities as well as actual behavioral engagement. The systematic review conducted by Hagger & Chatzisarantis (2016) showed their empirical support across multiple studies conducted in the PE setting highlighting significant relationships between perceived autonomy support and self-determined motivation in PE (i.e., first proposition), between self-determined motivation in PE and in PA (i.e., second proposition), and between self-determined motivation and intention toward PA and actual PA engagement (i.e., third proposition). Therefore, according to this model, a key target is for teachers' autonomy-supportive behaviors and active styles of teaching in PE to be transferred to an increase of self-determined motivation, future intention, and actual participation in PA during out-of-school contexts (Wang & Chen, 2019).

In addition to these above-mentioned individual theories related to students' PA promotion, the integration of some constructs from each of those theories could help to better understand the underlying mechanisms of behavior change using each theory's strengths (Biddle, Hagger, Chatzisarantis, & Lippke, 2007). Specifically, the increase of students' health-related knowledge proposed by the Social Cognitive Theory to promote PA practice, could also improve students' perception of competence (i.e., one of the basic psychological needs proposed by the SDT), which leads to greater levels of students' more self-determined motivation forms (Ryan & Deci, 2020). Furthermore, according to the TCM postulates, those high levels of students' autonomous motivation will work to increase students' intention and PA practice (Hagger & Chatzisarantis, 2016). Therefore, the increase of students' knowledge of the possibilities of PA practice offered by their environment along with an autonomy support context in PE, could be an effective way to promote students' motivation towards PA and ultimately, actual PA practice out-of-school.

Unfortunately, to our knowledge, no previous studies have examined the effect of two alternated teaching units on the environmental knowledge for practicing out-of-school PA, motivation, and PA participation transferred from the PE class to real-life. Additionally, although several correlational studies have shown the TCM to be a useful framework for understanding how autonomy-supportive behaviors during PE lessons can promote PA practice during out-of-school contexts, the current research evidence of school-based interventions studies according to the postulates of the TCM is still scarce (e.g., Mavropoulou, Barkoukis, Douka, Alexandris, & Hatzimanouil, 2018). Furthermore, there is a lack of intervention studies based on the TCM examining PA participation using objectively measured PA. The present study was designed to address these needs from a holistic perspective by investigating the effectiveness of two alternated teaching units (i.e., inside and outside the school center) involving fitness tasks and games/sports, and promoting an autonomy-supportive teaching style compared with a traditional teaching unit delivered only inside the school center involving fitness tasks and games/sports and based on direct instruction methodologies. Thus, the main aim of the present study was to compare the effects of two alternated teaching units (i.e., inside and outside the school center) and a traditional teaching unit on students' knowledge of their environment for practicing PA. The secondary aim was to compare the effects of the above-mentioned teaching units on students' perceptions of autonomy support, autonomous and controlled motivation towards PA, intention to be physically active, and self-reported and objective PA levels. The main hypotheses were that students in the innovative program will obtain: (a) higher knowledge of their environment for practicing PA; (b) higher perceptions of autonomy support and, in consequence, higher autonomous and lower controlled motivation towards PA; and (c) higher intention to be physically active, and self-reported and objective PA levels as compared with students attending the traditional teaching unit.

Methods

Study design

The present study is reported according to the CONSORT for cluster randomized trials guidelines (Campbell, Piaggio, Elbourne, & Altman, 2012). The protocol conforms to the Declaration of Helsinki statements (64th WMA, Brazil, October 2013) and it was approved by the Ethical Committee for Human Studies at the University of Granada. For practical reasons and due to the nature of the present study (i.e., pre-established classes in a school setting), a cluster randomized controlled trial design was used.

Participants

The principal and the PE teachers of a public high-school center (chosen by convenience) from an urban area situated in Granada were contacted and informed about the project, and provided the approval to carry out the study. All 179 students (52.5% females) from the eighth to ninth grades (i.e., aged 13–15 years old) were invited to participate in the present study. Students and their legal guardians were fully informed about the project features. Participants' signed written informed assent and their legal guardians' signed written informed consent before taking part in the study.

The following inclusion criteria were considered: a) being enrolled in the eighth to ninth grade at the secondary education level; b) being free of any health disorder that would make them unable to engage in PA normally; and c) presenting the corresponding signed written consent/assent by legal guardians/students. The exclusion criterion was not having performed the dependent variables correctly at the beginning and/or at the end of the intervention program following the administration rules (being removed only for incomplete variables and not for the overall study).

Sample size

A priori sample size calculation was estimated with the Optimal Design Plus Empirical Evidence Software Version 3.01 for Windows. Parameters were set as follows: significance level $\alpha = 0.05$, number of participants per cluster $n = 25$, effect size $\delta = 0.50$, intra-class correlation coefficient $\rho = 0.01$, and statistical power $(1 - \beta) = 0.80$. A total number of six clusters (150 participants) was estimated.

Randomization

Before the pre-intervention evaluation was administered, the six pre-established available classes were, balanced by grade, randomly assigned by an independent and blinded research into the Traditional Group (TG) or the Innovative Group (IG). Randomization was conducted at the class-level, using a computerized random number generator. However, according to the education rules, the school center assigned the students randomly and balanced by gender to each class, before starting the scholar year.

Intervention

Before the intervention, the guidelines for correctly delivering the lessons of each teaching unit were designed by the researchers and given to the PE teachers. Figure 1 shows the general scheme of the intervention. Both groups carried out a four-week intervention program (eight PE lessons, two lessons a week). Each PE lesson lasted approximately 55 minutes and consisted of: a 10-minute warm-up, a 40-minute main part and a five-minute cool-down.

Innovative group

The IG students carried out two alternated teaching units (Viciana & Mayorga-Vega, 2016) for the practice of inside and outside school PA. It consisted of delivering one inside lesson (i.e., in school teaching unit, using conventional school facilities like a gym or sports courts) followed by another outside lesson in the immediate environment (i.e., out-of-school teaching unit, using outside installations and features, green zones, or a municipal sport center) during the whole program (four of each modality). The main part of each lesson was divided into two portions of approximately the same length (i.e., 20 min each). The first one was focused on physical fitness and health tasks (e.g., CrossFit-teens). The second part was focused on traditional and alternative games and sports (e.g., basketball). During both in-school and out-of-school lessons of the same week, students performed the same contents and the main difference was the material and the space used in each of them. This structure was developed to establish learning transference from the PE class to the out-of-school context in order to promote PA practice in the immediate environment.

The lessons teaching approach was based on three main characteristics: (a) to teach general concepts about PA; (b) to significantly develop students' knowledge of the environment for practicing PA; and (c) the use of a teaching methodology focused on motivational strategies and students' autonomy. These features were included based on previous evidence (Cheon, Reeve, Lee, & Lee, 2018; Hagger & Chatzisarantis, 2016; Wang & Chen, 2019) in order to encourage PA participation through the increase of autonomous motivation. These strategies are completely detailed in Table 1.

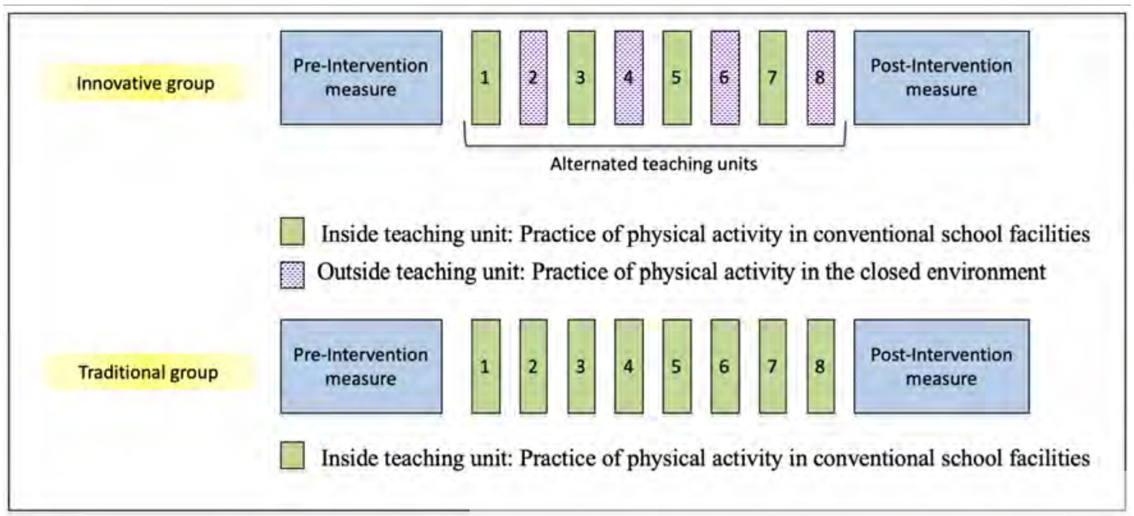


Figure 1. Intervention programs. Alternated inside-outside teaching units versus Traditional teaching unit (adapted from Viciana & Mayorga-Vega, 2016).

Table 1. Strategies applied in the Innovative Group during the intervention and fidelity checklist.

Explanation of concepts about PA
Main concepts related to PA (e.g., frequency, levels of intensity, physical inactivity, or SB)
International PA guidelines (i.e., daily goal of 60 min of MVPA or 10.000 steps/day)
Kinds of activities (aerobic, muscle strengthening, or bone-strengthening) and examples to reach the daily goal
Health-related benefits of PA and the harmful effect of physical inactivity in school-aged children
Strategies to solve barriers towards PA practice proposed by students
Principles for designing a PA plan with an example week (e.g., selection of frequency, intensity, length, or kind of activity)
Debate about possible students' PA habits modification during the program, where students give their reasons for and/or against this change
Review of the main concepts explained during the previous lessons
Knowledge of the environment for practicing PA
Highlighting the similarities between tasks and how these can be done in the other context (i.e., inside or outside)
Students do situational PA practice in out-of-school installations (i.e., green zones, urban parks or sports center)
Publicizing the structured PA offered by the community (e.g., modalities and timetables offered by sports clubs)
Teacher offers a free trial session at the sports club of students' choice
Teacher hands a weekly physical fitness plan individualized to three levels of difficulty and using outside installations and features
Use of demonstration videos about tasks developed in outside/in-school contexts, in order to have students think about how to carry them in the other context.
Teaching methodology
Students' capability is considered in order to individualize the tasks according to the students' level
Creating opportunities for choice (e.g., choosing activities from a list of what they liked most with the same goal)
Using of interrogative feedback in order to make students reflect on the relationship between the two contexts
Using of an informational and non-controlling language (i.e., avoidance of directives and commands)
Providing explanatory and meaningful rationales (e.g., points out the importance or benefits of a task)
Teacher actively listens to students' PA concerns and has an open attitude towards resolving them
Encouraging students to improve their PA levels to reach the guidelines
Proposing small-group activities with the opportunity to choose activity partners to support their relatedness
Teacher offers hints on how to make progress on tasks
Innovative contents and resources (i.e., an augmented reality mobile app) are included

Note. PA = Physical Activity; SB = Sedentary behavior; MVPA = Moderate-to-vigorous PA.

Traditional group

The TG students developed a similar lesson structure, contents, and tasks as the IG. Similar to the IG, the first portion of each lesson was focused on improving physical fitness and the second was focused on traditional and alternative games and sports. However, unlike the IG, students only received in-school lessons using conventional school facilities and materials. Furthermore, the TG group did not receive any of the specific strategies developed in the IG (i.e., PA concepts, knowledge of the environment nor specific motivational strategies). Therefore, no transference of learning from the PE context to the out-of-school context in the immediate environment was promoted. The main strategies applied are detailed in Supplementary Table 1.

Intervention programs fidelity

The main researcher observed carefully the fidelity of each teaching unit. Specifically, and according to the checklists reported in Table 1 for the IG and in Supplementary Table 1 for the TG, the observer registered if the behavior described by each item was present (i.e., it was used at least once during the lesson) or not during the lesson. Afterward, the percentage of fidelity was calculated. Firstly, if the evaluated behavior (i.e., the item) was present in at least the 75% of the lessons, it was considered that this specific strategy had been faithfully applied according to its design. Secondly, the percentage of overall fidelity for each category was calculated by dividing the number of items that met that criterion by the total number of items in the category and multiplied by 100. Therefore, regarding IG teaching unit fidelity, 100% of fidelity was obtained in the categories of general concepts about PA and knowledge of the environment for practicing PA. However, in the teaching methodology category, 90% fidelity was obtained (i.e., nine out of the 10 items observed were applied correctly) as the item "Teacher provides explanatory and meaningful rationales" was not correctly completed in three out of the eight lessons. Regarding TG teaching unit fidelity, the main researcher registered 100% fidelity in the three categories observed.

Measures

Knowledge test about the environment to practice physical activity

Students' knowledge was measured through the *Knowledge about the Environment to Practice PA in schoolchildren test* (CEPAF, Casado-Robles et al., 2021). It consisted of 30 questions with four possible answers where only one was correct, divided into three knowledge dimensions: (a) Declarative (i.e., *to know* the *what* of something); (b) Procedural (i.e., *to know how* to do something); and (c) Causal (i.e., to *know why* something happens). Furthermore, it was also divided into two main contents (i.e., physical fitness and games/sports). The CEPAF test has shown adequate reliability and validity among high-school students (ICC = 0.72; adequate difficulty and discrimination indices; discriminant validity $p < 0.001$, $d = 1.54$; Casado-Robles et al., 2021).

Perceived autonomy support

The PE teacher autonomy-support was assessed through the Spanish version of the Perceived Autonomy Support Scale for Exercise Settings (PASSES, Moreno, Parra, & González-Cutre, 2008). It consists of 12 items that evaluate a single factor of autonomy support in a 10-point Likert-type scale. The Spanish version of PASSES has shown adequate psychometric properties among high-school students (CFI = 0.92; IFI = 0.92; TLI = 0.90; SRMR = 0.04; RMSEA = 0.08; $\alpha = 0.91$; Moreno et al., 2008).

Self-determined motivation towards physical activity

Students' motivation towards PA was measured using the Spanish version of the Behavioral Regulation in Exercise Questionnaire (BREQ-3, González-Cutre, Sicilia, & Fernández, 2010). It consists of 23 items distributed into six dimensions (intrinsic motivation, demotivation, integrated, identified, introjected and external regulation). A 10-point Likert-type scale was also used. Subsequently, the autonomous (i.e., intrinsic, integrated and identified regulation) and controlled (i.e., introjected and external) motivation was calculated using the average of the corresponding dimensions (Chemolli & Gagné, 2014). The Spanish version of the BREQ-3 has shown adequate psychometric properties among high-school students (CFI = 0.91; IFI = 0.91; RMSEA = 0.06; SRMR = 0.06; $\alpha = 0.66\text{-}0.87$; González-Cutre et al., 2010).

Intention to be physically active

The Spanish version of the Intention to partake in leisure-time PA questionnaire was applied to measure the students' intention to be physically active in their free time (Granero-Gallegos, Baena-Extremera, Pérez-Quero, Ortiz-Camacho, & Bracho-Amador, 2014). It is composed of three items measured in a 10-point Likert-type scale. The Spanish version of this questionnaire has shown adequate psychometric properties among high-school students (GFI = 1.00; CFI = 1.00; RMSEA = 0.03; α = 0.93; Granero-Gallegos et al., 2014).

Self-reported habitual physical activity and sedentary behavior

Self-reported PA and sedentary behavior (SB) profiles were assessed with the Spanish version of the Youth PA Profile (YAP, Saint-Maurice & Welk, 2015). It includes 15 items referring to PA and SB in different contexts. A 5-point Likert-type scale from 0 to 4 was used. Subsequently, six categories were calculated using the average of the scores (0 to 4) obtained for the corresponding dimensions: (a) PA during weekdays (items 1 to 8); (b) PA in after-school time (items 6 to 8); (c) PA during weekends (items 9 to 10); (d) SB (items 11 to 15); (e) Total leisure-time PA (items 6 to 10); (f) Whole week PA (items 1 to 10). The YAP has shown adequate validity to estimate PA and SB among high-school students in comparison with objective instruments (e.g., MVPA, r = 0.58; SB, r = 0.75; Saint-Maurice & Welk, 2015).

Objectively measured habitual physical activity and sedentary behavior

Students' PA and SB were objectively measured by a GT3X+ accelerometer (ActiGraph, LLC, Pensacola, FL, USA). Students were asked to wear accelerometers on their right hip for eight consecutive days from waking to bedtime, except when they engaged in aquatic activities (Chinapaw et al., 2014). All data were downloaded and analyzed using the ActiLife Lifestyle Monitoring System Software version 6.13.3.

In order to avoid participants' reactivity biases, the first day was considered as a familiarization day and it was not used for statistical analysis (Mattocks et al., 2008). Accelerometers were initialized with a sample ration of 30 Hz (Migueles et al., 2017). Since short bursts of rapidly changing activity characterize schoolchildren's behavior, a one-second *epoch* was used (Migueles et al., 2017). A minimum wear time of 480 min per day and 240 min per after-school time was set (Chinapaw et al., 2014). Non-wear periods were set with a minimum length of 60 min of consecutive zero-count *epochs* with up to two minutes of spike tolerance (Chinapaw et al., 2014). Regarding the data scoring, Evenson's cut-off points were used to obtain the percentage of time engaged in SB and MVPA (i.e., ≤ 100 counts/min and ≥ 2,296 counts/min, respectively; Trost, Loprinzi, Moore, & Pfeiffer, 2011). To calculate students' PA during the whole week, a minimum criterion of two weekdays and one weekend day with valid time was established (Mattocks et al., 2008). To calculate the students' PA during weekdays/after-school time and weekend days, a minimum criterion of two days and one day with valid time, respectively, was established. ActiGraph accelerometer-measured PA and SB has shown a high reliability and validity among school-aged children (e.g., SB, ROC-AUC = 0.90, Se = 1.00, Sp = 0.79; MVPA, ROC-AUC = 0.90, Se = 0.88, Sp = 0.92; Trost et al., 2011).

Statistical analysis

Mean (± standard deviation; continuous variables) or percentage (categorical variables) for the general characteristics of the participants and dependent variables were calculated. The one-way analyses of variance (ANOVA; continuous variables) and the chi-squared test (categorical variables) were conducted to examine potential differences in terms of general characteristics and attendance rate between the two groups. The internal consistency of the dependent variables measured by the questionnaires was examined with the Cronbach's alpha.

All the participants were included in the statistical analyses regardless of adherence to the protocol (i.e., intention-to-treat approach). However, since the implementation of the missing data requires strong assumptions that are hard to justify, "complete case" analyses including only those whose outcomes were known were used (Campbell et al., 2012). Missing data was low (4.5-5.6%), except for the objectively measured habitual PA/SB (32.4-51.4%). Because the unit of intervention was the class, the effect of the teaching units on the dependent variables was examined using a Multilevel Linear Model with participants nested within classes and measures nested within participants as random effects, and with the between-subjects factor *group* (TG, IG) and the within-subject factor *time* (pre-intervention, post-intervention) as fixed effects (i.e., two-way mixed nested ANOVA/ANCOVA; Li, Xiang, Chen, Xie, & Li, 2017). All the potential confounding variables (i.e., gender, age, grade, body mass, body height, body mass index, and intervention attendance) were explored and used as covariables when necessary (see Tables 1-3 footnotes). The maximum likelihood estimation method was used. The *post-hoc* within-subject pairwise comparisons with the Bonferroni adjustment was carried out. Effect sizes were estimated using the Cohen's d. All statistical analyses were performed using the SPSS version 25.0 for Windows (IBM® SPSS® Statistics). The statistical significance level was set at $p \leq 0.05$.

Results

Final sample and general characteristics

Although all the invited 179 students (52.5% female) agreed to participate and met the inclusion criteria, the number of students that satisfactorily passed the exclusion criterion to become the final sample group was different depending on each dependent variable (N = 87-171) (Figure 2). The results of the one-way ANOVA and the chi-square test did not show statistically significant differences in terms of general characteristics between the two groups (p > 0.05; TG, age = 14.1±1.1 years, gender = 52.3%/47.7% females/males, grade = 64.8/35.2% eighth/ninth, body mass = 61.2±14.8 kg, body height = 164.5±8.4 cm, body mass index = 22.5±4.6 kg/m^2, weight status = 65.5/34.5% overweight/obese; IG, age = 14.0±0.8 years, gender = 52.7%/47.3% females/males, grade = 64.8/35.2% eighth/ninth, body mass = 58.4±11.7 kg, body height = 164.1±8.0 cm, body mass index = 21.6±3.6 kg/m^2, weight status = 69.2/30.8% overweight/obese). In the sample of the present study, the internal consistency of all the dependent variables measured by dimensional questionnaires was above 0.80. Regarding the attendance rate, the IG and TG participants obtained an average of 91.5% and 89.3%, respectively (overall = 90.4%). The results of the one-way ANOVA and the chi-square test did not show statistically significant differences in attendance rate between the two groups (F = 0.908, p = 0.342; χ^2 = 0.769, p = 0.380).

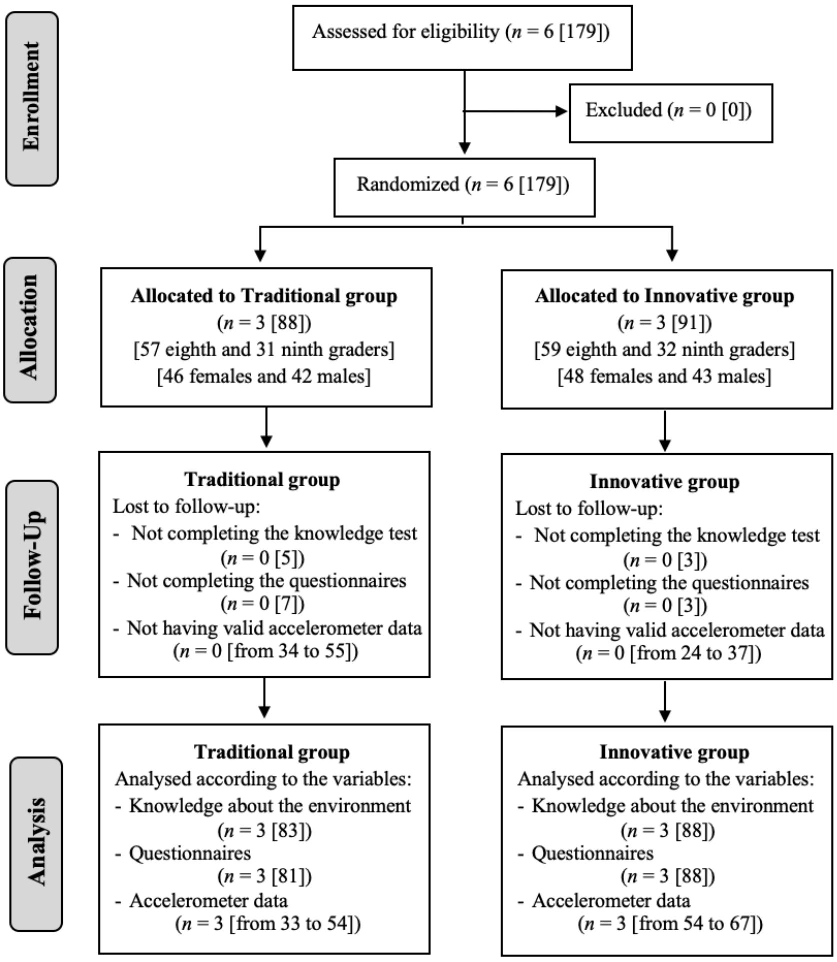


Figure 2. Flow chart of the school classes and students of the present study. All numbers are reported as school classes [students].

Knowledge about the environment for practicing physical activity

The MLM results showed statistically significant interaction effects between *group* and *time* in all the knowledge dimensions ($p < 0.001$; Table 2). Subsequently, the pairwise comparisons showed that the IG participants statistically significantly improved all their knowledge dimensions from pre- to post-intervention ($p < 0.001$; $\partial = 0.90\text{-}1.97$). However, while the TG participants statistically significantly improved their procedural knowledge ($p < 0.05$), they reduced their declarative and causal knowledge ($p < 0.05$), and statistically significant differences were not found for the overall knowledge ($p > 0.05$) from pre- to post-intervention.

Table 2. Effect of the inside-outside alternated teaching units on the knowledge of the environment for practicing physical activity

Variable	Group	Pre-intervention	Post-intervention	Multilevel Lineal Model			ES
		M (SE)	M (SE)	- 2LL	*F*	*p*	∂
Knowledge of the environment							
Declarative	Innovative	3.6 (0.2)	5.5 (0.2)***	1344.674	59.763	< 0.001	1.63
	Traditional	3.8 (0.2)	3.0 (0.2)**				
Procedural	Innovative	3.2 (0.2)	6.4 (0.2) ***	1306.334	57.307	< 0.001	1.57
	Traditional	3.2 (0.2)	3.9 (0.2)**				
Causal[a]	Innovative	4.4 (0.2)	5.6 (0.2)***	1341.860	26.940	< 0.001	0.90
	Traditional	4.3 (0.2)	3.8 (0.2)*				
Overall	Innovative	11.3 (0.4)	17.5 (0.4)***	1819.421	130.584	< 0.001	1.97
	Traditional	11.3 (0.4)	10.7 (0.4)				

ES = Effect size; M = Mean; SE = Standard error; - 2LL = -2 log-likelihood; ∂ = Cohen's ∂ effect size; Innovative, *n* = 88, Traditional, *n* = 83. *Post-hoc* pairwise comparisons with the Bonferroni adjustment from pre- to post-intervention (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$). Covariables: Gender[a].

Perceived autonomy support and self-determined motivation toward physical activity

The MLM results showed statistically significant interaction effects between *group* and *time* on perceived autonomy support and controlled motivation toward PA ($p < 0.01$; Table 3). Moreover, regarding the autonomous motivation toward PA a tendency towards statistical significance was found ($p < 0.10$). Subsequently, the pairwise comparisons showed that the IG participants statistically significantly improved their perceived autonomy ($p < 0.001$; $\partial = 1.03$) and autonomous motivation ($p < 0.05$; $\partial = 0.22$) from pre- to post-intervention. Additionally, the IG participants statistically significantly decreased their controlled motivation ($p < 0.05$; $|\partial| = 0.42$). However, the TG participants statistically significantly reduced their perceived autonomy ($p < 0.001$), while for the rest of comparisons statistically significant differences were not found ($p > 0.05$).

Table 3. Effect of the inside-outside alternated teaching units on perceived autonomy support and self-determined motivation toward physical activity.

Variable	Group	Pre-intervention	Post-intervention	Multilevel lineal model			ES
		M (SE)	M (SE)	- 2LL	*F*	*p*	*∂*
Perceived autonomy support							
Autonomy	Innovative	6.8 (0.2)	7.8 (0.3)***	1454.621	33.695	< 0.001	1.03
	Traditional	7.0 (0.2)	5.8 (0.3)***				
Self-determined motivation							
Autonomous[a]	Innovative	7.3 (0.2)	7.7 (0.2)*	1347.578	2.813	0.095	0.22
	Traditional	6.9 (0.2)	6.8 (0.2)				
Controlled[b]	Innovative	3.1 (0.2)	2.6 (0.2)**	1199.214	9.580	0.002	-0.42
	Traditional	2.3 (0.2)	2.5 (0.2)				

ES = Effect size; M = Mean; SE = Standard error; - 2LL = -2 log-likelihood; *∂* = Cohen's *∂* effect size; Innovative, n = 88, Traditional, n = 81. *Post-hoc* pairwise comparisons with the Bonferroni adjustment from pre- to post-intervention (*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$). Covariables: Gender[a]; Body height[b]

Intention to be physically active, physical activity, and sedentary behavior

The MLM results showed statistically significant interaction effects between *group* and *time* on the intention to be physically active, self-reported PA during weekdays, the after-school time, total leisure time and the whole week ($p \leq 0.05$; Table 4). Subsequently, the pairwise comparisons showed that the IG participants statistically significantly improved all the above-mentioned variables from pre- to post-intervention ($p < 0.001$; $\partial = 0.27$-0.36). On the other hand, regarding the PA levels during weekend days and overall SB statistically significant interaction effects were not found ($p > 0.05$). However, for the pairwise comparisons with the TG participants, statistically significantly differences were not found from pre- to post-intervention ($p > 0.05$). Regarding the objectively measured PA and SB, the MLM results did not show statistically significant interaction effects between *group* and *time* on habitual MVPA and SB during weekdays, weekend, after-school time, nor whole week ($p > 0.05$; Supplementary Table 2).

Table 4. Effect of the inside-outside alternated teaching units on the intention to be physically active and self-reported physical activity and sedentary behavior.

Variable	Group	Pre-intervention	Post-intervention	Multilevel Lineal Model			ES
		M (SE)	M (SE)	- 2LL	*F*	*p*	*∂*
Intention to be physically active							
Intention[a]	Innovative	6.9 (0.3)	8.1 (0.3)***	1566.147	4.461	0.036	0.30
	Traditional	6.9 (0.3)	7.2 (0.3)				
Self-reported physical activity							
Weekdays (PA)[a]	Innovative	2.7 (0.1)	2.9 (0.1)***	525.792	4.607	0.033	0.29
	Traditional	2.7 (0.1)	2.7 (0.1)				
After-school time (PA)[a]	Innovative	2.4 (0.1)	2.8 (0.1)***	877.395	6.336	0.013	0.36
	Traditional	2.4 (0.1)	2.5 (0.1)				
Weekend days (PA)[a]	Innovative	2.8 (0.1)	3.0 (0.1)	992.854	0.355	0.552	0.09
	Traditional	2.7 (0.1)	2.8 (0.1)				
Overall sedentary[a]	Innovative	2.6 (0.1)	2.4 (0.1)	605.993	0.342	0.559	-0.07
	Traditional	2.6 (0.1)	2.5 (0.1)				
Total Leisure time[a]	Innovative	2.5 (0.1)	2.9 (0.1)***	799.945	4.097	0.045	0.30
	Traditional	2.5 (0.1)	2.6 (0.1)				
Whole week[a]	Innovative	2.7 (0.1)	3.0 (0.1)***	532.175	3.836	0.052	0.27
	Traditional	2.7 (0.1)	2.7 (0.1)				

ES = Effect size; M = Mean; SE = Standard error; - 2LL = -2 log-likelihood; PA = Physical activity; *∂* = Cohen's *∂* effect size; Innovative, *n* = 88, Traditional, *n* = 81. *Post-hoc* pairwise comparisons with the Bonferroni adjustment from pre- to post-intervention (***$p < 0.001$). Covariable: Gender[a].

Discussion

The first hypothesis of the present study was focused on examining if two alternated teaching units (i.e., inside and outside) promoting an autonomy-supportive teaching style are more effective than a traditional teaching unit for increasing students' knowledge of their environment for practicing PA. The results of the study showed that the innovative program improved the students' knowledge from pre to post-intervention. These findings are meaningful because if students do not have enough knowledge about their environment, it is unlikely they will be competent enough to practice PA autonomously out-of-school as it is an enabling factor for PA practice (Bandura, 2004). Therefore, according to the Social Cognitive Theory (Bandura, 2004), the IG students' acquisition of the knowledge about *what* kind of PA tasks to do in their environment, *how* to do them, and *why* to practice them, could positively influence their perception of capability for practicing PA masterfully on their own (i.e., satisfying their need for competence) and being the first step toward generating a behavior change. Consequently, a positive effect of this knowledge variable on the students' behavior change towards PA is expected (Wang & Chen, 2019).

Moreover, meeting with the goal of the transferability of learning from the PE class to students' daily life (Spanish Ministry of Education, Culture and Sport, 2015; Viciana & Mayorga-Vega, 2018), these inside-outside alternated teaching units improved students' knowledge, making them capable of using the resources offered by the environment and structure their own PA program during out-of-school time. In this line, some previous studies proved the importance of enhancing health-related knowledge during school-based interventions for voluntary PA participation in leisure time promotion (e.g., Chen, Chen, Sun, & Zhu, 2013). However, those studies were focused on fitness knowledge (i.e., training principles), and as far as we know, no previous research has inquired on the students' knowledge about their environment for practicing PA. On the other hand, conflictive findings have been found in the TG students, with them improving their procedural knowledge but decreasing their declarative and causal knowledge. These results seem to indicate that the development of the same contents as the innovative teaching unit (i.e., physical fitness tasks and games/sports) improves students' knowledge about *how* to do specific tasks in the environment. However, the realization of only inside PE lessons working on physical fitness and games/sports seems not enough to increase students' overall knowledge about their environment for practicing PA.

A secondary hypothesis was focused on examining if the innovative program is more effective than a traditional teaching unit for increasing students' perceptions of autonomy support and self-determined motivation towards PA. The results of the present study showed that the innovative program increased students' perceived autonomy in PE classes from pre- to post-intervention, and that may be due to: (a) the PA counseling included which provided explanations as to why the behavior is truly worth the students' effort (Dobbins, Husson, DeCorby, & LaRocca, 2013); (b) the teaching methodology based on nurturing students' inner motivation and autonomy (see Table 1 for detailed strategies; Wang & Chen, 2019; Cheon et al., 2018); and (c) the use of out-of-school contexts to deliver PE classes, which provide students with authentic performances they can apply to their daily life's competencies (Viciana & Mayorga-Vega, 2018). As a result of these applied autonomy support strategies, the present study has also shown a positive effect on students' autonomous motivation towards PA and a decrease in students' controlled motivation towards PA, from pre- to post-intervention.

It is in line with the SDT (Ryan & Deci, 2020) and past evidence, which imply that perceived autonomy support in the educational context would have an indirect effect on promoting students' autonomous motivation toward PA in out-of-school contexts (Cheon et al., 2018). According to Wang & Chen (2019), this increase in autonomous motivation towards PA may also be related to an increment in students' competence and autonomy, as a result of the improvement of the knowledge about their environment for practicing PA. Furthermore, these findings are also in line with previous Education Outside programs which have shown to be effective in nurturing students' autonomous motivation through relocating their PE lessons from the conventional classroom to places outside the school center (e.g., Bølling et al., 2018). Therefore, although further research is needed to confirm the influence of each strategy used in the present study on students' positive behavior towards PA, it is possible to conclude that combining the above-mentioned strategies with the alternated teaching units is an effective PE-based intervention to engage students with the real context where the PA would be implemented in their leisure time. Conversely, students' perceived autonomy decreased in the TG students while their self-determined motivation did not change. This may be due to the teaching methodology applied in the TG, where students could not elect any aspect of their class, the tasks were not individualized according to their skill levels, and there was a predominance of instructional feedback used by the PE teacher.

Finally, the third hypothesis was based on examining if the innovative program is more effective than a traditional teaching unit for increasing students' intention to be physically active, and self-reported and objective PA levels. The present results support an increase in the intention to be physically active after experiencing an autonomy-supportive program from pre- to post-intervention. These findings provide further evidence to the SDT which supports students' autonomous motivation as being an influential factor related to the acquisition of PA habits in students (Teixeira et al., 2012). Moreover, these results are also in line with previous intervention programs based on the TCM. For instance, Mavropoulou et al. (2018) found higher students' intention to be physically active in leisure time in the group where PE teachers adopted an autonomy-supportive methodology (e.g., explaining the benefits of PA or giving positive feedback) after similar length intervention programs (i.e., six weeks).

Regarding the actual PA participation, students in the IG increased self-reported PA during weekdays, the after-school time, total leisure time, and the whole week from pre- to post-intervention. These findings are so valuable because the main goal of this innovative program seems to be achieved (i.e., an increase in students' PA levels in most of the self-reported variables). Furthermore, it is in line with the third proposition of the TCM which hypothesizes that autonomous motivation toward PA will predict future intention to engage in similar activities as well as actual behavioral engagement (Hagger & Chatzisarantis, 2016). Nevertheless, it should be noted that no effect was found for students' objectively-measured habitual MVPA and SB levels. Therefore, although the TCM suggests that autonomy-supportive contexts will lead to students' PA participation (Hagger & Chatzisarantis, 2016), in previous literature, inconsistent results have been registered in the last steps trying to link the improvement of psychological variables (e.g., perceived autonomy support or intention to be active) with an increment in out-of-school PA after an intervention program, and most of the time their results depend on the measurement. Self-reported PA measurements had confirmed the TCM hypotheses (e.g., Hagger & Chatzisarantis, 2016; Yli-Piipari, Lavne, Hinson & Irwin, 2018), whereas objectively measurements had not (Viciana et al., 2019).

Therefore, the feasibility of the TCM in predicting PA out-of-school seems not to be fully supported for objective PA, and more interventions are needed to test its effectiveness. Moreover, the absence of improvements in habitual objectively-measured PA levels could be obtained due to the short length of the intervention (four weeks), since previous literature suggests longer school-based PA interventions (around 12 weeks and up) in order to achieve changes in PA behavior (Dobbins et al., 2013). Moreover, complementing the intervention with a proposal of an extracurricular PA plan for weekends might have also helped to achieve significant results because, although they learned about how to use their environment for practicing PA, an appropriately designed and structured PA plan to increase students' PA levels (in periods of time where doing PA is voluntary) seems necessary (Dobbins et al., 2013).

Strengths and limitations

To our knowledge, this is the first study that examines the effect of two alternated teaching units on the students' environmental knowledge for practicing out-of-school PA and their objectively measured PA. Furthermore, the accelerometer-measured PA improves the validity and quality of the results, taking a step ahead regarding previous studies carried out with self-reported measures. The use of a cluster-randomized and controlled research design also adds quality to this study. Finally, the evaluation of the effect of the program with a Mixed Multilevel Linear Model with participants nested within classes and measures nested within participants represents an advance with respect to the commonly applied analyses in previous literature (Li et al., 2017).

However, some limitations should also be considered. First of all, due to human, time, and material resource restrictions, a larger sample could not be examined, limiting the generalizability of the obtained outcomes to the particular studied setting. Moreover, due to some resource constraints such as time available for evaluation sessions, and participant's attention and patience during these sessions, the students' self-determined motivation in the educational context (i.e., a variable presented in the first and second proposition of the TCM) could not be measured in the present study. However, students' motivation towards PA in general, which is more directly linked to the ultimate goal of promoting actual daily PA in out-of-school contexts, was measured instead. Regarding the contents, the present innovative program was developed with physical fitness and traditional and alternative games and sports contents, which dominate the students' PA experiences globally in PE (Hardman et al., 2014) and could be the most applicable to students' free-time (individually or in small groups with friends), but these effects should also be studied with other PE contents. Moreover, the program length may have been a limitation to achieving greater effects on the objective PA. However, given the large volume of objectives that have to be developed throughout the school year with a very limited time for the PE subject (Hardman et al., 2014), the present study was adjusted to the mean length of PE teaching units. Finally, current results should be taken with caution due to the non-compliance rate in objectively measured PA variables, especially during weekend days. However, as analyses showed no differences in students' general characteristics in the valid subsample for each particular dependent variable, the non-compliance rate may not bias the findings of the present study.

Conclusion

The results of the present study showed that it is possible to improve students' knowledge about the environment for practicing PA, perceived autonomy, autonomous motivation towards PA, and the intention to be physically active. However, outcomes do not support the effectiveness of the program on actual objective PA, questioning its usefulness in promoting healthy PA habits if applied as it was. Further studies are necessary to deeply analyze this complex relationship between psychological variables and actual PA.

Acknowledgments

The authors want to thank all the participating students and the Physical Education teachers without whom the present study could not be carried out. We also thank Aliisa Hatten for the English revision of the manuscript. This work was supported by the Spanish Ministry of Universities [grant number: Carolina Casado-Robles, FPU16/03314; Santiago Guijarro-Romero, FPU15/02387].

Supplementary Files

Scan or click here to see Supplementary Files [Escanea el código o pincha aquí para ver los Archivos Suplementarios]

References

Bandura, A. (2004). Health promotion by social cognitive means. *Health education & behavior, 31*(2): 143-164. doi:10.1177/1090198104263660

Becker, C., Lauterbach, G., Spengler, S., Dettweiler, U., & Mess, F. (2017). Effects of regular classes in outdoor education settings: A systematic review on students' learning, social and health dimensions. *International Journal of Environmental Research and Public Health, 14*(5), 485. doi:10.3390/ijerph14050485

Biddle, S. J. H., Hagger, M. S., Chatzisarantis, N. L. D., & Lippke, S. (2007). Theoretical frameworks in exercise psychology. In G. Tenenbaum & R. C. Eklund (Eds.), Handbook of sport psychology (3rd ed., pp. 537–559). New York, NY: Wiley

Bølling, M., Otte, C. R., Elsborg, P., Nielsen, G., & Bentsen, P. (2018). The association between education outside the classroom and students' school motivation: Results from a one-school-year quasi-experiment. *International Journal of Educational Research, 89,* 22-35. doi:10.1016/j.ijer.2018.03.004

Campbell, M. K., Piaggio, G., Elbourne, D. R., & Altman, D. G. (2012). Consort 2010 statement: Extension to cluster randomised trials. *BMJ, 345:* e5661. doi:10.1136/bmj.e5661

Casado-Robles, C., Viciana, J., Guijarro-Romero, S., & Mayorga-Vega, D. (2021). Conocimiento del entorno para la práctica de actividad física en escolares (CEPAF): Desarrollo y validación de una prueba escrita objetiva de elección múltiple. *Journal of Sport and Health Research, 13*(2).

Chemolli, E., & Gagné, M. (2014). Evidence against the continuum structure underlying motivation measures derived from Self-Determination Theory. *Psychological Assessment, 26*(2): 575–585. doi:10.1037/a0036212

Chen, S., Chen, A., Sun, H., & Zhu, X. (2013). Physical activity and fitness knowledge learning in physical education: Seeking a common ground. *European Physical Education Review, 19*(2): 256-270. doi:10.1177/1356336X13486058

Cheon, S. H., Reeve, J., Lee, Y., & Lee, J. W. (2018). Why autonomy-supportive interventions work: Explaining the professional development of teachers' motivating style. *Teaching and Teacher Education, 69:* 43-51. doi:10.1016/j.tate.2017.09.022

Chinapaw, M. J., de Niet, M., Verloigne, M., De Bourdeaudhuij, I., Brug, J., & Altenburg, T. M. (2014). From sedentary time to sedentary patterns: accelerometer data reduction decisions in youth. *PLoS One, 9*(11): e111205. doi:10.1371/journal.pone.0111205

Deci, E. L., & Ryan, R. M. (2008). Facilitating optimal motivation and psychological wellbeing across life's domains. *Canadian Psychology, 49,* 14–23. doi:10.1037/0708-5591.49.1.14

Dobbins, M., Husson, H., DeCorby, K., & LaRocca, R. L. (2013). School-based physical activity programs for promoting physical activity and fitness in children and adolescents aged 6 to 18. *Cochrane Database of Systematic Reviews, 2*: CD007651.

González-Cutre, D., Sicilia, A., & Fernández, A. (2010). Hacia una mayor comprensión de la motivación en el ejercicio físico: medición de la regulación integrada en el contexto español. *Psicothema, 22*(4): 841-847.

Granero-Gallegos, A., Baena-Extremera, A., Pérez-Quero, F. J., Ortiz-Camacho, M. M., & Bracho-Amador, C. (2014). Validación española del «intention to partake in leisure-time physical activity». *Retos, 26*: 40-45.

Guthold, R., Stevens, G. A., Riley, L. M., & Bull, F. C. (2020). Global trends in insufficient physical activity among adolescents: a pooled analysis of 298 population-based surveys with 1.6 million participants. *The Lancet Child & Adolescent Health, 4*(1): 23-35. doi:10.1016/S2352-4642(19)30323-2

Hagger, M. S., & Chatzisarantis, N. L. (2016). The trans-contextual model of autonomous motivation in education: Conceptual and empirical issues and meta-analysis. *Review of Educational Research, 86*(2): 360-407. doi:10.3102/0034654315585005

Hardman, K., Murphy, C., Routen, A. C., & Tones, S. (2014). *UNESCO-NWCPEA: World-wide survey of school Physical Education*. París: United Nations Educational, Scientific and Cultural Organization.

Li, W., Xiang, P., Chen, Y. J., Xie, X., & Li, Y. (2017). Unit of analysis: Impact of Silverman and Solmon's article on field-based intervention research in Physical Education in the U.S.A. *Journal of Teaching in Physical Education, 36*(2): 131–141. doi:10.1123/jtpe.2016-0169

Mattocks, C., Ness, A., Leary, S., Tilling, K., Blair, S. N., Shield, J., ... & Wells, J. (2008). Use of accelerometers in a large field-based study of children: Protocols, design issues, and effects on precision. *Journal of Physical Activity & Health, 5*(1): S98-111. doi:10.1123/jpah.5.s1.s98

Mavropoulou, A., Barkoukis, V., Douka, S., Alexandris, K., & Hatzimanouil, D. (2018). The role of autonomy supportive activities on students' motivation and beliefs toward out-of-school activities. *The Journal of Educational Research, 112*(2): 223-233. doi:10.1080/00220671.2018.1503580

Migueles, J. H., Cadenas-Sanchez, C., Ekelund, U., Nyström, C. D., Mora-Gonzalez, J., Löf, M., ... & Ortega, F. B. (2017). Accelerometer data collection and processing criteria to assess physical activity and other outcomes: a systematic review and practical considerations. *Sports medicine, 47*(9): 1821-1845. doi:10.1007/s40279-017-0716-0

Moreno, J. A., Parra, N., & González-Cutre, D. (2008). Influencia del apoyo a la autonomía, las metas sociales y la relación con los demás sobre la desmotivación en educación física. *Psicothema, 20*(4): 636-641.

Poitras, V. J., Gray, C. E., Borghese, M. M., Carson, V., Chaput, J. P., Janssen, I., ... & Sampson, M. (2016). Systematic review of the relationships between objectively measured physical activity and health indicators in school-aged children and youth. *Applied Physiology, Nutrition, and Metabolism, 41*(6): S197-239. doi:10.1139/apnm-2015-0663

Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology, 61*: 101860. doi:10.1016/j.cedpsych.2020.101860

Saint-Maurice, P.F., & Welk, G. J. (2015). Validity and calibration of the youth activity profile. *PloS one, 10*(12): e0143949. doi:10.1371/journal.pone.0143949

Schneller, M. B., Duncan, S., Schipperijn, J., Nielsen, G., Mygind, E., & Bentsen, P. (2017). Are children participating in a quasi-experimental education outside the classroom intervention more physically active? *BMC Public Health*, *17*(1), 523. doi:10.1186/s12889-017-4430-5

Sheeran, P., Klein, W. M., & Rothman, A. J. (2017). Health behavior change: Moving from observation to intervention. *Annual review of Psychology, 68*: 573-600. doi:10.1146/annurev-psych-010416-044007

Spanish Ministry of Education, Culture, and Sport. (2015). Royal Decree 1105/2014, laying down the curriculum for Secondary Education and Baccalaureate. *Goverment Gazette, 3:* 480–486.

Teixeira, P. J., Carraça, E. V., Markland, D., Silva, M. N., & Ryan, R. M. (2012). Exercise, physical activity, and self-determination theory: a systematic review. *International Journal of Behavioral Nutrition and Physical Activity, 9*(1): 78. doi:10.1186/1479-5868-9-78

Trost, S. G., Loprinzi, P. D., Moore, R., & Pfeiffer, K. A. (2011). Comparison of accelerometer cut points for predicting activity intensity in youth. Medicine *Science and Sports Exercise, 43*(7): 1360-1368. doi:10.1249/MSS.0b013e318206476e

Viciana, J., & Mayorga-Vega, D. (2016). Innovative teaching units applied to Physical Education–changing the curriculum management for authentic outcomes. *Kinesiology, 48*(1), 142–152.

Viciana, J., & Mayorga-Vega, D. (2018). The three-axes model of planning in physical education. *Retos, 33*, 313–319.

Viciana, J., Mayorga-Vega, D., Martínez-Baena, A., Hagger, M.S., Liukkonen, J., & Yli-Piipari, S. (2019). Effect of self-determined motivation in physical education on objectively measured habitual physical activity: A trans-contextual model *Kinesiology, 51*(1), 140–145.

Wang, Y., & Chen, A. (2019). Two pathways underlying the effects of physical education on out-of-school physical activity. *Research Quarterly for Exercise and Sport,* 1-12. doi:10.1080/02701367.2019.1656325

World Health Organization. (2018). *Physical Activity Factsheets for the 28 European Union Member States of the Who European Region*. Copenhagen: World Health Organization

Yli-Piipari, S., Layne, T., Hinson, J., & Irwin, C. (2018). Motivational Pathways to Leisure-Time Physical Activity Participation in Urban Physical Education: A Cluster-Randomized Trial. *Journal of Teaching in Physical Education, 37*(2), 123–132. doi:10.1123/jtpe.2017-0099

Paper VI

PROMOCIÓN DE LA ACTIVIDAD FÍSICA DE LOS ESCOLARES MEDIANTE UNIDADES DIDÁCTICAS INTERMITENTES EN EDUCACIÓN FÍSICA

Casado-Robles, C., Guijarro-Romero, S., & Mayorga-Vega, D.

Innovación en la docencia e investigación de las Ciencias Sociales y de la Educación

2021, *In press*. ISBN: 978-84-1377-590-6

Paper VI

ABSTRACT

Physical Education subject is an ideal context for promoting physical activity practice among school-aged children. However, numerous limitations make the design and development of adequate physical activity promotion programs for students a complex task. For instance, the large volume of contents to be developed throughout the school year, the difficulty of controlling students' PA practice outside the school, or the low students' motivation towards physical activity. Therefore, it is necessary to seek feasible and effective strategies to overcome these constraints and awaken the students' motivation. For this purpose, the "intermittent teaching units" proposed by Viciana and Mayorga-Vega (2016) are based on devoting only a specific period of the Physical Education lesson (e.g., the cool-down part) to a particular curricular objective, leaving the rest of the class for another aim, proposing a solution to the limited curricular time of Physical Education subject. In addition, the activity wristbands are very popular among young people, positioning themselves as a valid strategy to be used in the educational context, increasing students' motivation, and giving the teacher the possibility of monitoring the students' physical activity outside the school. Consequently, this study develops a proposal for an intermittent teaching unit in Physical Education incorporating activity wristbands to promote students' PA practice and reduce SB in their leisure time.

KEYWORDS

Innovative teaching unit. Students. School-based program.

Consumer-wearable activity trackers.

Introducción

La práctica regular de actividad física, así como la reducción de conducta sedentaria, están relacionadas con numerosos beneficios sobre la salud de los niños en edad escolar (Organización Mundial de la Salud, 2020). Entre ellos, destacan beneficios en la salud física, mental y social, mejora de la calidad de vida, así como otros marcadores importante como el rendimiento académico y cognitivo (Organización Mundial de la Salud, 2020). Por este motivo, la Organización Mundial de la Salud (2020) ha establecido unas recomendaciones internacionales de actividad física y conducta sedentaria. Para los escolares, estas recomendaciones consisten en acumular, de media, un mínimo de 60 minutos diarios de actividad física de intensidad moderada a vigorosa, así como reducir el tiempo empleado en conductas sedentarias, especialmente el tiempo dedicado a la visualización de pantallas de forma recreativa (Organización Mundial de la Salud, 2020).

Desafortunadamente, los resultados de estudios previos constatan que la mayoría de niños en edad escolar a nivel mundial no cumplen con estas recomendaciones diarias de actividad física. Concretamente, la tasa de inactividad física a nivel mundial entre los niños en edad escolar se sitúa en el 81%, siendo entre los escolares españoles del 84% (Guthold *et al.*, 2020). Además, los niños en edad escolar emplean gran parte del tiempo del día en conducta sedentaria (Ruiz *et al.*, 2011). Por ejemplo, viendo la televisión, jugando a videojuegos o estudiando. Estos datos son preocupantes si tenemos en cuenta los numerosos efectos negativos de la inactividad física sobre la salud de los escolares, situándose como el cuarto factor de riesgo de mortalidad a nivel mundial, además de los elevados costes sanitarios que supone (más de 990 millones de euros en costes directos cada año) (International Sport and Culture Association, 2015).

Debido a esta evidencia, la promoción de unos niveles adecuados de actividad física y la reducción de la conducta sedentaria es una cuestión prioritaria de salud pública y un importante reto educativo a nivel mundial (Gillis *et al.*, 2013; Organización Mundial de la Salud, 2018). En concreto, España se ha comprometido a reducir al menos un 15% la prevalencia de inactividad física en adolescentes para el año 2030 (Organización Mundial de la Salud, 2018). Entre los diferentes contextos desde los que se puede llevar a cabo esta promoción de hábitos saludables entre los niños en edad escolar para disminuir las tasas de inactividad física, la escuela, por su carácter obligatorio y estar guiada por un profesional capacitado en el campo de la actividad física y la salud, es señalada como un contexto ideal (Organización Mundial de la Salud, 2018; Sevil-Serrano *et al.*, 2019). Dentro de la escuela, la asignatura de Educación Física es considerada como el mejor instrumento para ayudar al cumplimiento de las recomendaciones de práctica de actividad física de los escolares (Association for Physical Education, 2020; Comisión Europea/EACEA/Eurydice, 2013). De hecho, estudios previos han demostrado que los días que los escolares practican Educación Física presentan mayores niveles de actividad física que los días que no tienen estas clases, ayudando al cumplimiento de las recomendaciones internacionales (Viciana *et al.*, 2019; Mayorga-Vega *et al.*, 2018). Además, la Ley Orgánica para la Mejora de la Calidad Educativa (Ministerio de Educación, Cultura y Deporte, 2013) señala al profesorado de Educación Física como el encargado de promocionar la actividad física en las diferentes etapas educativas (Reales Decretos 126 y 1105 por el que se desarrollan los currículos básicos de Educación Primaria y Educación Secundaria para todo el territorio español, respectivamente).

Sin embargo, son numerosas las limitaciones en el ámbito educativo que impiden al profesorado de Educación Física llevar a cabo programas eficaces para promocionar la práctica de actividad física regular de los escolares. Entre ellas, se pueden destacar: (a) El gran número de contenidos y objetivos educativos a desarrollar en un curso académico; (b) el limitado tiempo curricular dedicado a la asignatura de Educación Física (en general, 2 sesiones a la semana en España); (c) la imposibilidad de controlar la actividad física del alumnado fuera del centro educativo, cuando el profesor/a no está presente; o (d) la motivación de los escolares para practicar actividad física en el tiempo de ocio (Comisión Europea/EACEA/Eurydice, 2013; Hardman *et al.*, 2014). Lamentablemente, esto provoca que se apliquen unidades didácticas tradicionales muy cortas e insuficientes para la promoción de la práctica de actividad física. Ante este panorama general, se hace necesario buscar estrategias en el ámbito educativo que sean viables y efectivas, así como que despierten la motivación de los escolares, para solucionar estos problemas y conseguir resultados significativos en la promoción de la práctica de actividad física en la edad escolar (Viciana y Mayorga-Vega, 2016).

Con el fin de superar todas estas limitaciones presentes en la asignatura de Educación Física, surgen las Unidades Didácticas Innovadoras propuestas por Viciana y Mayorga-Vega (2016). Estos autores se basan en los principios de la flexibilidad, eficacia y basada en objetivos de la planificación educativa (Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, UNESCO, 1962), y plantean una "distribución innovadora del tiempo" en función del objetivo educativo perseguido (Viciana y Mayorga-Vega, 2016). Entre ellas, se encuentran las Unidades Didácticas Intermitentes, las cuales se basan en dedicar únicamente una franja de tiempo de la sesión de Educación Física en vez de las sesiones completas al objetivo curricular propuesto dedicando el resto de la sesión a otro objetivo. Es decir, las sesiones tradicionales donde se dedica la sesión completa a un único objetivo curricular, se sustituyen por fragmentos (por ejemplo, de 5 a 15 minutos) en cada sesión durante un periodo prolongado de sesiones (por ejemplo, un trimestre académico). De este modo, los escolares alcanzan un aprendizaje progresivo, ya que trabajan sobre el mismo objetivo durante un número de sesiones mayor. Además, puede parecer que los escolares tienen un tiempo de práctica muy reducido dedicando únicamente una franja pequeña de la sesión, pero realmente, si se suman todos estos períodos cortos se está brindando a los escolares las mismas o más oportunidades de aprendizaje que una unidad didáctica tradicional.

Por otro lado, estudios previos han reconocido que, para provocar un cambio en la conducta de los escolares hacia el incremento de la actividad física, la monitorización autónoma del comportamiento activo (*real-time feedback* o *self-monitoring*) es una estrategia fundamental (Bort-Roig *et al.*, 2014; Michie *et al.*, 2013). Por este motivo, dentro de estas unidades didácticas innovadoras, puede incorporarse el uso de tecnologías portátiles de usuario o *wearables* (por ejemplo, los relojes inteligentes o las pulseras de actividad física) que aportan *feedback* en tiempo real a los escolares sobre sus hábitos de actividad física y conducta sedentaria. Estas tecnologías son cada vez más populares entre los escolares, viéndose reflejado en un incremento en el número de ventas en los últimos años (International Data Corporation, 2021). El empleo de estas nuevas tecnologías se está posicionando como una estrategia válida para su uso en el ámbito educativo, mediante la aplicación de programas de intervención que han conseguido aumentos en la motivación y en los niveles de práctica de actividad física de los escolares (por ejemplo, Duncan *et al.*, 2012; Evans *et al.*, 2017; Eyre *et al.*, 2015; Grao-Cruces *et al.*, 2016; Hardman *et al.*, 2011). Al mismo tiempo, estos dispositivos permiten solucionar algunas de las limitaciones que encontramos en la asignatura de Educación Física como, por ejemplo, dar respuesta a la imposibilidad de controlar la actividad física que realiza el alumnado fuera del centro educativo cuando el profesor/a no está presente. Por tanto, usarlos para el fomento de la actividad física en el tiempo de ocio de los escolares parece algo viable y deseable. Igualmente, para el diseño de estas intervenciones, además del empleo de las tecnologías portátiles de usuario para la auto-monitorización de los niveles de actividad física, deben tenerse en cuenta las variables psicológicas mediadoras de la práctica de actividad física, ya que son factores clave en cualquier cambio de comportamiento (Lubans *et al.*, 2016; Rhodes *et al.*, 2017). Entre ellas, la motivación de los escolares, la satisfacción de las necesidades psicológicas básicas (es decir, autonomía, competencia y relación) y la autoeficacia percibida son variables influyentes a tener en cuenta en este complejo proceso de adquisición de conductas activas saludables (Bandura, 2004; Ryan y Deci, 2020; Hagger y Chatzisarantis, 2016).

Consecuentemente, el objetivo del presente trabajo es presentar una propuesta de Unidad Didáctica Intermitente llevada a cabo durante la vuelta a la calma de las sesiones de Educación Física que incorpora el uso de las pulseras de actividad física, para para el fomento de la práctica de actividad física y la reducción de la conducta sedentaria en el tiempo de ocio de los estudiantes de Educación Primaria y Educación Secundaria Obligatoria.

Propuesta de unidad didáctica intermitente con pulseras de actividad física

La presente propuesta va dirigida al alumnado de último ciclo de Educación Primaria Obligatoria y cualquier curso de la etapa de Educación Secundaria Obligatoria. Sin embargo, sería especialmente útil realizarla con alumnado de 4º de Educación Secundaria Obligatoria, ya que en este curso es donde se solicita que el alumnado adquiera hábitos adecuados de actividad física regular de forma autónoma, lo que pueden conseguir con ayuda de las pulseras de actividad física (Ministerio de Educación, Cultura y Deporte, 2015).

La propuesta se desarrolla durante un trimestre académico completo (aproximadamente 10 semanas). Durante este tiempo, los escolares llevan una pulsera de actividad física como estrategia para aumentar la actividad física habitual y reducir la conducta sedentaria. Además, al mismo tiempo se aplica una Unidad Didáctica Intermitente durante la vuelta a la calma de las sesiones de Educación Física (últimos 10 minutos) para analizar el *feedback* de la información registrada por la pulsera de actividad física. Por otro lado, para diseñar esta Unidad Didáctica Intermitente se ha tenido en cuenta la incorporación de diversas técnicas de cambio comportamental (Michie *et al.*, 2013; Rhodes *et al.*, 2017). Igualmente, para el diseño de esta propuesta nos hemos basado en teorías psicológicas tales como la Teoría Cognitivo Social (Bandura, 2004), la Teoría de la Autodeterminación (Ryan y Deci, 2020), y el Modelo Trans-Contextual (Hagger y Chatzisarantis, 2016), para promover un cambio comportamental en los hábitos de actividad física.

Entre las estrategias que incorpora este programa intermitente, se pueden destacar: (a) Asesoramiento sobre el uso de las pulseras de actividad física, así como la explicación de los elementos fundamentales del *feedback* que aportan éstas (pasos, distancias recorridas o frecuencia cardiaca), la activación de alertas de conducta sedentaria y avisos de consecución del objetivo propuesto; (b) sesiones educativas sobre las recomendaciones de actividad física, beneficios de la práctica regular de actividad física, posibilidades de práctica saludables para el tiempo libre y actividades sedentarias a evitar; (c) establecimiento de metas diarias de actividad física (por ejemplo, 10.000 pasos diarios) y su registro diario en la aplicación móvil de la pulsera de actividad física; y (d) establecimiento de retos competitivos y seguimiento diario de la actividad física de los escolares mediante un blog educativo propio como estrategia de apoyo social del programa.

Además, de forma previa al inicio de la Unidad Didáctica el profesorado realizaría una reunión con los padres, madres y/o tutores legales del alumnado para explicar la importancia de esta Unidad Didáctica sobre la promoción de hábitos de actividad física y conducta sedentaria saludables en los escolares. En esta reunión, se les solicitaría la colaboración necesaria para que el alumnado lleve las pulseras de actividad física, así como, les permita la utilización de la aplicación móvil correspondiente y su participación en el blog educativo privado donde se compartirá información con el alumnado como estrategias de apoyo social y donde la familia podrá ver sus progresos y el material explicado. Para el alumnado que ya disponga de sus propias pulseras de actividad física, se tendrán en cuenta los diferentes modelos para atender a ellos en la unidad didáctica, y para aquellos que no tengan se les pediría a las familias su adquisición como material escolar.

A continuación, se muestra detalladamente el desarrollo de las 20 sesiones que podrían llevarse a cabo mediante una Unidad Didáctica Intermitente durante la vuelta a la calma (aproximadamente los 10 últimos minutos) de las sesiones de Educación Física. Cabe destacar, que además del objetivo principal de cada sesión, en todas las sesiones el profesorado debería interesarse por preguntas y dudas que pueda tener el alumnado sobre el funcionamiento de las pulseras. Así como, se insistiría en que lleven la pulsera de actividad física todos los días para tener un conocimiento real de sus hábitos de práctica.

Fase 1. Introducción al uso de las pulseras de actividad física y sus características principales (Sesiones 1 - 2)

Sesión 1

En primer lugar, se mostrarían las características principales de las pulseras de actividad física. Principalmente, se enseñaría cómo utilizar la pulsera y los diferentes menús disponibles en la pantalla, así como, la información más relevante que aporta respecto a los niveles de actividad física (es decir, número de pasos, minutos de actividad física, kilómetros recorridos, calorías gastadas, frecuencia cardíaca, etc.). Además, se explicaría al alumnado otras funcionalidades para que puedan personalizar la interfaz de la pulsera de actividad física según sus propios gustos o la activación de notificaciones, con el fin de aumentar la motivación del alumnado. Por último, se recordarían aspectos a tener en cuenta para su correcto cuidado y carga del dispositivo.

Sesión 2

Esta sesión se dedicaría a explicar el funcionamiento de la aplicación móvil asociada a la pulsera de actividad física, así como al procedimiento a seguir para sincronizar los datos de la pulsera de actividad física con el móvil. Para aquellos alumnos/as que no dispongan de teléfono móvil propio, en la reunión previa se les pediría a los padres/madres que les permitan el uso de su teléfono durante el desarrollo de la propuesta, con el fin de realizar un seguimiento más completo de sus niveles habituales de actividad física y conducta sedentaria. Dentro de las diversas funcionalidades de la aplicación móvil, la explicación se centraría en: (a) El establecimiento de objetivos de pasos o minutos de actividad física moderada-vigorosa a alcanzar diariamente; y (b) en las opciones de activación de alertas de inactividad (es decir, aviso sonoro o vibratorio cuando están demasiado tiempo en conducta sedentaria) y de cumplimiento de objetivo de actividad física (es decir, aviso sonoro o vibratorio al alcanzar el objetivo de actividad física establecido).

Fase 2. Creación y explicación de un blog educativo como estrategia de apoyo social (Sesión 3)

Sesión 3

Previo al desarrollo de esta sesión, el profesorado crearía un blog educativo para el desarrollo de esta unidad didáctica. En esta sesión se explicaría cómo utilizar el blog educativo, así como qué funcionalidades va a tener a lo largo de la unidad didáctica. En primer lugar, el blog es el lugar donde el profesor/a colgaría todos los elementos trabajados en clase (por ejemplo, el material sobre los beneficios de la práctica de actividad física, noticias relevantes o ejemplos de actividades que pueden practicar en su tiempo libre). Además, se utilizaría como medio para realizar recordatorios semanalmente al alumnado sobre: (a) No olvidar llevar la pulsera de actividad física puesta todo el día; (b) cargar su batería y sincronizar los datos de la pulsera de actividad física con la aplicación móvil regularmente; o (c) alcanzar el objetivo de actividad física propuesto para esa semana de pasos o minutos de actividad física. Por otro lado, el blog educativo se utilizaría como foro de discusión, en el cual los alumnos/as puedan subir fotografías de sus progresos (por ejemplo, una fotografía de la pulsera de actividad física con la meta conseguida, o una fotografía haciendo actividad física con sus compañeros/as en su tiempo libre). De este modo, el profesorado y los propios compañeros/as podrían dar *feedback* y felicitar al alumnado por sus logros.

Fase 3. Explicación de la importancia de los hábitos de vida saludables. Beneficios de la práctica de actividad física y perjuicios de la conducta sedentaria (Sesiones 4 - 6)

Sesión 4

En esta sesión, el alumnado ya llevaría más de una semana utilizando la pulsera de actividad física y recibiendo *feedback* inmediato de la cantidad de práctica de actividad física que están realizando, así como del tiempo que pasan en conducta sedentaria diariamente. Por lo tanto, durante esta sesión, se realizaría una ronda de preguntas general sobre los hábitos de práctica de actividad física del alumnado en su tiempo libre. Por ejemplo, preguntas acerca de qué tipo de actividades realizan, con qué frecuencia las practican, o cuántas horas pasan sentados diariamente. De este modo, se pretende generar un debate con el alumnado donde reflexionen acerca del tiempo que pasan en conducta sedentaria en comparación con el tiempo dedicado a la práctica de actividad física al día, y si creen que son niveles adecuados.

Sesión 5

Esta sesión se dedicaría a despejar las dudas generadas en el debate de la sesión anterior. Para ello, se explicarían las recomendaciones diarias de actividad física y conducta sedentaria establecidas por la Organización Mundial de la Salud (2020). Estas recomendaciones son, de media, realizar 60 minutos diarios de actividad física de intensidad moderada-vigorosa, así como limitar el tiempo en conducta sedentaria, especialmente el tiempo dedicado a la visualización de pantallas de forma recreativa. Además de este indicador de actividad física, se podrían explicar las recomendaciones diarias basadas en número de pasos, ya que es un indicador más fácil de entender para el alumnado (por ejemplo, 10.000 pasos al día; Mayorga-Vega *et al.*, 2021; Tudor-Locke *et al.*, 2011). Tras su explicación, se debería incidir en que el alumnado debe proponerse el cumplimiento de estas recomendaciones como un objetivo diario, el cual pueden comprobar con ayuda de la pulsera de actividad física si lo están consiguiendo o no. Para toda esta explicación, se utilizaría un póster educativo (Figura 1) que el profesorado imprimiría para colgarlo en el aula y en el gimnasio donde se realizan las clases de Educación Física. Además, después de esta sesión, el profesor/a subiría el póster al blog educativo para que los padres/madres y los propios alumnos/as en su tiempo libre puedan consultarlo.

Sesión 6

Durante esta sesión, siguiendo con el póster educativo utilizado en la sesión previa, el profesorado explicaría los beneficios que aporta la práctica regular de actividad física, así como los perjuicios de la conducta sedentaria, en su salud. Entre ellos, se destacarían los beneficios a nivel académico, ya que la propuesta se está llevando a cabo en el entorno escolar. Además, el profesorado propondría ejemplos de actividades físicas que el alumnado puede realizar y qué tipo de intensidad implican (es decir, ligera, moderada o vigorosa), para ayudar al alumnado a identificar la intensidad en función de la actividad practicada.

Figura 1. Póster educativo sobre recomendaciones y beneficios de la práctica de actividad física

Fase 4. Revisión y resolución de problemas (Sesiones 7 y 12)

Sesiones 7 y 12

Estas sesiones se dedicarían a revisar los hábitos del alumnado, así como a resolver problemas que se puedan estar encontrando a lo largo de la unidad didáctica. Por ejemplo, se resolverían dudas sobre la utilización de las pulseras o la aplicación móvil, o se comprobaría si el alumnado está cumpliendo con el objetivo de pasos y las actividades deportivas escogidas para realizar cada semana. Si los estuviesen cumpliendo, el profesor/a felicitaría e incidiría que hay que mantenerlos y seguir progresando. En caso contrario, el profesor/a preguntaría el motivo por el cuál no se está cumpliendo el objetivo propuesto y se debatiría sobre las barreras que el alumnado se puede encontrar para la práctica de actividad física (por ejemplo, no estar motivado para hacerlo, no tener tiempo, inclemencias meteorológicas, etc.) para ayudarles a superarlas. Además, se podría revisar la meta de pasos establecida en las sesiones previas, y si es muy difícil de alcanzar se podría considerar bajarla con el fin de individualizar la progresión al alumnado.

Fase 5. Establecimiento de metas personales (Sesiones 8 - 9)

Sesión 8

Esta sesión se dedicaría a explicar al alumnado la importancia de establecer una meta de actividad física, referida en concreto a un número de pasos diarios a alcanzar. Además, se consensuaría con el alumnado un plan adaptativo de pasos a cumplir cada semana. Para ello, se propondrían diferentes categorías según el número de pasos diarios que realizan actualmente (es decir, nivel basal o de partida) y se establecería una progresión hasta intentar alcanzar las recomendaciones diarias al finalizar la Unidad Didáctica Intermitente. Cabe destacar la importancia de que los alumnos/as estén de acuerdo en realizar ese plan e incluir todas las modificaciones que sean oportunas para individualizarlo lo máximo posible a las necesidades del alumnado.

Como se ha comentado previamente, las metas propuestas al alumnado deberían basarse en las recomendaciones (por ejemplo, 10.000 pasos diarios), pero a la vez ser realistas. Es decir, se debería establecer una meta asumible de alcanzar de forma progresiva, de forma que suponga un reto al alumnado pero que no sea demasiado difícil de alcanzar. También cabe resaltar que la meta propuesta debería establecerse sólo como un mínimo a superar, pero que cuantos más pasos realicen diariamente los escolares mayores serán los beneficios en su salud. En la Figura 2 se incluye una propuesta de progresión semanal en las metas establecidas.

Sesión 9

En esta sesión el profesorado explicaría cómo configurar en la aplicación móvil de la pulsera de actividad física la meta de pasos establecida para cada semana, según el plan explicado en la sesión previa. Además, se le pediría al alumnado que active la notificación sonora o vibratoria al alcanzar dicho objetivo de actividad física diario, para que los felicite tras su consecución como estrategia motivacional. Posteriormente, con el objetivo de conseguir el compromiso de los alumnos/as para alcanzar dichas metas, se les pediría que firmen el contrato "Mejoro mis niveles de actividad física". En dicho contrato, se detallaría el plan adaptativo de número de pasos para cada semana que el alumno/a se compromete a alcanzar durante toda la unidad didáctica, así como el profesor/a también se comprometería a apoyar y ayudar al alumnado en lo que sea necesario para alcanzar dicho objetivo.

Nivel de partida	Semana 5	Semana 6	Semana 7	Semana 8	Semana 9	Semana 10	Incremento total
5.000 pasos/día	6.000 (+1.000)	7.000 (+1.000)	8.000 (+1.000)	9.000 (+1.000)	10.000 (+1.000)	11.000 (+1.000)	+6.000
6.000 pasos/día	7.000 (+1.000)	8.000 (+1.000)	9.000 (+1.000)	9.500 (+500)	10.500 (+1.000)	11.000 (+500)	+4.500
7.000 pasos/día	8.000 (+1.000)	8.500 (+500)	9.000 (+500)	10.000 (+1.000)	10.500 (+500)	11.000 (+500)	+4.000
8.000 pasos/día	8.500 (+500)	9.000 (+500)	9.500 (+500)	10.000 (+500)	10.500 (+500)	11.000 (+500)	+3.000
9.000 pasos/día	9.500 (+500)	10.000 (+500)	10.000 (Igual)	10.500 (+500)	11.000 (+500)	11.000 (Igual)	+1.500

Figura 2. Progresión semanal de metas de pasos en función del nivel de partida

Fase 6. Entrega de material educativo con estrategias para incrementar la práctica de actividad física (Sesiones 10 - 11)

Sesión 10

En la presente sesión, el profesorado explicaría el decálogo "10 maneras para incrementar la actividad física y reducir el tiempo en conducta sedentaria". Este decálogo incluiría consejos para que el alumnado pueda incrementar sus niveles de actividad física y reducir el tiempo que emplean en conducta sedentaria con pequeños cambios en sus hábitos diarios. Este decálogo podría ser el siguiente:

- Utiliza los cambios de clase para levantarte de la silla y caminar algunos pasos antes de la llegada del próximo profesor/a.
- Utiliza las escaleras en lugar del ascensor siempre que puedas. Si ya lo haces, prueba a subir los escalones de dos en dos para aumentar la intensidad.
- Evita quedarte varias horas sentado o tumbado en una misma postura. Utiliza los avisos de la pulsera de actividad física o el teléfono móvil para levantarte frecuentemente, camina unos pocos pasos alrededor y estírate.
- Utiliza formas de desplazamiento activo para venir al centro educativo (por ejemplo, caminando o en bicicleta) en lugar de venir en transporte motorizado (por ejemplo, coche, motocicleta o patinete eléctrico).
- ¿Utilizas siempre el mismo camino para volver a casa? Aprovecha para variar de camino, realiza cambios en el ritmo de la marcha y busca, si las hay, pendientes que subir para incrementar la intensidad.
- Mientras estás sentado delante del ordenador o la televisión realiza ejercicios sencillos cada cierto tiempo (por ejemplo, levántate y siéntate en la silla 20 veces cada hora).
- Acompaña a tu familia a realizar la compra y carga las bolsas desde el supermercado hasta casa, te servirá como un entrenamiento de fuerza.
- Cambia la silla que utilizas para leer, ver la televisión o jugar a videojuegos por una pelota de estabilidad (fitball). Sentarte en una silla no requiere ninguna actividad muscular, pero sentarte en la pelota te obliga a contraer tus músculos para mantener el equilibrio.
- Cambia una charla por Whatsapp con tus amigos por una actividad deportiva al aire libre (por ejemplo, dar un paseo caminando o en bicicleta).

- ¿Has ordenado ya tu habitación? Levántate y limpia tu habitación al ritmo de tu música favorita.
- Al terminar la sesión, esta información podría ser colocada en el aula y el gimnasio donde se realicen las sesiones de Educación Física, así como subirla al blog educativo para su consulta.

Sesión 11

En esta sesión, el profesorado propondría una lista de actividades deportivas que el alumnado puede realizar para aumentar su actividad física, tanto de forma individual como en grupo, así como apuntándose a actividades extraescolares organizadas en su tiempo libre. Es importante que la lista no sea cerrada, sino que a modo de debate se vayan añadiendo las actividades que los alumnos/as propongan para que el profesorado se ajuste a sus intereses. Además, se le solicitaría al alumnado que cada semana escoja al menos dos o tres actividades de la lista que se ha ido creando colaborativamente, creándose ellos mismos un pequeño plan individualizado como elemento facilitador de la práctica de actividad física. La lista inicial propuesta por el profesorado podría incluir, entre otras, las siguientes opciones:

- Posibilidades de práctica de actividad física en el entorno local: Dar información de las actividades extraescolares del centro educativo o del área de deportes de la localidad.
- Proponer actividades orientadas a correr/caminar, pero también a formas jugadas.
- Proponer actividades que puedan realizar en casa con material casero, así como en el entorno próximo de su barrio.
- Proponer variedad de actividades, que sean adecuadas a diferentes niveles de competencia e interés.
- Proponer que busquen amigos con preferencias similares para que compartan un objetivo común, fomentando de este modo las relaciones sociales.

Al terminar la sesión, esta información podría ser colocada en el aula y el gimnasio donde se realicen las sesiones de Educación Física, así como subirla al blog educativo para su consulta.

Fase 7. Creación de una competición en clase (Sesiones 13, 15, 17 y 19)

Sesión 13

En esta sesión el profesorado plantearía un reto competitivo que puede estar basado en una estrategia de gamificación. Por ejemplo, se puede plantear el reto "¿En cuánto tiempo seréis capaces de recorrer España?". Para este reto se establecería el punto de partida (es decir, la localidad donde se encuentre el centro educativo) y se irían estableciendo ciudades clave, marcando la distancia en número de pasos entre dichas ciudades. Además, el reto sería colaborativo por grupos de cinco o seis alumnos/as. Como objetivo final, se establecería una ciudad de destino común para todos los grupos que implique realizar 10.000 pasos/día cada alumno/a durante el resto del programa (por ejemplo, Granada-Barcelona), para considerar ganador al grupo que más cerca se quede de la ciudad de destino. Una estrategia podría ser la siguiente:

- Granada-Jaén: 200.000 pasos → En grupos de 5 alumnos/as, realizando 10.000 pasos por día, alcanzaréis Jaén en 4 días.
- Granada-Madrid: 600.000 pasos → En grupos de 5 alumnos/as, realizando 10.000 pasos por día, alcanzaréis Madrid en 12 días.
- Granada-Barcelona: 1.000.000 pasos → En grupos de 5 alumnos/as, realizando 10.000 pasos por día, alcanzaréis Barcelona en 20 días.
- ....

Sesiones 15, 17 y 19

Para fomentar las relaciones sociales y cohesión del grupo creado para la competición, se realizaría un ranking de clase. Para ello, cada grupo con ayuda de las pulseras de actividad física recopilaría el número de pasos que ha hecho esa semana y se establecería una clasificación semanal para ver cuánto han avanzado y en qué ciudad se encuentran en cada momento, así como cuánta actividad física necesitarían hacer para alcanzar la ciudad de destino. En la sesión 19 (última sesión de esta fase) se establecería la clasificación final, para determinar el grupo ganador y conocer qué ciudad ha conseguido alcanzar cada grupo. Este ranking semanal y definitivo se colgaría en el blog educativo para la consulta por parte del propio alumnado y de las familias.

Fase 8. Visualización de vídeos y noticias relacionadas con actividad física y motivación (Sesiones 14 y 16)

Sesiones 14 y 16

De forma previa a la sesión, el profesorado colgaría en el blog educativo videos muy cortos (de 3 o 4 minutos de duración) y noticias relacionadas con la práctica de actividad física y la motivación, o el espíritu de superación. El profesor/a, les pediría a los alumnos/as que consulten este material antes de clase, para dedicar estas sesiones a debatir sobre ellos. Por ejemplo, se podrían incluir vídeos sobre personajes públicos que cambiaron sus hábitos de vida radicalmente (desde una vida sedentaria hasta una vida activa) o noticias sobre la alta tasa de inactividad física y obesidad que hay en la población española. Para generar debate el profesor/a podría realizar preguntas tales como: "¿Qué recomendarías a una persona sedentaria para qué cambie su estilo de vida?"; "Con respecto a las estrategias que hemos hablado para introducir la actividad física en tu vida diaria, ¿Sabes cómo organizarte para estudiar y también participar en actividad física todos los días?"; "¿Crees que los medios de comunicación tienen una influencia para qué hagas más o menos actividad física?

Fase 9. Comentarios sobre las publicaciones de actividad física colgadas en el blog educativo por el alumnado (Sesión 18)

Sesión 18

Esta sesión se dedicaría a comentar y dar *feedback* sobre las publicaciones que han ido realizando los alumnos/as a lo largo de toda la unidad didáctica en el blog educativo. En este momento el profesor/a intentaría dar *feedback* de tipo afectivo, se interesaría por los hábitos de actividad física diaria que han alcanzado los escolares, mostrando preocupación por las barreras que puedan encontrar durante su práctica y proponiendo soluciones, así como los animaría a seguir practicando y mejorando sus niveles diarios de actividad física en el tiempo libre. Por otro lado, también se les preguntaría acerca de lo que les está pareciendo la experiencia, así como qué incorporarían o eliminarían en futuras aplicaciones.

Fase 10. Conclusión de la propuesta (Sesión 20)

Sesión 20

La sesión final estaría destinada a hacer una recopilación de todos los contenidos trabajados en la unidad didáctica, con el fin de mantener las mejoras alcanzadas. Además, se realizaría la entrega de premios por la competición establecida o la consecución de objetivos individuales. Por ejemplo, esos premios podrían ser:

- Premio a los tres alumnos/as que han alcanzado su meta de pasos más días a la semana durante toda la unidad didáctica.
- Premio al equipo con mayor distancia recorrida (es decir, número de pasos) durante la competición
- Premio conjunto "Uno para todos y todos para uno" si todos los alumnos/as de clase superan la meta propuesta durante la última semana de la intervención.

Conclusiones

Como conclusión de este trabajo, podemos resaltar que con el desarrollo de esta propuesta de Unidad Didáctica Intermitente (Viciana y Mayorga-Vega, 2016) se pretende repercutir positivamente en los hábitos de práctica de actividad física y en la reducción de la conducta sedentaria de los escolares utilizando las pulseras de actividad física. Esta propuesta podría ayudar al profesorado de Educación Física a planificar y llevar a cabo programas efectivos para la consecución del objetivo educativo de incrementar la actividad física de los escolares en su tiempo libre fuera del centro educativo.

Agradecimientos

Carolina Casado-Robles recibe una ayuda FPU del Ministerio de Universidades (Referencia: FPU16/03314).

Referencias bibliográficas

Association for Physical Education, A. (2020). *Health Position Paper*. Association for Physical Education.

Bandura, A. (2004). Health promotion by social cognitive means. *Health Education & Behavior, 31*(2), 143–164. https://doi.org/10.1177/1090198104263660

Bort-Roig, J., Gilson, N. D., Puig-Ribera, A., Contreras, R. S., y Trost, S. G. (2014). Measuring and influencing physical activity with smartphone technology: a systematic review. *Sports Medicine, 44*(5), 671-686

Comisión Europea/EACEA/Eurydice. (2013). *Physical Education and sport at school in Europe Eurydice Report*. Oficina de Publicaciones de la Unión Europea

Duncan, M., Birch, S., y Woodfield, L. (2012). Efficacy of an integrated school curriculum pedometer intervention to enhance physical activity and to reduce weight status in children. *European Physical Education Review, 18*(3), 396-407. https://doi.org/10.1177/1356336X12450799

Evans, E. W., Abrantes, A. M., Chen, E., y Jelalian, E. (2017). Using Novel Technology within a School-Based Setting to Increase Physical Activity: A Pilot Study in School-Age Children from a Low-Income, Urban Community. *BioMed research international, 2017,* 4271483. https://doi.org/10.1155/2017/4271483

Eyre, E. L., Cox, V. M., Birch, S. L., y Duncan, M. J. (2016). An integrated curriculum approach to increasing habitual physical activity in deprived South Asian children. *European journal of sport science, 16*(3), 381–390. https://doi.org/10.1080/17461391.2015.1062565

Gillis, L., Tomkinson, G., Olds, T., Moreira, C., Christie, C., Nigg, C., Cerin, E., Van Sluijs, E., Stratton, G., Janssen, I., Dorovolomo, J., Reilly, J. J., Mota, J., Zayed, K., Kawalski, K., Andersen, L. B., Carrizosa, M., Tremblay, M., Chia, M., Hamlin, M., … Van Mechelen, W. (2013). Research priorities for child and adolescent physical activity and sedentary behaviours: an international perspective using a twin-panel Delphi procedure. *The international journal of behavioral nutrition and physical activity, 10,* 112. https://doi.org/10.1186/1479-5868-10-112

Grao-Cruces, A., Ruiz-López, R., Moral-García, J. E., Ruiz-Ariza, A., y J Martínez-López, E. (2016). Effects of a steps/day programme with evaluation in physical education on body mass index in schoolchildren 11-12 years of age. *Kinesiology, 48*(1.), 132-141. https://doi.org/10.26582/k.48.1.2

Guthold, R., Stevens, G. A., Riley, L. M., y Bull, F. C. (2020). Global trends in insufficient physical activity among adolescents: a pooled analysis of 298 population-based surveys with 1· 6 million participants. *The Lancet Child & Adolescent Health, 4*(1), 23-35. https://doi.org/10.1016/S2352-4642(19)30323-2

Hagger, M. S., y Chatzisarantis, N. L. (2016). The Trans-Contextual Model of Autonomous Motivation in Education: Conceptual and Empirical Issues and Meta-Analysis. *Review of educational research, 86*(2), 360–407. https://doi.org/10.3102/0034654315585005

Hardman, C. A., Horne, P. J., y Fergus Lowe, C. (2011). Effects of rewards, peer-modelling and pedometer targets on children's physical activity: a school-based intervention study. *Psychology and Health, 26*(1), 3-21. https://doi.org/10.1080/08870440903318119

Hardman, K., Murphy, C., Routen, A., y Tones, S. (2014). *UNESCO-NWCPEA: World-wide survey of school Physical Education.* United Nations Educational, Scientific and Cultural Organization.

International Data Corportation. (2021). *Consumer Enthusiasm for Wearable Devices Drives the Market to 28.4% Growth in 2020, According to IDC*. Recuperado el 19 julio 2021, de https://www.idc.com/getdoc.jsp?containerId=prUS47534521

International Sport and Culture Association. (2015). *The economic cost of physical inactivity in Europe. An ISCA/ Cebr report*. ISCA/ Cebr Office

Lubans, D. R., Smith, J. J., Morgan, P. J., Beauchamp, M. R., Miller, A., Lonsdale, C., Parker, P., y Dally, K. (2016). Mediators of Psychological Well-being in Adolescent Boys. *The Journal of adolescent health : official publication of the Society for Adolescent Medicine, 58*(2), 230–236. https://doi.org/10.1016/j.jadohealth.2015.10.010

Mayorga-Vega, D., Martínez-Baena, A., y Viciana, J. (2018). Does school physical education really contribute to accelerometer-measured daily physical activity and non sedentary behaviour in high school students? *Journal of sports sciences, 36*(17), 1913–1922. https://doi.org/10.1080/02640414.2018.1425967

Mayorga-Vega, D., Casado-Robles, C., López-Fernández, I., y Viciana, J. (2021) A comparison of the utility of different step-indices to translate the physical activity recommendation in adolescents. *Journal of Sports Sciences, 39*(4), 469-479. https://doi.org/10.1080/02640414.2020.1826667

Michie, S., Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., Eccles, M. P., Cane, J., y Wood, C. E. (2013). The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Annals of behavioral medicine: a publication of the Society of Behavioral Medicine, 46*(1), 81–95. https://doi.org/10.1007/s12160-013-9486-6

Ministerio de Educación, Cultura y Deporte. (2013). Ley 8/2013 Ley Orgánica para la Calidad de la Mejora Educativa. BOE, 295, de martes 10 de diciembre de 2013. *Boletín Oficial Del Estado, 295,* 97858-97921.

Ministerio de Educación, Cultura y Deporte. (2015). Real Decreto 1105/2014 de 26 de diciembre, por el que se establece el currículo básico de la Educación Secundaria Obligatoria y del Bachillerato. *Boletín Oficial Del Estado, 3,* 169–546.

Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO). (1962). *New methods and techniques in education. Report of a meeting of experts*. Recuperado el 19 julio 2021, de: https://unesdoc.unesco.org/ark:/48223/pf0000126329

Organización Mundial de la Salud (2020). *Guidelines on physical activity and sedentary behaviour*. Organización Mundial de la Salud.

Organización Mundial de la Salud. (2018). *Global action plan on physical activity 2018–2030: More active people for a healthier world*. Organización Mundial de la Salud

Rhodes, R. E., Janssen, I., Bredin, S. S. D., Warburton, D. E. R., y Bauman, A. (2017). Physical activity: Health impact, prevalence, correlates and interventions. *Psychology Health, 32,* 942–975. https://doi.org/10.1080/08870446.2017.1325486

Ruiz, J. R., Ortega, F. B., Martínez-Gómez, D., Labayen, I., Moreno, L. A., De Bourdeaudhuij, I., Manios, Y., Gonzalez-Gross, M., Mauro, B., Molnar, D., Widhalm, K., Marcos, A., Beghin, L., Castillo, M. J., Sjöström, M., y HELENA Study Group (2011). Objectively measured physical activity and sedentary time in European adolescents: the HELENA study. *American journal of epidemiology, 174*(2), 173–184. https://doi.org/10.1093/aje/kwr068

Ryan, R. M., y Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology, 61*, 101860. https://doi.org/10.1016/j.cedpsych.2020.101860

Sevil, J., García-González, L., Abós, Á., Generelo, E., y Aibar, A. (2019). Can High Schools Be an Effective Setting to Promote Healthy Lifestyles? Effects of a Multiple Behavior Change Intervention in Adolescents. *The Journal of adolescent health: official publication of the Society for Adolescent Medicine, 64*(4), 478–486. https://doi.org/10.1016/j.jadohealth.2018.09.027

Tudor-Locke, C., Craig, C. L., Beets, M. W., Belton, S., Cardon, G. M., Duncan, S., ... y Blair, S. N. (2011). How many steps/day are enough? for children and adolescents. *International Journal of Behavioral Nutrition and Physical Activity, 8*(1), 1-14. https://doi.org/10.1186/1479-5868-8-78

Viciana, J., y Mayorga-Vega, D. (2016). Innovative teaching units applied to physical education - Changing the curriculum management for authentic outcomes. *Kinesiology, 48*(1), 142–152. https://doi.org/10.26582/k.48.1.1

Viciana, J., Mayorga-Vega, D., y Parra-Saldías, M. (2019). Adolescents' physical activity levels on physical education and non-physical education days according to gender, age, and weight status. *European Physical Education Review, 25*(1), 143-155. https://doi.org/10.1177/1356336X17706683

EFFECTS OF CONSUMER-WEARABLE ACTIVITY TRACKER-BASED PROGRAMS ON OBJECTIVELY MEASURED DAILY PHYSICAL ACTIVITY AND SEDENTARY BEHAVIOR AMONG SCHOOL-AGED CHILDREN: A SYSTEMATIC REVIEW AND META-ANALYSIS

Casado-Robles, C., Viciana, J., Guijarro-Romero, S., & Mayorga-Vega, D.

Sports Medicine-Open

Under review

Paper VII

ABSTRACT

Objective: To estimate the effects of consumer-wearable activity tracker-based programs on daily objectively measured physical activity (PA) and sedentary behavior (SB) among school-aged children, as well as to compare the influence of participants' and programs' characteristics.

Methods: Eligibility criteria were: (1) Participants: apparently healthy school-aged children (< 18 years old); (2) Intervention: aimed to promote PA and/or to reduce SB incorporating consumer-wearable activity trackers; (3) Comparator: Baseline measurements and/or a control/traditional group; (4) Outcomes: objectively measured daily PA and/or SB levels; (5) Study design: pre-experimental, quasi-experimental, and true-experimental trials. Relevant studies were searched from eight databases up to December 2020, as well as from four alternative modes of searching. Based on the Cochrane Risk-of-bias tool 2, the risk of bias was assessed following four domains: (1) Randomization process; (2) missing outcome data; (3) measurement of the outcomes; and (4) selection of the reported results. Based on a comprehensive systematic review, meta-analyses of the Cohen's standardized mean difference (∂) and 95% confidence interval (CI) with a random-effects model were conducted to estimate the overall effects, as well as the within- and between-study moderator effects, of the programs on daily total steps, moderate-to-vigorous PA (MVPA), total PA and SB.

Results: 45 studies were included in the systematic review (5,620 unique participants; mean age = 12.85 ± 2.84 years) and 41 studies in the meta-analysis. Programs had a mean length of 11.78 ± 13.17 weeks and most used a waist-worn consumer-wearable activity tracker (77.78% waist-worn; 22.22% wrist-worn). Programs characteristics were: Goal-setting strategies (64.06%); participants' logbooks (56.25%); counseling sessions (62.50%); reminders (28.13%); motivational strategies (42.19%); and exercise routine (17.19%). Results showed a statistically significant moderate positive effect on daily total steps (∂ = 0.612, 0.477-0.746), small positive effect on daily MVPA (∂ = 0.220, 0.134-0.307), trivial positive effect on daily total PA (∂ = 0.151, 0.038-0.264) and trivial negative effect on daily SB (∂ = 0.172, 0.039-0.305). Subgroups analyses showed a higher effect for daily total steps and daily MVPA levels in females and the physically inactive for daily total steps (p = 0.003-0.044). Programs with educational counseling and/or goal-setting strategies, as well as a greater number of strategies, were more effective for improving children's daily total steps, and wrist-worn activity trackers were more effective than waist-worn trackers for improving their daily MVPA levels (p = 0.001-0.021).

Conclusions: Consumer-wearable activity tracker-based programs seem to be effective in promoting school-aged children's daily total steps and MVPA levels, especially for females and those that are physically inactive. These programs should include specific goal-setting, educational counseling, and wrist-worn trackers as especially effective strategies. However, due to the certainty of evidence being from "Low" to "Moderate", future well-designed primary research studies about the topic are needed. **Prospero:** CRD42020222363.

KEYWORDS

Wearable activity tracker. Activity monitor. Fitness tracker. Behavior change. Self-monitoring. Intervention. Goal-setting. Children. Adolescents.

1. Background

Childhood and adolescence are considered to be sensitive periods of life as they are the stages in which healthy lifestyle behaviors can be formed and become established [1]. Moreover, these behaviors could influence future adult health status and behavior [2]. Two important lifestyle behaviors during the school-aged children's awake time are physical activity (PA) (i.e., any waking behavior consisting of bodily movement that requires energy expenditure, which can be categorized into a *continuum* from light to vigorous intensity) [3,4], and sedentary behavior (SB) (i.e., any waking behavior characterized by a low energy expenditure) [4,5]. Both PA and SB are independent key indicators of health and quality of life among school-aged children [4]. Specifically, regular moderate-to-vigorous PA (MVPA) is positively associated with several health-related markers in young people [4]. Moreover, school-aged children's daily total PA levels (i.e., light-to-vigorous PA) and daily total steps have also been demonstrated to be potential health indicators [6]. Independently of daily PA levels, long periods of SB is considered a risky behavior associated with health problems in school-aged children [7]. The World Health Organization (WHO) [4] recommends that school-aged children should achieve, on average, at least 60 minutes daily of MVPA across the week. Although there is no evidence about a specific cut-off point, these guidelines also indicate that school-aged children should limit the amount of time spent being sedentary [4]. Unfortunately, nowadays worldwide most school-aged children (approximately 81%) are physically inactive (i.e., do not meet the PA recommendations) [8], and they spend most of their time in SB [9]. These levels are worrying because physical inactivity is one of the main risk factors for non-communicable disease and death worldwide, which makes the promotion of health-enhancing PA and SB levels (i.e., high levels of PA and low levels of SB) among young people a paramount priority public health challenge [10].

As a result of the high importance of this topic in public health, a large number of intervention programs have been carried out to promote school-aged children's health-enhancing PA and SB levels, which include a wide variety of behavior change techniques [11,12]. Among others, self-monitoring behavior is an essential technique for PA practice promotion [13], and consumer-wearable activity trackers (e.g., smartwatches, activity wristbands, or pedometers) could be ideal devices to track school-aged children's PA and SB levels, providing them this important real-time feedback [14]. Furthermore, these consumer-wearable activity trackers have become increasingly popular over recent years being reflected in a great increase in sales and strong demand from society worldwide every year [15]. Moreover, these devices are considered the most plausible activity monitors to be used in public health [16]. The popularity of these devices has led the scientific community to carry out effective studies that integrate consumer-wearable activity trackers promoting health-enhancing school-aged children's PA and SB levels [17,18].

Every single primary study about the effectiveness of consumer-wearable activity trackers for promoting PA and/or reducing SB levels only constitutes a specific portion of the total evidence. Moreover, conflicting findings can often be found between different primary studies (e.g., Eyre et al. [19] found a positive effect, while Lubans et al. [20] found a negative or null effect for improving PA levels; or Jago et al. [21] which found a positive effect in one wave, but a null effect in the other wave for reducing SB levels). By contrast, systematic reviews and meta-analyses allow for an objective analysis of all the available evidence about a specific topic, making sense of the often conflicting results found and providing estimation with greater power and precision than each individual primary study [22,23]. In this sense, systematic reviews and meta-analyses are considered the lens for appraising, synthesizing, and applying scientific evidence [24].

The number of reviews focused on the use of wearable activity trackers to increase PA and reduce SB levels has grown exponentially in recent years. However, most of them have been carried out in adults [25–27] or clinical populations [28,29], leaving little evidence on the effectiveness of wearable activity trackers in apparently healthy school-aged children. To our knowledge, only Ridgers et al. [30], Böhm et al. [31] and Cajita et al. [32] reviews included intervention studies carried out with apparently healthy school-aged children for promoting PA with consumer-wearable activity trackers. However, those reviews have considerable limitations. Firstly, regarding inclusion criteria, the three systematic reviews included restrictions on the language of publication (only in English) and they also exclude some grey literature sources (e.g., conference abstracts, dissertations, or pilot studies) which could bias the results due to interventions with significantly larger effects being more likely to be published in English and journal papers [33–35]. Furthermore, the Böhm et al.

[31] and Cajita et al. [32] reviews included date restrictions in the search (only from 2012 and 2009, respectively), and the Cajita et al. [32] review also presented a limitation regarding electronic databases selected, since they did not include main bibliographic databases (e.g., Web of Science, Scopus, or SPORTDiscus) [22,35]. These limitations imply that the systematic reviews were not carried out considering all the scientific evidence, and therefore, the estimation of the effect is not completely good.

Moreover, these systematic reviews included studies with self-reported measured PA levels (i.e., questionnaires). However, previous evidence has demonstrated that self-reported measures of children and adolescents are poorly correlated with objectively measured PA and SB levels [36,37]. Lastly, both the Ridgers et al. [30] and Böhm et al. [31] systematic reviews did not include pedometer-based studies as they did not consider them as consumer-wearable activity trackers. However, following common consumer-wearable activity trackers' scientific definitions, pedometers should be included because they are electronic devices that can be worn on the body as an accessory [38], and they are used for monitoring and recording daily PA [28], providing users real-time feedback on their PA levels via the monitor display [39]. All the above-mentioned reasons could explain the limited evidence found (i.e., low number of publications) in their reviews, due to the inclusion of only five [30,31] or two studies in their analyses [32]. Moreover, it should be noted that Cajita et al. [32] conducted only a scoping review, whose objective was only to provide an overview of the available research evidence without producing a summary answer to a discrete research question [40]. Finally, as far as we know, there is no previous meta-analysis examining the effects of consumer-wearable activity tracker-based programs on daily objectively assessed PA and SB among school-aged children. Conducting a meta-analysis could summarize the effectiveness of those interventions in an overall statistical synthesis (effect size) rather than taking the results of each primary study separately. Therefore, it allows for improving the precision of the results by the estimation of the effect size and direction, and knowing whether or not the effect size is consistent across studies. Thus, also based on study design and risk of bias assessment, the meta-analysis allows for assessing the strength of the evidence [35,41].

Consequently, the main purpose of the present systematic review and meta-analysis was to estimate the effects of consumer-wearable activity tracker-based programs on daily objectively measured PA and SB among apparently healthy school-aged children. A secondary purpose was to compare the influence of school-aged children's characteristics (i.e., gender, age group, and PA status) and the intervention programs characteristics (i.e., duration, worn-type activity tracker, goal-setting, diary, counseling, reminders, motivational strategies, and exercise) on the effects of the consumer-wearable activity tracker-based programs on daily objectively measured PA and SB.

2. Methods

2.1. Protocol and registration

The review protocol was registered with the International Prospective Register for Systematic Reviews (PROSPERO, CRD42020222363, https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=222363). The present systematic review and meta-analysis was based on the methodology described in previous reference literature such as the PRISMA guidelines [42] (see Supplementary File 1 for the PRISMA checklist) and Cochrane Handbook for Systematic Reviews of Interventions [35], among other important references [22,41]. Firstly, a reproducible, transparent, and comprehensive systematic review (qualitative synthesis) was performed to identify, select, and synthesize all the relevant studies. Then, a meta-analysis (quantitative synthesis) was performed to provide more precise estimates of the effects than those derived from the primary studies.

2.2. Eligibility criteria

The eligibility criteria for including the retrieved studies in the systematic review were the following: (1) Participants: apparently healthy children and adolescents (i.e., populations with diagnosed diseases/conditions were excluded) aged under 18 years old (although the study was also included if the participants were up to 24-years-old but the average age was under 18, as they are considered "young people" by the WHO [43]); (2) Intervention: studies that examined the effect of programs to promote PA and/or to reduce SB incorporating consumer-wearable activity trackers (i.e., pedometers, smartwatches, fitness wristbands, or similar;

smartphones applications were not included) alone or combined with other strategies (i.e., goal-setting, diaries, counseling, reminders, motivational strategies, or exercise routine) with a duration of at least three weeks were included [44]; (3) Comparator: The design must include at least one measure to compare intervention effects (i.e., baseline measurements in single-group designs and/or a control group with no intervention or with usual treatment without consumer-wearable activity trackers). Thus, interventions with only one experimental group and only post-intervention measures were not included; (4) Outcomes: studies that assessed the effect of the programs on the objectively measured daily PA and/or SB levels were included (i.e., objective measures such as the same consumer-wearable activity trackers used in the intervention or research-grade activity trackers such as accelerometers, excluding self-reported measures). Furthermore, only whole day (awake) time-based timeframes such as a whole week, weekdays, and/or weekend days were included (i.e., part-time days such as Physical Education lessons, school recess, or leisure time were not included); (5) Study design: Any kind of experimental designs including, but not limited to, pre-experimental trials (e.g., non-controlled trials with one-group pre-post-intervention design); quasi-experimental trials (e.g., cluster-randomized controlled trials or non-randomized controlled trials), and true-experimental trials (e.g., randomized controlled trials) were included.

2.3. Data sources and search strategy

The databases search following the search strategies and the download and collection in the reference manager was completed in December 2020. This search included the following eight electronic bibliographic databases: Web of Science™ (all databases), Scopus, PubMed, SPORTDiscus with Full Text, CINAHL, Cochrane Library, ProQuest Social Sciences Premium Collection, and ProQuest Dissertations & Theses Global™. The searches were carried out in the search field type "Title, abstract, and keywords" or equivalent. The search terms used were based on three concepts: (1) consumer-wearable activity tracker; (2) intervention program, and (3) PA/SB. The terms of the same concept were combined with the Boolean operator "OR" and then the three concepts were combined using the Boolean operator "AND". The keywords with more than one word were enclosed in quotes. No publication status, language, or date restrictions were imposed [22]. All search strategies are available in Supplementary File 2.

Then, additional studies were identified as follows (i.e., "snowballing"): (1) Searching the reference lists of original studies, as well as some related study reviews and study protocols; (2) examining the reference citations and the researchers' publications (first authors) in the Web of Science™ and Scopus databases; (3) contacting with the corresponding authors by email, and (4) screening the researchers' personal lists in ResearchGate and Google Scholar (first authors). Any time a new study was found, all of these modes of searching were repeated until no new study appeared.

2.4. Study selection

After eliminating duplicates, the first author (CCR) undertook the potentially eligible records selection based on the screening of titles and abstracts to identify relevant studies. After that, two independent reviewers assessed the full texts for inclusion following the above-mentioned eligibility criteria (CCR/SGJ). Any disagreements regarding the inclusion of studies were resolved by consensus with a third reviewer (DMV). The inter-rater agreement between coders was substantial to almost perfect (proportion of agreement = 0.89; Cohen's Kappa = 0.78).

2.5. Data extraction

From each selected study, data were coded using an *ad hoc* coding form developed by the research group and previously tested with a pilot sample of studies at the beginning of the review. This form included data about: (1) Study characteristics (i.e., reference, publication date, date of the data collection, study design, sequence generation, suspicion of selective outcomes, and initial and final sample size); (2) participant characteristics (i.e., gender and age); (3) outcome measures pertaining to PA and/or SB (i.e., measurement moment, measurement time, kind of measurement instrument, and measurement score and units); (4) intervention characteristics (intervention length, kind of consumer-wearable activity tracker, kind of goal-setting, diary, counseling, reminders, motivational strategies, and exercise routine); and (5) results of the intervention for each group (i.e., initial and final group size, pre- and post-intervention standard deviation, and pre- and post-intervention means score or pre-post-intervention mean difference score). Complete coding form is available in Supplementary File 3.

If a study consisted of two or more study arms of which one of the intervention arms did not meet the inclusion criteria, data were only extracted from the study arms that met the inclusion criteria. In the event that the studies did not report some study feature, corresponding authors were contacted to retrieve it. If means and standard deviation were not retrieved, the scores were estimated and converted by the standard error, confidence intervals, F, t or p values [35]. Since median and interquartile range are often used when the data is asymmetrical, these values were not converted [35]. If any other study feature was not retrieved, the information was omitted. The sample size of each group, the mean scores of the pre- and/or post-intervention or mean difference scores of each group, and the measurement score of the dependent variable/s were considered to be critical for including the selected studies (qualitative synthesis) in the meta-analysis (quantitative synthesis). In order to avoid removing studies from the quantitative synthesis, numerical data were extracted from their figures using the WebPlotDigitizer software [35], as was done with three studies [45–47]. Coding studies were carried out independently by two researchers (CCR/DMV). When doubt or disagreement occurred, a consensus was always achieved through discussion.

2.6. Risk of bias and certainty of the evidence

Based on the Cochrane Risk-of-bias tool version 2 [35], the following methodological domains were assessed: (1) Risk of bias arising from the randomization process; (2) risk of bias due to missing outcome data; (3) risk of bias in measurement of the outcomes, and (4) risk of bias in the selection of the reported results. Due to the nature of the selected studies (i.e., self-monitoring interventions to promote healthy habits with objective measurements), allocation concealment, and blinding of participants, personnel and outcomes assessment were not considered. Supplementary File 4 shows the algorithms followed for assessing methodological risk of bias in each domain.

Domains were judged and classified as “Low risk”, “Some concerns” or “High risk” of bias. Finally, overall risk of bias judgment was obtained as follows: “Low risk” if low risk of bias was obtained for all domains; “Some concerns” if at least one domain was judged as having some concerns, but not to be at high risk of bias for any domain; and "High risk" if at least one domain was judged as high risk or if two or more domains were judged as having some concerns [35].

Additionally, the overall certainty of the evidence was rated as "High", "Moderate", "Low", or "Very low" using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach [48]. This assessment was based on the following five domains: Risk of bias, inconsistency, imprecision, indirectness, and publication bias. A domain was classified as "No limitation" if no reason for downgrading the evidence was found, but the domain was classified as "Serious" if a reason was found for downgrading the evidence (thus, downgrading the certainty rating by one level). The reasons for considering the domains as "Serious" were: (a) Risk of bias: Most of the studies (i.e., > 50%) were classified as "High" risk of bias; (b) Inconsistency: High level of heterogeneity (i.e., I^2 > 75%) was found; (c) Imprecision: The confidence interval was wide including both the possibility of trivial effect (i.e., ∂ < 0.200) and large effect (i.e., ∂ ≥ 0.800); (d) Indirectness: Most studies (i.e., > 50%) addressed a restricted version of the main review question in terms of population, intervention, comparator or outcomes; and (e) Publication bias: Egger's test was statistically significant and the impact of publication bias was large (i.e., the number of additional studies with no effect that would be needed to increase the ∂ value was low following the fail-safe N analyses; or Trim and Fill method trimmed many studies with an adjusted value for effect size different to the observed values).

2.7. Data analyses

The meta-analyses were performed using the software Comprehensive Meta-Analysis version 3.3.070 for Windows (Biostat, Englewood, USA). The significance level was set at $p < 0.05$.

If a single study reported data for the whole sample and separately by different subsamples (e.g., children and adolescents), only the whole sample was used. Moreover, when in the same study there were different options for the same outcome (e.g., data from consumer-wearable activity trackers and research-grade activity trackers, or different measurement moments during the program) only the best option was selected (e.g., research-grade activity tracker or the last measurement as post-intervention). When a study had more than one activity tracker-based intervention group, each group was included in the analysis individually. The studies carried out with a small sample (defined as less than 10 participants per group) were not included in the meta-analysis [49].

2.7.1. *Effects sizes computation*

A detailed description of the data analyses carried out in the present meta-analysis can be found elsewhere [41]. Meta-analyses of the Cohen's d standardized mean difference and 95% confidence interval with a random-effects model were conducted to obtain the intervention program effects on: (a) Daily total steps; (b) MVPA; (c) Total PA; and (d) SB. Moreover, since daily total steps and MVPA in minutes work with meaningful scales, the Mean Difference (D) with a random-effects model was also conducted [35,41]. A positive d and D value referred to the program increasing participants' daily PA and SB.

2.7.2. *Publication bias*

Firstly, an exhaustive systematic review was carried out to avoid availability bias. Afterward, a deep examination of the selected studies was carried out to avoid any potential duplication of the information retrieved. Similarities between publications of the same authors, with the same values and/or the same sample size were examined. When the selected publications had full or partial duplicated information, these particular values were not analyzed. Then, to visually identify the impact of any potential publication bias, the funnel plots and the Egger's test [50] were carried out for daily total steps and MVPA. Moreover, for assessing the impact of any potential publication bias, the Orwin's fail-safe N analyses [51] (criterion for a "trivial" d = 0.100; mean d in missing studies, d = 0.000) [52], the Duval and Tweedie's Trim and Fill method [53] (assuming missing studies in the left of the mean), and a cumulative meta-analysis sorted by larger study size was computed [41]. Due to the limited number of studies found for total PA and SB (k = 8), the publication bias analyses could not be carried out for these variables [49].

2.7.3. *Heterogeneity and moderator analyses*

The presence of statistical heterogeneity in the estimation of the effect sizes of the program was examined with the I^2 statistic. The thresholds for its interpretation were: Values up to 40% were considered not important, up to 75% moderate, and more than 75% high heterogeneity [35].

Based on *a priori* hypothesized moderators, subgroups analyses were also carried out to test the effect of the intervention regarding: (a) Individuals' characteristics (i.e., gender, age, and PA status); and (b) intervention program characteristics (i.e., duration, activity tracker type, goal-setting, kind of goal-setting, diary, counseling, reminders, motivational strategies, exercise routine, and number of behavior change strategies). Due to the limited number of studies found for total PA (k = 8) and SB (k = 8), the above-mentioned subgroups analyses were only carried out for the daily steps and MVPA variables following a partial hierarchical analysis approach. All subgroups analyses were carried out for between-study meta-analysis, while for within-study meta-analysis, only those with at least two units of analysis to compare with were performed. Finally, the influence of continuous covariates (i.e., age, PA status, duration, and number of behavior change strategies) on the intervention effect was also evaluated using meta-regression analyses.

2.7.4. Sensitivity analysis

Finally, in order to evaluate the robustness of the main results, the following sensitivity analyses were performed: Cohen's d with a fixed-effect model, Hedges' g with a random-effects model, and Cohen's d with a random-effects model separately for randomized controlled trial design or not. However, sensitivity analyses separately for studies classified by the overall risk of bias could not be carried out due to the low number of studies classified as "Low risk" (k = 1 for daily total steps; and k = 2 for MVPA, total PA, and SB).

3. Results

3.1. Study selection

Figure 1 shows the flow diagram of the study selection process. The search strategy identified 39,625 potentially relevant studies (19,864 studies after removing duplicates). Afterward, as a result of the studies of the Boolean-based database search, 183 additional records were identified through other sources. From the 1,498 records retrieved for a more detailed evaluation, 47 publications were a priori selected for meeting the selection criteria. However, only 44 unique publications were finally included in the qualitative synthesis to avoid duplicated information. Finally, due to carrying out the study with a small sample or not reporting the critical values, only 40 publications were included in the meta-analysis (quantitative synthesis).

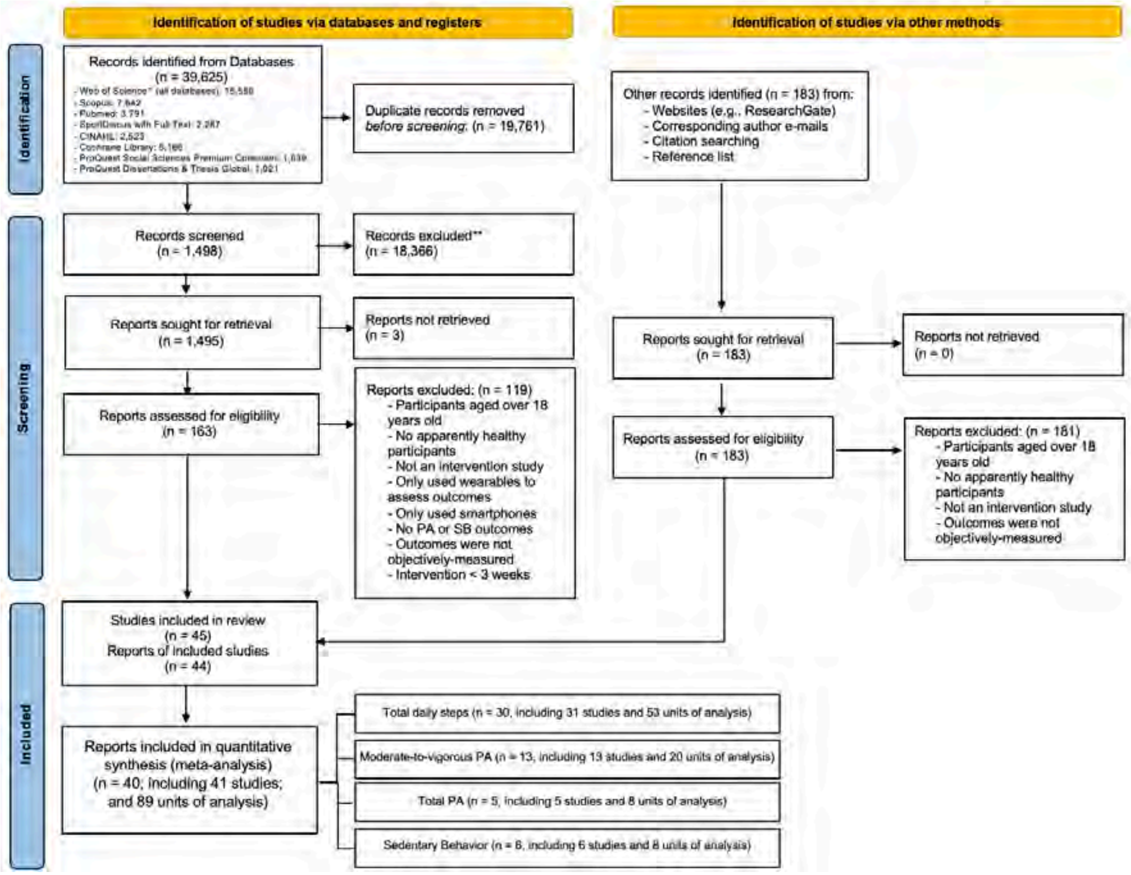


Figure 1. Flow diagram of the study selection process. PA = Physical activity. *Note*. The sum of publications included in the meta-analysis for each variable is greater than 40 because some studies reported outcomes for different variables.

3.2. Study characteristics

Table 1 presents a summary of study characteristics included in the qualitative synthesis. Across the 45 included studies (i.e., 44 publications), in total 5,620 unique participants were included in all studies (3,914 = intervention; 1,706 = control). Samples were composed of a median of 84.50 initial participants, ranging from 20 [54] to 496 [55]. The participants' mean age was 12.85 ± 2.84 years (range: 5-18 years). Regarding gender, 32 studies were conducted with males and females, while eight studies included only females and five studies included only males.

Table 1. Characteristics of included studies in the systematic review.

Study	Participants characteristics			Study design	Intervention characteristics						
	Number of participants[a]	Age[b]	Gender		Tracker	Duration[c]	Group	Goal-setting	Diary	Counseling	Re
Aksornsri [62][e]	EG1: 40 (39)	11.4 (8-13)	Males and females	Non-CT	Wrist-worn	4	EG1	Static goal	No	Yes	Y
Baldursdottir [63][e]	CG: 57 (54); EG1: 61 (57); EG2: 58 (57)	16.0 (15-16)	Males and females	Cluster-RCT	Hip-worn	3	EG1	Adaptive goal	Yes	Yes	N
							EG2	Adaptive goal	No	Yes	N
Baldursdottir et al. [56][e]	CG: 46 (27); EG1: 38 (26)	16.0 (16-16)	Males and females	Cluster-RCT	Hip-worn	3	EG1	Adaptive goal	Yes	Yes	Y
Bronikowski et al. [17][e]	EG1: 100 (100); EG2: 93 (93)	14.8 (10-17)	Males and females	Non-CT	Wrist-worn	8	EG1	Static goal	No	No	N
							EG2	No goal	No	No	N
Corepal et al. [64][e]	CG: 82 (64); EG1: 142 (126)	13.0 (12-14)	Males and females	Cluster-RCT	Hip-worn	22	EG1	Adaptive goal	Yes	Yes	N
Corr et al. [65][e]	EG1: 31 (17)	14.0 (12-16)	Females	Non-CT	Hip-worn	6	EG1	Adaptive goal	Yes	Yes	Y
Corr & Murtagh [66][e]	EG1: 31 (10)	16.0 (15-17)	Females	Non-CT	Hip-worn	6	EG1	Adaptive goal	Yes	Yes	N
Duck et al. [67][e]	CG: 18 (14); EG1: 17 (13)	9.2 (9-10)	Males and females	Cluster-RCT	Wrist-worn	10	EG1	No goal	No	No	N
Duncan et al. [68][e]	EG1: 59 (59)	10.7 (10-11)	Males and females	Non-CT	Hip-worn	4	EG1	Static goal	Yes	Yes	N
Ermetici et al. [57][e]	EG1: 181 (167)	12.5 (11-15)	Males and females	Non-CT	Hip-worn	84	EG1	No goal	No	Yes	Y
Evans et al. [69][e] - Study 1	EG1: 32 (32)	10.0 (10-11)	Males and females	Non-CT	Hip-worn	4	EG1	Static goal	Yes	Yes	N
Evans et al. [69][e] - Study 2	CG: 10 (10); EG1: 19 (19); EG2: 13 (13)	12.3 (11-12)	Males and females	Non-RCT	Wrist-worn	6	EG1	Static goal	Yes	Yes	N
							EG2	No goal	No	No	N
Eyre et al. [19][e]	CG: 40 (30); EG1: 94 (55)	10.0 (9-12)	Males and females	Cluster-RCT	Hip-worn	6	EG1	Static goal	Yes	Yes	N
Finkelstein et al. [70][e]	CG: 138 (89); EG1: 147 (145)	8.2 (6-12)	Males and females	Cluster-RCT	Hip-worn	36	EG1	Static goal	No	Yes	N
Galy et al. [71][e]	EG1: 24 (15)	11.9 (12-14)	Males and females	Non-CT	Wrist-worn	4	EG1	No goal	No	Yes	N
Gaudet et al. [45][e]	CG: 23 (16); EG1: 23 (16)	13.0 (13-14)	Males and females	Cluster-RCT	Wrist-worn	7	EG1	No goal	No	No	N
Grao-Cruces et al. [72][e]	EG1: 142 (66)	11.4 (11-12)	Males and females	Non-CT	Hip-worn	6	EG1	Static goal	Yes	No	N
Groffik et al. [73][e]	EG1: 71 (64)	17.9 (17-18)	Males and females	Non-CT	Hip-worn	3	EG1	No goal	Yes	No	N

Study	Participants characteristics			Study design	Intervention characteristics						
	Number of participants[a]	Age[b]	Gender		Tracker	Duration[c]	Group	Goal-setting	Diary	Counseling	Re
Guagliano et al. [18][e]	CG: 25 (23); EG1: 24 (18); EG2: 24 (15)	9.3 (7-11)	Males and females	Cluster-RCT	Hip-worn	8	EG1	No goal	No	Yes	N
							EG2	Adaptive goal	Yes	Yes	Y
Hardman et al. [74]	CG: 18 (15); EG1: 14 (14)	10.6 (10-11)	Females	Cluster-RCT	Hip-worn	12	EG1	Adaptive goal	Yes	No	N
Hardman et al. [75][e]	EG1: 81 (51); EG2: 119 (67); EG3: 186 (118)	9.0 (7-11)	Males and females	Non-CT	Hip-worn	17	EG1	No goal	No	No	N
							EG2	Adaptive goal	Yes	Yes	Y
							EG3	Adaptive goal	Yes	Yes	Y
Horne et al. [46][e]	CG: 53 (51); EG1: 47 (38)	10.0 (9-11)	Males and females	Cluster-RCT	Hip-worn	16	EG1	Adaptive goal	Yes	Yes	Y
Jago et al. [21][e] - Phase 1	CG: 55 (51); EG1: 65 (64)	13.0 (10-14)	Males	Cluster-RCT	Hip-worn	9	EG1	Static goal	No	Yes	N
Jago et al. [21][e] - Phase 2	CG: 130 (88); EG1: 103 (86)										
Jauho et al. [76]	CG: 139 (38); EG1: 137 (25)	17.9 (17-18)	Males	RCT	Wrist-worn	12	EG1	No goal	No	No	N
Kantanista et al. [77][e]	EG1: 26 (26); EG2: 56 (56)	17.2 (16-18)	Females	Non-CT	Hip-worn	7	EG1	No goal	No	No	N
							EG2	Adaptive goal	No	No	N
Kerner et al. [78][e]	EG1: 62 (28)	14.5 (14-15)	Males and females	Non-CT	Wrist-worn	5	EG1	No goal	No	No	N
Larsen et al. [79][e]	EG1: 21 (19)	14.7 (12-18)	Females	Non-CT	Hip-worn	12	EG1	No goal	Yes	Yes	Y
Leinonen et al. [55][e]	CG: 246 (80); EG1: 250 (87)	17.8 (17-18)	Males	RCT	Wrist-worn	26	EG1	Static goal	No	Yes	N
Linck [80][e]	CG: 22 (17); EG1: 22 (18)	16.6 (14-18)	Females	RCT	Hip-worn	12	EG1	Adaptive goal	Yes	No	N
Lubans & Morgan [81][e]	CG: 66 (52); EG1: 50 (45)	14.2 (14-15)	Males and females	Cluster-RCT	Hip-worn	8	EG1	Adaptive goal	Yes	Yes	N
Lubans et al. [82][e]	CG: 66 (52); EG1: 58 (50)	14.1 (13-18)	Males and females	Cluster-RCT	Hip-worn	24	EG1	Adaptive goal	Yes	Yes	Y
Lubans et al. [20][e]	CG: 50 (50); EG1: 50 (50)	14.3 (14-15)	Males	Cluster-RCT	Hip-worn	24	EG1	Adaptive goal	Yes	Yes	N
Macias-Cervantes et al. [83]	CG: 38 (30); EG1: 38 (32)	7.8 (6-9)	Males and females	RCT	Hip-worn	12	EG1	Adaptive goal	No	Yes	N
Morris et al. [84][e]	CG: 72 (31); EG1: 82 (52)	9.9 (9-10)	Males and females	Cluster-RCT	Hip-worn	6	EG1	Adaptive goal	No	No	N
Newton et al. [85][e]	EG1: 14 (14); EG2: 13 (13)	8.7 (6-10)	Males and females	Non-CT	Hip-worn	12	EG1	Adaptive goal	No	No	N
							EG2	Adaptive goal	No	Yes	Y

Remmert et al. [54]	EG1: 10 (6); EG2: 10 (9)	12 (12-12)	Males and females	Non-CT	Wrist-worn	12	EG1	No goal	No	No	N
							EG2	Adaptive goal	No	Yes	N
Routen et al. [47][e]	EG1: 17 (14); EG2: 26 (20); EG3: 25 (16)	11.2 (11-12)	Males and females	Non-CT	Hip-worn	3	EG1	No goal	No	No	N
							EG2	Adaptive goal	Yes	No	N
							EG3	Adaptive goal	Yes	No	N
Schofield et al. [86][e]	CG: 30 (24); EG1: 27 (23); EG2: 28 (21)	15.8 (15-18)	Females	Cluster-RCT	Hip-worn	12	EG1	Adaptive goal	Yes	Yes	Y
							EG2	Adaptive goal	Yes	Yes	Y
Shapiro et al. [87][e]	EG1: 18 (13); EG2: 18 (7)	8.7 (5-13)	Males and females	Non-CT	Hip-worn	8	EG1	Static goal	Yes	Yes	Y
							EG2	Static goal	Yes	Yes	N
Shimon & Petlichkoff [88][e]	CG: 62 (36); EG1: 72 (43); EG2: 60 (34)	13.1 (12-14)	Males and females	Cluster-RCT	Hip-worn	4	EG1	Adaptive goal	Yes	Yes	N
							EG2	No goal	Yes	No	N
Shore et al. [89][e]	EG1: 57 (46); EG2: 56 (46)	11.9 (11-12)	Males and females	Non-CT	Hip-worn	6	EG1	Static goal	Yes	Yes	Y
							EG2	No goal	Yes	No	N
Smith et al. [90][e]	CG: 180 (154); EG1: 181 (139)	12.7 (12-14)	Males	Cluster-RCT	Hip-worn	20	EG1	Adaptive goal	Yes	Yes	Y
Thompson et al. [91][e]	CG: 40 (34); EG1: 40 (36); EG2: 40 (31); EG3: 40 (37)	15.2 (14-17)	Males and females	RCT	Hip-worn	12	EG1	No goal	No	Yes	N
							EG2	No goal	No	Yes	Y
							EG3	No goal	No	Yes	Y
Wang [92][e]	EG1: 34 (24); EG2: 32 (22)	13.5 (13-14)	Females	Non-CT	Hip-worn	6	EG1	Adaptive goal	Yes	Yes	N
							EG2	No goal	Yes	No	N
Zizzi et al. [93][e]	EG1: 84 (68); EG2: 81 (60)	16 (14-18)	Males and females	Non-CT	Hip-worn	3	EG1	No goal	Yes	Yes	N
							EG2	No goal	Yes	No	N

Abbreviations: CG = Control group; EG = Experimental group; CT = Controlled trial; RCT = Randomized controlled trial; P/
vigorous physical activity; SB = Sedentary behavior; ACL = Research accelerometer.

[a] Number of participants are reported as "Group: Initial number (Final number)"; [b] Age is reported as "Mean (Minimum-M
number of weeks; [d] Wrist-worn and hip-worn wearable correspond to the same consumer-wearable activity trackers used
correspond to research-grade activity trackers; [e] Studies included in the meta-analysis (quantitative synthesis).

Regarding the design of the studies, five were true-experimental trials (i.e., a randomized controlled trial design), 19 were quasi-experimental trials (i.e., a cluster-randomized controlled design or non-randomized controlled design), and 21 were pre-experimental trials (i.e., non-controlled design). Regarding the measurement instrument, 27 used a hip-worn consumer-wearable activity tracker, five a wrist-worn consumer-wearable activity tracker, and 11 a research-grade activity tracker, while two studies reported both wrist-worn consumer-wearable activity tracker and research-grade activity tracker measures.

The consumer-wearable activity tracker-based programs included had a mean length of 11.78 ± 13.17 weeks, but varying considerably from three [56] to 84 weeks [57]. Specifically, 44.44% of the included studies last more than eight weeks. Most of them used a waist-worn consumer-wearable activity tracker as a motivational tool in the program (n = 35 waist-worn; and n = 10 wrist-worn). Regarding the intervention characteristics, from the 64 experimental groups formed (along with the 45 studies), they included: A goal-setting strategy (the 20.31% and the 43.75% of experimental groups used a static goal and an adaptive goal, respectively); participants' logbooks (the 56.25% of experimental groups); educational counseling sessions (the 62.50% of experimental groups); reminders to persuade participants to move (the 28.13% of experimental groups); motivational strategies (the 42.19% of experimental groups); or exercise routines (the 17.19% of experimental groups).

3.3. Risk of bias and certainty of the evidence

Supplementary File 5 shows the risk of bias assessment for each included unit of analysis. For total daily steps, 60.38% (n = 32) of units of analysis were assessed as overall "High risk", 37.74% (n = 20) as "Some concerns", and only 1.88% (n = 1) as overall "Low risk". Regarding MVPA, 80.00% (n = 16) was assessed as overall "High risk", 10.00% (n = 2) as "Some concerns", and 10.00% (n = 2) as overall "Low risk". Concerning total PA, 75.00% (n = 6) was assessed as overall "High risk" and 25.00% (n = 2) as "Some concerns". Finally, for SB, 62.50% (n = 5) was assessed as overall "High risk", 12.50% (n = 1) as "Some concerns" and 25.00% (n = 2) as overall "Low risk".

Particularly, the risk of bias arising from the randomization process was the most problematic domain, with 21.35% of units of analysis being classified as "High risk" for not being a random design and the 44.94% classified as "Some concerns" due to being a non-controlled design or presenting imbalances regarding sample size or PA status between control and experimental groups. Moreover, the bias in measurement of the outcome domain was also problematic with half of the unit of analysis classified as "High risk" (the 13.48%) or "Some concerns" (the 37.08%) due to inappropriate method of measuring (e.g., studies using consumer-wearables as measurement instrument instead of research-grade).

Moreover, based on the GRADE assessment (Supplementary File 6), the effect size of daily total steps was classified as "Low" certainty of evidence due to both the risk of bias and inconsistency domains presenting serious limitations (i.e., most studies were classified as "High risk" of bias, and a substantial level of heterogeneity was found, respectively). Regarding MVPA, total PA and SB, the effect size was classified as "Moderate" certainty of evidence due to serious limitations in the risk of bias domain (most studies were classified as "High risk" of bias).

3.4. Effects sizes

The results of the effects sizes showed that consumer-wearable activity tracker-based programs had a statistically significant moderate positive effect on daily total steps (∂ = 0.612, 95% CI = 0.477-0.746, $p < 0.001$; Figure 2), small positive effect on daily levels of MVPA (∂ = 0.220, 95% CI = 0.134-0.307, $p < 0.001$; Figure 3), and trivial positive effect on daily levels of total PA (∂ = 0.151, 95% CI = 0.038-0.264, $p = 0.009$; Figure 4). Moreover, the results of effect sizes carried out with Mean Difference showed significant positive effects on daily total steps (D = 1,692.792, 95% CI = 1,322.522-2,063.061, $p < 0.001$) and MVPA (D = 5.583, 95% CI = 2.527-8.640, $p < 0.001$). However, the programs had a statistically significant trivial negative effect on daily SB (∂ = 0.172, 95% CI = 0.039-0.305, $p = 0.011$; Figure 5).

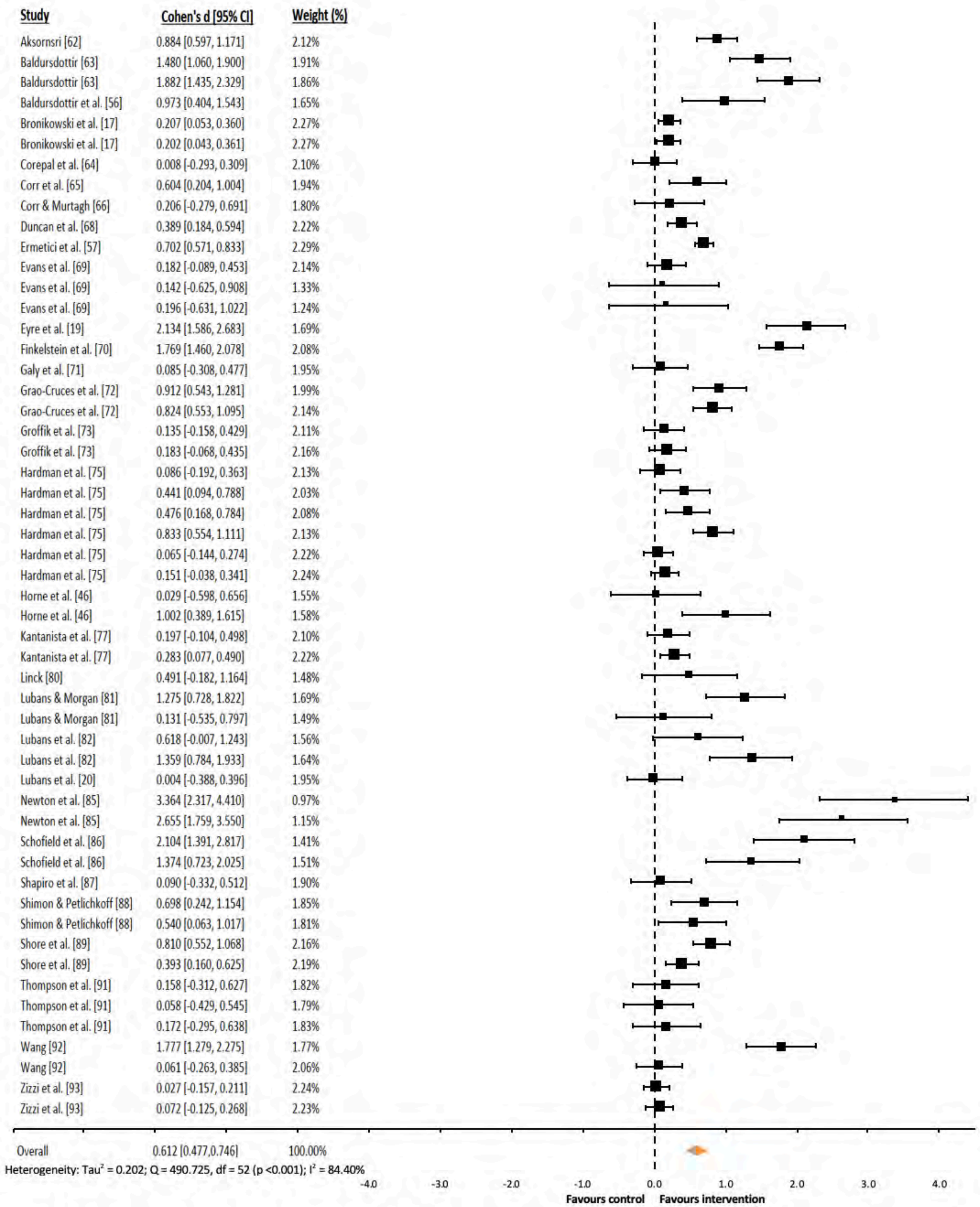


Figure 2. Forest plot of standardized mean differences (∂) comparing consumer-wearable activity tracker-based programs effects on daily total steps.

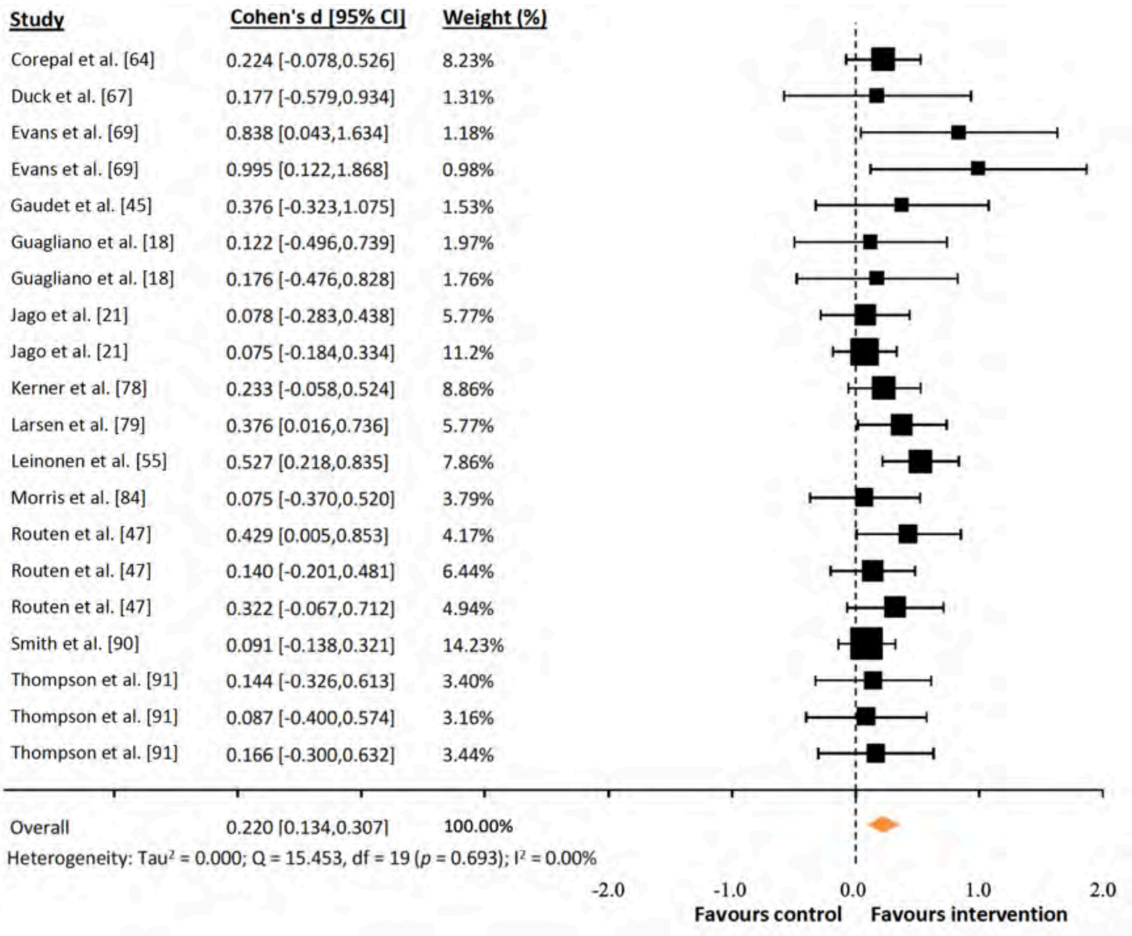


Figure 3. Forest plot of standardized mean differences (*∂*) comparing consumer-wearable activity tracker-based programs effects on daily moderate-to-vigorous physical activity.

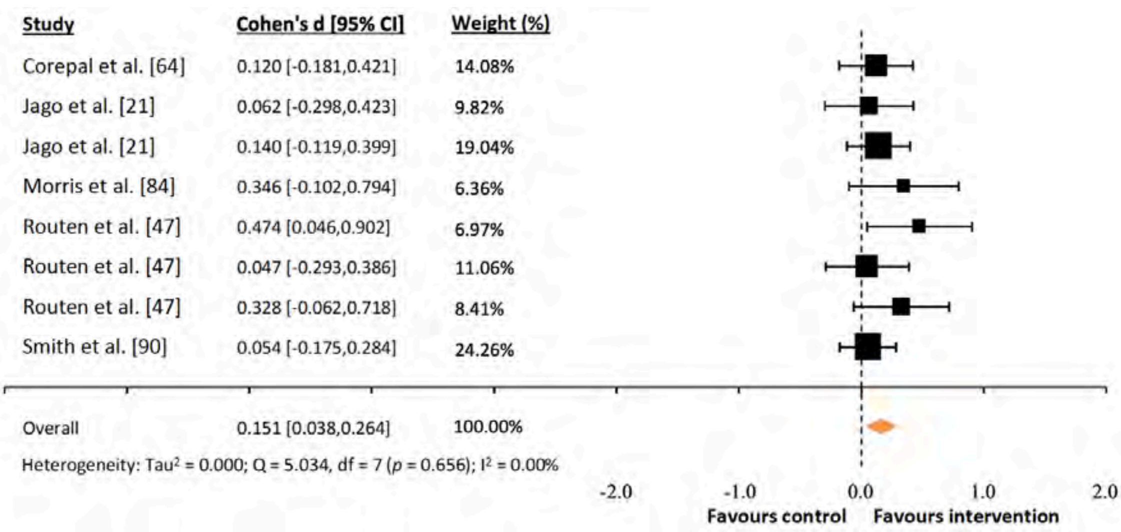


Figure 4. Forest plot of standardized mean differences (*∂*) comparing consumer-wearable activity tracker-based programs effects on daily total physical activity.

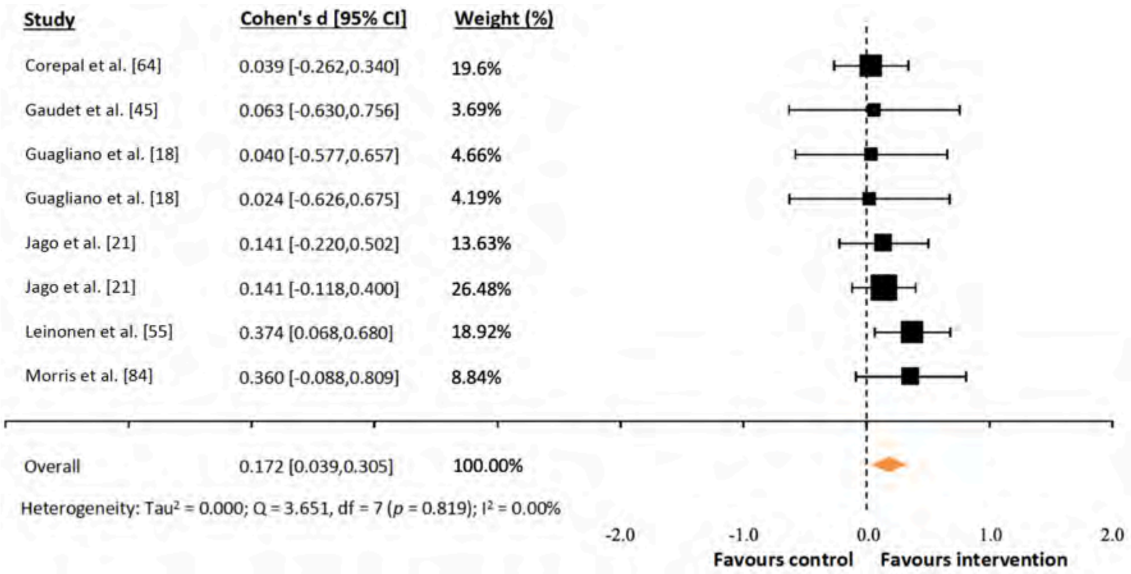


Figure 5. Forest plot of standardized mean differences (∂) comparing consumer-wearable activity tracker-based programs effects on daily sedentary behavior.

Furthermore, regarding the heterogeneity of the results, for daily MVPA (only for ∂ analysis), total PA and SB heterogeneity was not found (I^2 = 0.00%). However, high heterogeneity was found for daily total steps (I^2 = 89.40% and 93.36% for ∂ and D analyses, respectively), and moderate heterogeneity for daily MVPA (I^2 = 54.23%, only for D analysis). Therefore, together with the fact that the number of studies with daily total PA and SB was low (k = 8), follow-up moderator analyses were conducted only for daily total steps and MVPA.

3.5. Publication bias

3.5.1. Avoiding duplicated information

Although three publications met the selection criteria [58–60], they were not included in the systematic review and meta-analysis because the same data had been reported in other publications.

3.5.2. Identifying publication bias

The visual assessment of funnel plots suggested that there was publication bias for daily total steps and MVPA (Figure 6). In this sense, the results of the Egger's test were statistically significant for daily total steps (p = 0.001) and MVPA (p = 0.045).

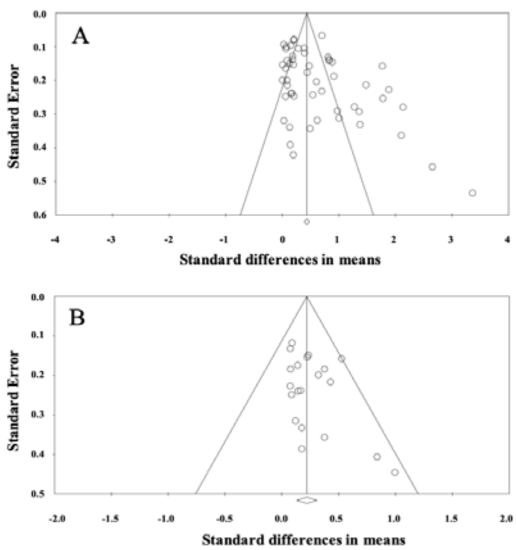


Figure 6. Funnel plots of standard error by standard differences in means (∂) comparing consumer-wearable activity tracker-based programs effects on (a) daily total steps, and (b) moderate-to-vigorous physical activity.

3.5.3. Assessing the impact of publication bias

The results of Orwin's fail-safe *N* analyses showed that the number of missing studies needed to bring the mean ∂ values under a trivial value was unlikely, especially if the percentage of unlocated/located studies was considered, being 177 (333.96%) for daily total steps and 25 (125.00%) for daily MVPA. Furthermore, the results of Duval and Tweedie's Trim and Fill method did not trim any study for daily steps and only two studies for MVPA in the standardized mean difference ∂ analyses. Regarding Mean Difference *D* analyses, no study for daily steps and only one study for daily MVPA were trimmed. Moreover, the adjusted value for MVPA was similar to the observed values (∂ = 0.220 vs. 0.213; *D* = 5.583 vs. 4.824). Finally, regarding the cumulative meta-analysis plots sorted by larger study (Supplementary File 7), after some fluctuations in the first studies which may be due to chance [61], a fairly constant estimate of the effect over sample size was observed for total steps and MVPA. Although a large effect in the first primary study was found, the summary value was decreased after the 14^{th}-to-15^{th} study and the 2^{nd}-to-3^{rd} study for daily total steps and MVPA, respectively. Furthermore, neither a transient loss of formal significance nor a complete reversal of the initial association was found. Finally, it is worth mentioning that the addition of new primary studies did not materially change the estimates, so the final effect size values for these intervention programs seem to be quite robust.

3.6. Moderator analyses

3.6.1. Moderator analyses for daily total steps

Supplementary File 8 shows the results of the within-study moderator analyses for the effect of the consumer-wearable activity tracker-based programs on the daily total steps among school-aged children. The initial values of school-aged children's accomplishment with PA recommendations was the only individuals' characteristic that influenced on the effect of the programs, being more effective in those physically inactive than in those physically active (d = 1.206 vs. 0.107; p < 0.001). Moreover, heterogeneity analyses showed that the effect of the program separately for accomplishment with PA recommendations was homogeneous (I^2 = 0.00%).

Furthermore, Supplementary File 9 shows the results of the between-study moderator analyses. Regarding the influence of the individuals' characteristics, subgroup analysis showed that the program had statistically significantly more effect in females than in males (d = 0.636 vs. 0.266; p = 0.044) and in physically inactive (i.e., not meeting the daily PA recommendations) than in those that were physically active (d = 0.795 vs. 0.397; p = 0.003) for increasing daily total steps. As regards the influence of the intervention characteristics, programs with some kind of counseling were more effective than those without counseling (d = 0.711 vs. 0.407; p = 0.003), likewise, programs that included goal-setting were more effective than those that did not (d = 0.770 vs. 0.243; p < 0.001) for increasing school-aged children's daily total steps. However, for the rest of the between-study subgroups comparisons no statistically significant differences were found (p > 0.05). According to heterogeneity analyses, the effect of these programs separately for gender, PA status, goal-setting and counseling was still moderate-to-high heterogeneous (I^2 = 66.08-92.54%).

Moreover, meta-regression analysis showed that the effect of the intervention program was statistically significantly associated with PA status and number of strategies, being the effect higher for daily total steps in less active school-aged children (Q = 10.83; p = 0.001; R^2 = 0.07; I^2 = 88.58%), and when a greater number of strategies were included in programs (Q = 5.19; p = 0.023; R^2 = 0.02; I^2 = 89.02%) (Figure 7). However, no statistically significant associations were found by intervention duration (Q = 0.38; p = 0.536; R^2 = 0.00; I^2 = 89.04%) nor age (Q = 2.17; p = 0.141; R^2 = 0.01; I^2 = 89.17%) and school-aged children's daily total steps.

Furthermore, meta-regression analysis of a more complex model including both significant explanatory variables together (i.e., PA status and number of strategies included), showed a statistically significant association with intervention effect on daily total steps, explaining a higher percentage of the variance than both explanatory variables alone (Q = 16.55, p < 0.001; R^2 = 0.09; I^2 = 89.40%).

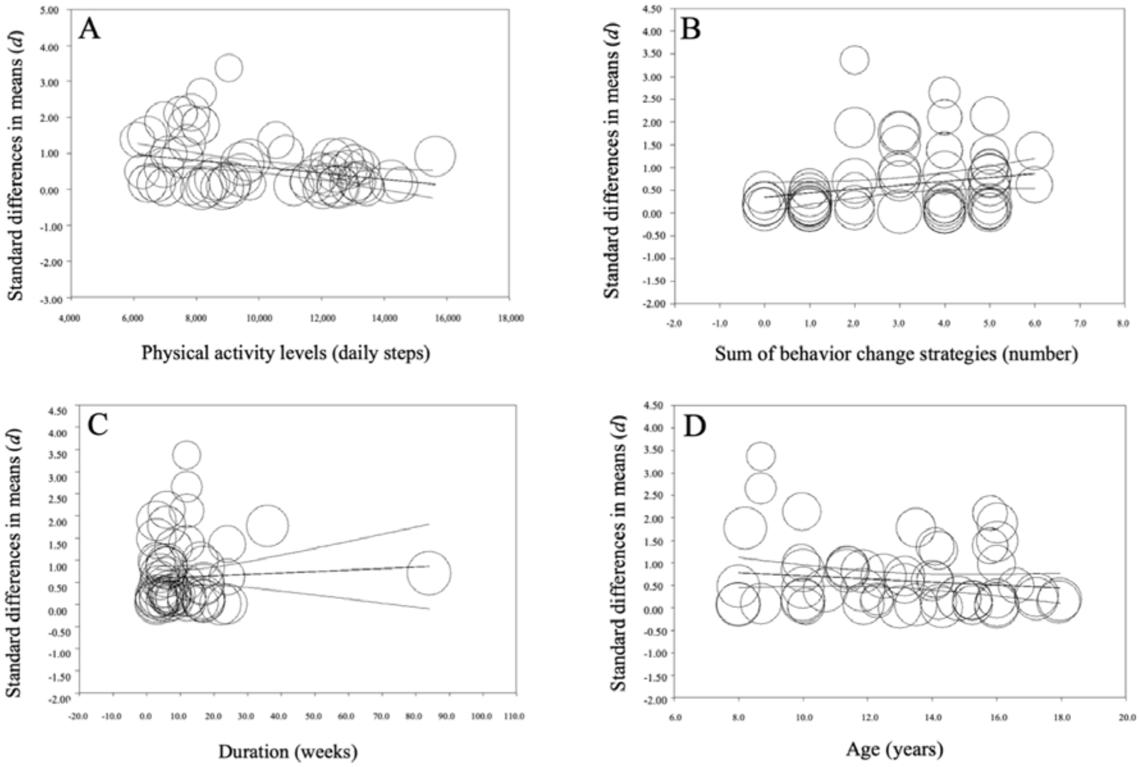


Figure 7. Meta-regression scatterplots of standard differences in means (*d*) comparing consumer-wearable activity tracker-based programs effects on daily total steps regarding (a) school-aged children's physical activity levels, (b) strategies included in the program, (c) duration, and (d) school-aged children's age.

3.6.2. Moderator analyses for daily moderate-to-vigorous physical activity

Supplementary File 10 shows the results of the within-study moderator analyses for the effect of the consumer-wearable activity tracker-based programs on the daily MVPA levels among school-aged children. Results showed no statistically significant differences for any subgroups comparisons (i.e., gender, goal-setting, diary, and reminders).

Furthermore, Supplementary File 11 shows the results of the between-study moderator analyses. Regarding the influence of the school-aged children's characteristics, subgroup analyses showed that the program had significantly more effect in physically inactive participants than in those physically active (∂ = 0.404 vs. 0.170; p = 0.046) for increasing daily MVPA levels. As regards the influence of the intervention programs' characteristics, programs carried out with a wrist-worn activity tracker were more effective than those with a waist-worn activity tracker (∂ = 0.413 vs. 0.167; p = 0.021), moreover, programs that did not include any exercise routine were more effective than those that did (∂ = 0.283 vs. 0.083; p = 0.036). However, for the rest of the between-subgroups comparisons no statistically significant differences were found (p > 0.05). According to heterogeneity analyses, results by subgroups showed homogeneous results (I^2 = 0.00-29.06%).

Moreover, meta-regression analyses did not show statistically significant associations between PA status (Q = 0.65; p = 0.420; R^2 = 0.00; I^2 = 0.00%), number of strategies (Q = 0.06; p = 0.805; R^2 = 0.00; I^2 = 0.00%), intervention duration (Q = 1.55; p = 0.213; R^2 = 0.00; I^2 = 0.00%), nor participants' age (Q = 1.84; p = 0.175; R^2 = 0.00; I^2 = 0.00%) and school-aged children's daily MVPA levels (Figure 8).

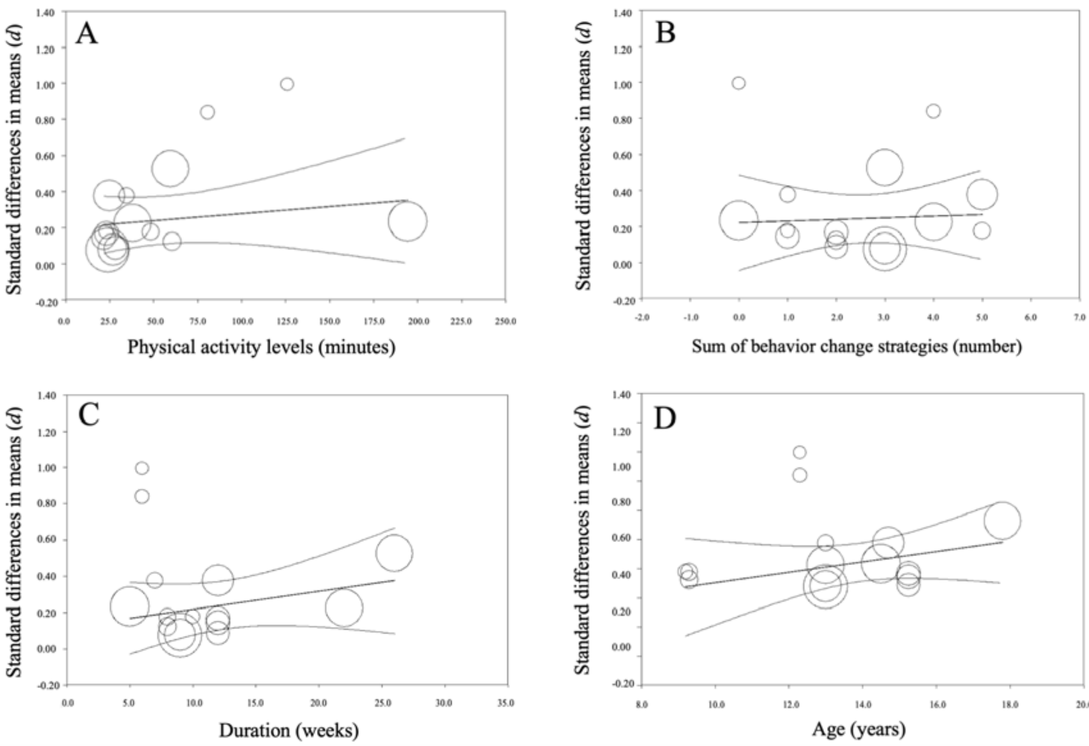


Figure 8. Meta-regression scatterplots of standard differences in means (∂) comparing consumer-wearable activity tracker-based programs effects on daily moderate-to-vigorous physical activity regarding (a) school-aged children's physical activity levels, (b) strategies included in the program, (c) duration, and (d) school-aged children's age.

3.7. Sensitivity analyses

The results of the sensitivity analyses for the overall effects sizes carried out with Hedges' g random-effects model for daily total steps (g = 0.600, 95% CI = 0.467-0.732, p < 0.001), MVPA (g = 0.218, 95% CI = 0.133-0.302, p < 0.001), total PA (g = 0.150, 95% CI = 0.039-0.261, p = 0.008) and SB (g = 0.171, 95% CI = 0.038-0.303, p = 0.011) showed the same results as the main analysis carried out with Cohen's d with a random-effects model. Similarly, overall effects sizes carried out with Cohen's d with a fixed-effect model for daily total steps (d = 0.434, 95% CI = 0.392-0.475, p < 0.001), MVPA (d = 0.220, 95% CI = 0.134-0.307, p < 0.001), total PA (d = 0.151, 95% CI = 0.038 -0.264, p = 0.009) and SB (d = 0.172, 95% CI = 0.039-0.305, p = 0.011) also showed the same results as the main analysis with the random-effects model.

Finally, results of the Cohen's d with a random-effects model separately for randomized controlled trial design compared with non-randomized controlled trial design showed greater effects in randomized controlled trials for daily total steps (d = 0.859 vs. 0.465; p = 0.004). However, no statistically significant differences were found for MVPA (p = 0.104) nor total PA (p = 0.328). Regarding SB, all the included studies were randomized controlled trials and the comparison could not be performed.

4. Discussion

4.1. Overall effects of consumer-wearable activity tracker-based programs

The present systematic review and meta-analysis synthesizes the evidence to date about the effectiveness of consumer-wearable activity tracker-based programs on daily objectively measured PA and SB among apparently healthy school-aged children [17–21,45–47,54–57,62–93]. The overall results showed that the consumer-wearable activity tracker-based programs had significant moderate improvements in school-aged children's daily total steps, and small but significant improvements in daily MVPA levels after the intervention program. However, regarding daily levels of total PA the effect was trivial. Therefore, the use of a consumer-wearable activity tracker as a motivation tool for young people is kindly recommended to reduce the high levels of physical inactivity among school-aged children [8]. However, the intervention programs seem not to reduce the school-aged children's SB.

These results agree with similar previous meta-analyses carried out in adults [25–27]. Firstly, regarding daily total steps, all previous reviews found improvements, although they seem greater in the present systematic review (∂ = 0.612 vs. 0.240-0.449; D = 1,692.79 vs. 950.54). Moreover, with reference to MVPA, similar improvements in daily levels of MVPA were obtained (∂ = 0.220 vs. 0.270; D = 5.583 vs. 6.160). The greater improvement in school-aged children's daily total steps than in their MVPA levels may be due to the kind of goal established in the program. Most of the studies included in the systematic review with a goal-setting strategy set only a step-based goal (30 of 37 studies), while only five studies established both step-based and minutes of total PA-based goals [21,86] or SB-based goals [87,90]. Therefore, all the motivational strategies included, such as reminders [46,56,65], counseling sessions [69], or rewards [64,72] were carried out around this goal of increasing the number of steps. Moreover, the reason for relying mainly on the number of steps as the reference output for goal-setting may be due to steps having the advantage of being easier to understand and interpret by school-aged children compared to MVPA minutes [16]. In addition, in many studies, the consumer-wearable activity-tracker used were pedometers, which only show feedback about the number of steps. Therefore, similar to the evidence found with adults [25–27], consumer-wearable activity tracker-based programs seem to be effective for improving objectively measured daily PA, especially for daily total steps, although the effects seem to be greater for school-aged children. This greater effect may be due to the fact that school-aged children have a higher affinity to new technologies playing an important role in their daily life, therefore these kinds of technology-based interventions may be more interesting for them [94,95]. In addition, during the period of childhood and early adolescence, school-aged children are still forming their daily habits and they are more sensitive to changing their PA behavior which could explain these greater intervention effects, while in adulthood the stability of PA patterns is high and more difficult to be changed [96].

Regarding total PA, results showed a trivial positive effect (∂ = 0.151, 95% CI = 0.038-0.264). These results are curious because daily total steps is an indicator of total PA [97], and therefore, results were expected to be similar. However, it may be because very slow steps were included in daily total steps, but they do not reach the threshold to be considered light PA. Besides, regarding SB, similar to the trivial effect obtained in the present systematic review (∂ = 0.172; 95% CI = 0.039-0.305), no real differences were found in any previous review [25,27]. Firstly, this may be due to the fact that most of the programs that evaluate SB only used strategies to encourage and support PA behavior change (i.e., goals, tips and challenges, behavioral incentives or reinforcement messages based only on PA practice) and were not specifically designed to reduce SB [18,21,45,64,76,84]. Only Leinonen et al. [55] specifically included some SB-based strategies, such as feedback showing a thumb either up or down if the day included over two hours of sedentary (sitting) periods or not, and rewards regarding decrement in weekly school-aged children's sedentary time. The lack of specific focus on reducing SB may have contributed to this trivial effect. Moreover, these results may also present significant measurement bias due to all SB-studies (except Morris et al. [84]) analyzed raw time involved in SB per day instead of valid wear time-based standardized scores (e.g., percentage of time of each day engaged in SB of the total valid wear time; standardized mean SB in minutes) [98,99]. In addition, it must be considered that for accelerometer-based measures only a valid minimum time per day is established (normally 600 min), but that standardized values are not taken into account [100]. This is even more accentuated for consumer-based wearables whose valid wear time cannot be known and, therefore, it cannot be controlled if school-aged children wore the wearable for a long enough time, or if they were more motivated to wear it more time in the baseline or post-intervention measure. In this sense, it is easier for school-aged children with more valid time to have higher registered time involved in SB, so SB time outcomes could be especially affected by potential systematic valid wear-time variation between measurement moments (i.e., pre-post-intervention measures) or groups [98]. This could also affect daily total steps and daily MVPA levels although to a much lesser extent than for SB [98]. For instance, Gaudet et al. [45] showed that school-aged children in the experimental group had approximately 48 minutes more wear time per day in comparison with the control group, which may directly affect their differences regarding time involved in SB. However, most studies [18,21,64,76] only reported the minimum time per day needed to be included in the study but did not report the actual mean valid wear time per day, and it is important that it be reported in order to compare compliance between groups. Finally, it should be noted that there are very few studies measuring school-aged children's SB in comparison with studies measuring steps and MVPA, which implies a wider confidence interval and consequently greater uncertainty about the real value.

4.2. Influence of participants' characteristics

According to the results of the present meta-analysis, a significant relationship has been observed for PA status from the within-study moderator analysis, the intervention being much more effective for improving daily total steps in school-aged children who were physically inactive in the baseline measure than in those who were physically active. Furthermore, this positive influence was also found for improving objectively measured daily MVPA levels, although only from the between-study analysis (i.e., an observational relationship). A potential reason for this influence could be that school-aged children who already are physically active before the intervention are motivated enough for PA practice without the need for extra motivation with the proposed intervention [101]. Furthermore, it is difficult to further increase school-aged children's PA levels when the baseline levels are high, just as has occurred with the improvement of physical fitness levels [102,103]. Therefore, consumer-wearable activity tracker-based programs may be an especially appropriate strategy for less active school-aged children [81,84]. In line with the recommendation by Love's et al. [104] systematic review, future studies should analyze their results distinguishing by school-aged children's baseline PA profiles to correctly identify the intervention impact because all participants do not react in the same way to the intervention. Regarding school-aged children who already are physically active before the intervention, it may be necessary to study which specific strategies should be implemented in these interventions to help them maintain their PA levels, or even to continue increasing them, therefore obtaining greater health-related benefits [4]. Finally, regarding school-aged children's age, it does not seem to affect the intervention program effect.

Furthermore, programs seem to be more effective in females than in males for improving daily total steps, which could be related to PA status since females tend to be more physically inactive than males throughout childhood and adolescence [105,106]. Therefore, gender-specific interventions could be considered in future research like Böhm et al. [31] suggested, although the conclusions in relation to the effectiveness of gender-specific interventions should be taken with caution because the success of those programs was not obtained by the within-study moderator analysis (i.e., cause-effect relationships), but by the between-study analysis, which only establish observational relationships [17,46,68,72,73,82]. Furthermore, the between-study moderator analysis results still showed moderate-to-high heterogeneity in each subgroup (i.e., males and females), so it is likely that there were differences in other strategies of the intervention or participants' characteristics that could influence its effect, even more than gender itself. Additionally, these differences are even more accentuated because the analysis includes twice as many publications carried out with females than with males. However, the number of studies including this within-study gender comparison and which allows establishing cause-effect relationships is very limited (k = 7), and for this reason, the observational relationships of the between-study comparisons have been considered in the present meta-analysis.

4.3. Influence of intervention programs' characteristics

Firstly, overall meta-regression analysis results showed a higher effect for daily total steps when a greater number of strategies were included in the programs. Therefore, multi-dimensional interventions that include most of these strategies seem to be preferable for mediating PA behavior. That agrees with some psychological theories such as the Social Cognitive Theory [107], the Self-Determination Theory [108] or the Trans-Contextual Model [109] which includes most of those strategies as positive mediators influencing PA behavior. Furthermore, recent systematic reviews highlight that using multi-strategy approaches as behavior change techniques show better PA outcomes than singular change approaches [110,111]. For instance, the PA-related knowledge provided in the counseling sessions, the autonomy support environment by the consumer-wearable activity tracker feedback, the inclusion of additional motivational strategies like social networks, setting goals of moderate difficulty, or sending reminders with encouraging messages about PA practice have been shown to positively change school-aged children's PA behavior [112,113]. However, it should be noted that the explained variance was low (R^2 = 0.02), as well as the results being highly heterogeneous (I^2 = 89.02).

Furthermore, there was considerable heterogeneity in the strategies included in the reviewed studies. Most of the programs used a goal-setting strategy, participants' logbooks, educational counseling sessions, and/or some kind of motivational strategy, being the inclusion of reminders to persuade participants to move or exercise less frequent. The influence of these intervention program characteristics in school-aged children's daily total steps and MVPA has been analyzed in the present meta-analysis. However, it must be considered that only observational relationships have been obtained from the between-study moderator analyses due to the low number of studies compared in the within-study moderator analyses. Furthermore, the results regarding school-aged children's daily total steps separated by each subgroup, still showed a high level of heterogeneity which implies differences in other intervention characteristics. Therefore, future studies should include different intervention groups that compared some intervention characteristics (e.g., one experimental group including counseling and another without counseling) to establish causal-effect relationships between intervention characteristics and the effect of the programs.

Firstly, interventions including some kind of counseling [19,56,70] and/or goal-setting techniques [72,85,89] were highlighted as more effective than those without them for improving school-aged children's daily total steps. These results are in accordance with previous studies' recommendations about the inclusion of these explicit strategies (e.g., advice about PA benefits, strategies to reduce SB and increase PA, resolution of barriers to PA practice, or goal-setting strategies based on the international guidelines) which make students feel that they are making an informed decision about their health in any kind of program for PA promotion [11,13] and specifically in wearable-based programs [29,112].

Nevertheless, apparently contradictory results showed that programs which did not include any exercise routine seem to be more effective than those that included it for improving school-aged children's daily MVPA levels. However, analyzing the kind of exercise routine included, had some limitations. Firstly, Jago et al. [21] and Smith et al. [90] included a low frequency of supervised PA sessions (one 20-min PA session per week, and only six 20-min lessons in 20 weeks, respectively) with which it is very difficult to positively affect the school-aged children's daily PA levels. Secondly, most of the activities included by Jago et al. [21] did not have a direct relationship with increasing the school-aged children's number of steps or minutes involved in MVPA (i.e., stretching, technical drills, or strengthening tasks), and also it should be noted that despite including this exercise routine, they did not include other strategies that may be even more important than this one (e.g., reminders, diary, or motivational strategies). Finally, it is also important to denote that the analysis included 17 units of analysis without an exercise routine vs. three with an exercise routine [21,90], so given this marked difference in the sample, results should be interpreted with caution.

Regarding the kind of consumer-wearable activity tracker, waist-worn trackers such as pedometers are more frequent than wrist-worn trackers. This could mainly be due to the fact that waist-worn trackers such as traditional step counters with digital displays, have been around in scientific research since approximately 1996, being an accepted method for assessing PA and a tool for walking interventions [114]. On the contrary, wrist-worn trackers have burst onto the market in the last decade and, therefore, their scientific evidence is still scarce [14,15]. Moreover, the meta-analysis results showed that programs carried out with a wrist-worn activity tracker seem to be more effective than those carried out with waist-worn trackers for improving school-aged children's daily MVPA levels. This may be due to wrist-worn trackers having several advantages compared to waist-worn trackers, such as reporting real-time feedback that can be easily checked on their wrist or touch screens [14]. Moreover, unlike more traditional waist-based trackers, which only monitor and display simple feedback about PA levels, wrist-based trackers are much more interactive since the user is able to set reminders, notifications, or congratulatory messages upon reaching the proposed goal [115]. Finally, wrist devices have shown greater wear time compliance which could mean that if they wear it for a longer time school-aged children could interact more with its features [116].

4.4. Risk of bias and certainty of the evidence

Firstly, based on the methodological risk of bias assessment, most studies were classified as "High risk" or "Some concerns", leaving only two studies classified as overall "Low risk" [18,80]. Therefore, it may lead to a biased assessment of the intervention effect, underestimating or overestimating the true intervention effect, which supposed downgrading the GRADE certainty rating by one level for all outcomes (i.e., daily total steps, MVPA, total PA, and SB) regarding the risk of bias domain [35,48]. In reference to the study designs, it is interesting to highlight that only 51.11% of the included studies are true or cluster-randomized controlled trials, which are markedly far stronger interventions to demonstrate effect significance [24].

However, sensitivity analysis showed no differences for school-aged children's daily MVPA, total PA, and SB levels between study designs, but much greater effects were found in randomized controlled trials for school-aged children's daily total steps than non-randomized trials.

Moreover, regarding daily total steps outcomes, a substantial level of heterogeneity was found, even in the follow-up moderator analyses (except when separating by accomplishment with PA recommendations) and it supposed downgrading the GRADE certainty rating by another level regarding inconsistency domain for daily total steps. This is most likely because it is the PA outcome that includes the largest number of studies and, consequently, the greatest variety in the types of intervention applied when compared with other outcome measures. Finally, regarding publication bias, although the funnel plots and Egger's test suggested publication bias for daily total steps and daily MVPA levels, its impact seems to be very low given the unlikely number of "lost" studies suggested by the fail-safe *N* analyses. Furthermore, the Trim and Fill method did not trim any study for daily steps and only two studies for MVPA, having an adjusted value similar to the observed values (∂ = 0.220 vs. 0.213).

For all the above-mentioned reasons, it is important to highlight the "Low" certainty of evidence found for daily total steps, which supposes that the confidence in the effect estimate is limited and the true effect may be different from the estimated effect [117]. Regarding MVPA, total PA and SB outcomes, "Moderate" certainty of evidence was found, so the true effect is likely close to the estimated effect, but there is a possibility that it is substantially different [117]. Therefore, the findings of the present meta-analysis should be considered with caution and firmer conclusions should await the accumulation of a larger high-quality number of primary studies.

4.5. Strengths and limitations

Regarding the strengths of the present systematic review, numerous measures to avoid, or at least to reduce, publication bias were followed (e.g., the inclusion of a great range of bibliographic databases from different disciplines and complementary search strategies, or not restricting the search by the language, type or date of publication). Then, several exploratory analyses were conducted to identify and assess the impact of any potential publication bias (e.g., funnel plots, or Orwin's fail-safe N analyses), as well as sensitivity analyses (e.g., Hedges' *g* with a random-effects model or Cohen's ∂ with a random-effects model separately for randomized controlled trial design or not) to verify the robustness of the results. Furthermore, the present review was focused only on objective measurements, which have shown high validity to measure PA and SB levels in comparisons with self-reported measures [36,37]. Lastly, to our knowledge, to date this is the first systematic review and meta-analysis about the effects of consumer-based activity tracker-based programs on objectively measured PA and SB levels within apparently healthy school-aged children, including analyzing the influence of the intervention programs' characteristics and school-aged children's characteristics on the effects. This meta-analysis summarizes the effectiveness of those interventions in an overall statistical synthesis, improving the precision of the results by the estimation of the effect size and direction, and knowing whether or not the effect size is consistent across studies.

However, the present systematic review and meta-analysis is not without limitations. First, although randomized controlled trials have higher methodological quality, the present systematic review includes several study designs. As expected, there were not too many consumer-based activity tracker-based studies with a high level of quality design for improving the different PA-related behaviors. Therefore, a reason for including several designs is to provide evidence of the effects of interventions for which only a small number of randomized controlled trials are available, drawing on the "best available evidence" rather than the "highest tier" of evidence [35]. Nevertheless, sensitivity analyses were also performed comparing randomized controlled trials and non-randomized trials, showing no differences for school-aged children's daily MVPA, total PA, and SB levels. Furthermore, even greater effects in school-aged children's daily total steps were found in randomized controlled trials. Second, although the inclusion of a wide range of intervention types, populations, sample size, and study designs had some advantages regarding the generalizability of conclusions, it supposes a high level of heterogeneity. Therefore, it makes the independent contribution of any intervention features and, therefore, making strong conclusions from the intervention difficult to establish. However, in addition to the overall effect size, subgroup analyses and meta-regression of the *a priori* hypothesized moderators were also performed. Therefore, not only general effect results are provided, but also for each specific group based on the characteristics of the programs and school-aged children. Third, the present systematic review investigated effectiveness right at the end of the program (i.e., short-term), but future studies should investigate long-term effectiveness to assess actual behavioral changes some months after the program. However, due to the very limited evidence, it was not performed in the present systematic review. Finally, coding some study outcomes was problematic due to authors not reporting them. Although authors were contacted, many of them did not reply and the particular study outcome had to be omitted. However, this is a common problem in most systematic reviews [35], and a great effort was made in contacting authors, recalculating data, or estimating values from figures. Finally, in some cases, consumer-based activity trackers were both used as a motivational instrument during the intervention and to objectively measure PA, which could affect results by increasing their actual PA levels in the control group or during baseline assessments.

5. Conclusions

The present findings suggest that consumer-wearable activity tracker-based programs within school-aged children have a statistically significant moderate positive effect on daily total steps and a small, but positive effect on objectively measured daily levels of MVPA. However, the effect of the programs on school-aged children's objectively measured total PA levels and SB was trivial. The findings of this systematic review suggest that programs are more effective in females for increasing daily total steps, and in physically inactive subjects for increasing both daily total steps and daily MVPA levels. Moreover, the inclusion of a greater number of strategies in the programs had a higher effect on school-aged children's daily total steps. It should be highlighted that prompting specific goal-setting and the inclusion of educational counseling sessions are particularly useful strategies for being included in consumer-wearable activity tracker-based programs designed to promote school-aged children's daily total steps. Furthermore, regarding the kind of consumer-wearable activity tracker, programs were more effective to increase school-aged children's daily MVPA levels if a wrist-worn activity tracker was used. However, since only observational relationships have been obtained from the between-study moderator analyses due to the low number of studies, all the above-mentioned recommendations regarding intervention strategies should be taken with caution. However, due to the certainty of evidence being from "Low" to "Moderate", further primary research is needed to determine the effectiveness of these programs using robust designs with low risk of bias, and which compare the effect of different intervention characteristics in the same study. Consumer-wearable activity tracker-based programs (particularly those including goal-setting, educational counseling, and wrist-worn trackers) seem to be effective to promote school-aged children's daily total steps and MVPA, especially for females and those who are physically inactive.

Funding

Carolina Casado-Robles is supported by a research grant from the Spanish Ministry of Universities [grant number: FPU16/03314]. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements

Authors would like to thank all of the intervention authors for their willingness, time and efforts to contribute to this analysis through the provision of data. We also thank Aliisa Hatten for the English revision of the manuscript.

Supplementary Files

Scan or click here to see Supplementary Files [Escanea el código o pincha aquí para ver los Archivos Suplementarios]

References

1. Al-Hazzaa HM, Abahussain NA, Al-Sobayel HI, Qahwaji DM, Musaiger AO. Physical activity, sedentary behaviors and dietary habits among Saudi adolescents relative to age, gender and region. Int J Behav Nutr Phys Act. 2011;8:140. https://doi.org/10.1186/1479-5868-8-140

2. Longmuir PE, Colley RC, Wherley VA, Tremblay MS. Canadian Society for Exercise Physiology position stand: Benefit and risk for promoting childhood physical activity. Appl Physiol Nutr Metab. 2014;39:1271. https://doi.org/10.1139/apnm-2014-0074

3. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public Health Rep. 1985;100:126–31.

4. World Health Organization. WHO guidelines on physical activity and sedentary behaviour. Geneva: World Health Organization; 2020.

5. Tremblay MS, Aubert S, Barnes JD, Saunders TJ, Carson V, Latimer-Cheung AE, et al. Sedentary Behavior Research Network (SBRN) – Terminology Consensus Project process and outcome. Int J Behav Nutr Phys Act. 2017;14:75. https://doi.org/10.1186/s12966-017-0525-8

6. Poitras VJ, Gray CE, Borghese MM, Carson V, Chaput J-P, Janssen I, et al. Systematic review of the relationships between objectively measured physical activity and health indicators in school-aged children and youth. Appl Physiol Nutr Metab Physiol Appl Nutr Metab. 2016;41:S197-239. https://doi.org/10.1139/apnm-2015-0663

7. Carson V, Hunter S, Kuzik N, Gray CE, Poitras VJ, Chaput JP, et al. Systematic review of sedentary behaviour and health indicators in school-aged children and youth: an update. Appl Physiol Nutr Metab. 2016;41:S240–65. https://doi.org/10.1139/apnm-2015-0630

8. Guthold R, Stevens GA, Riley LM, Bull FC. Global trends in insufficient physical activity among adolescents: a pooled analysis of 298 population-based surveys with 1·6 million participants. Lancet Child Adolesc Health. 2020;4:23–35. https://doi.org/10.1016/S2352-4642(19)30323-2

9. Arundell L, Fletcher E, Salmon J, Veitch J, Hinkley T. A systematic review of the prevalence of sedentary behavior during the after-school period among children aged 5-18 years. Int J Behav Nutr Phys Act. 2016;13:93. https://doi.org/10.1186/s12966-016-0419-1

10. World Health Organization. Global action plan on physical activity 2018–2030: more active people for a healthier world. Geneva: World Health Organization; 2018.

11. Dobbins M, Husson H, DeCorby K, LaRocca RL. School-based physical activity programs for promoting physical activity and fitness in children and adolescents aged 6 to 18. Cochrane Database Syst Rev. 2013;2:CD007651. https://doi.org/10.1002/14651858.CD007651.pub2

12. Lonsdale C, Rosenkranz RR, Peralta LR, Bennie A, Fahey P, Lubans DR. A systematic review and meta-analysis of interventions designed to increase moderate-to-vigorous physical activity in school physical education lessons. Prev Med. 2013;56:152–61. https://doi.org/10.1016/j.ypmed.2012.12.004

13. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. Health Psychol. 2009;28:690–701. https://doi.org/10.1037/a0016136

14. Strath SJ, Rowley TW. Wearables for promoting physical activity. Clin Chem. 2018;64:53–63. https://doi.org/10.1373/clinchem.2017.272369

15. International Data Corportation. Shipments of Wearable Devices Leap to 125 Million Units, Up 35.1% in the Third Quarter, According to IDC. 2020. Available at: https://www.idc.com/getdoc.jsp?containerId=prUS47067820. Accessed 15 Jul 2021.

16. Tudor-Locke C, Craig CL, Beets MW, Belton S, Cardon GM, Duncan S, et al. How many steps/day are enough? for children and adolescents. Int J Behav Nutr Phys Act. 2011;8:78. https://doi.org/10.1186/1479-5868-8-78

17. Bronikowski M, Bronikowska M, Glapa A. Do they need goals or support? A report from a goal-setting intervention using physical activity monitors in youth. Int J Environ Res Public Health. 2016;13:9. https://doi.org/10.3390/ijerph13090914

18. Guagliano JM, Armitage SM, Brown HE, Coombes E, Fusco F, Hughes C, et al. A whole family-based physical activity promotion intervention: Findings from the

families reporting every step to health (FRESH) pilot randomised controlled trial. Int J Behav Nutr Phys Act. 2020;17:120. https://doi.org/10.1186/s12966-020-01025-3

19. Eyre ELJ, Cox VM, Birch SL, Duncan MJ. An integrated curriculum approach to increasing habitual physical activity in deprived South Asian children. Eur J Sport Sci. 2016;16:381–90. https://doi.org/10.1080/17461391.2015.1062565

20. Lubans DR, Morgan PJ, Aguiar EJ, Callister R. Randomized controlled trial of the Physical Activity Leaders (PALs) program for adolescent boys from disadvantaged secondary schools. Prev Med. 2011;52:239–46. https://doi.org/10.1016/j.ypmed.2011.01.009

21. Jago R, Baranowski T, Baranowski JC, Thompson D, Cullen KW, Watson K, et al. Fit for Life Boy Scout badge: outcome evaluation of a troop and Internet intervention. Prev Med. 2006;42:181–87. https://doi.org/10.1016/j.ypmed.2005.12.010

22. Cooper H, Hedges LV, Valentine JC. The Handbook of Research Synthesis and Meta-Analysis. 3rd ed. New York: Russell Sage Foundation; 2019.

23. Schmidt FL, Hunter JE. Methods of Meta-Analysis: Correcting Error and Bias in Research Findings. 3rd Edition. London: SAGE Publications; 2021.

24. Murad MH, Asi N, Alsawas M, Alahdab F. New evidence pyramid. Evid Based Med. 2016;21:125. http://dx.doi.org/10.1136/ebmed-2016-110401

25. Brickwood K-J, Watson G, O'brien J, Williams AD. Consumer-based wearable activity trackers increase physical activity participation: Systematic review and meta-analysis. JMIR MHealth UHealth. 2019;7. https://doi.org/10.2196/11819

26. Tang MSS, Moore K, McGavigan A, Clark RA, Ganesan AN. Effectiveness of wearable trackers on physical activity in healthy adults: Systematic review and meta-analysis of randomized controlled trials. JMIR MHealth UHealth. 2020;8. https://doi.org/10.2196/15576

27. Ringeval M, Wagner G, Denford J, Paré G, Kitsiou S. Fitbit-Based Interventions for Healthy Lifestyle Outcomes: Systematic Review and Meta-Analysis. J Med Internet Res. 2020;22:e23954. https://doi.org/10.2196/23954

28. Franssen WMA, Franssen GHLM, Spaas J, Solmi F, Eijnde BO. Can consumer wearable activity tracker-based interventions improve physical activity and cardiometabolic health in patients with chronic diseases? A systematic review and

meta-analysis of randomised controlled trials. Int J Behav Nutr Phys Act. 2020;17. https://doi.org/10.1186/s12966-020-00955-2

29. Ha L, Mizrahi D, Wakefield CE, Cohn RJ, Simar D, Signorelli C. The use of activity trackers in interventions for childhood cancer patients and survivors: A Systematic Review. J Adolesc Young Adult Oncol. 2020;10. https://doi.org/10.1089/jayao.2020.0099

30. Ridgers ND, McNarry MA, Mackintosh KA. Feasibility and effectiveness of using wearable activity trackers in youth: A systematic review. JMIR MHealth UHealth. 2016;4. https://doi.org/10.2196/mhealth.6540

31. Böhm B, Karwiese SD, Böhm H, Oberhoffer R. Effects of Mobile Health Including Wearable Activity Trackers to Increase Physical Activity Outcomes Among Healthy Children and Adolescents: Systematic Review. JMIR MHealth UHealth. 2019;7:e8298. https://doi.org/10.2196/mhealth.8298

32. Cajita MI, Kline CE, Burke LE, Bigini EG, Imes CC. Feasible but Not Yet Efficacious: A Scoping Review of Wearable Activity Monitors in Interventions Targeting Physical Activity, Sedentary Behavior, and Sleep. Curr Epidemiol Rep. 2020;7:25–38. https://doi.org/10.1007/s40471-020-00229-2

33. Egger M, Zellweger-Zähner T, Schneider M, Junker C, Lengeler C, Antes G. Language bias in randomised controlled trials published in English and German. Lancet. 1997;350:326–9. https://doi.org/10.1016/S0140-6736(97)02419-7

34. Moher D, Pham B, Lawson ML, Klassen TP. The inclusion of reports of randomised trials published in languages other than English in systematic reviews. Health Technol Assess. 2003;7:1–90. https://doi.org/10.3310/hta7410

35. Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. Cochrane Handbook for Systematic Reviews of Interventions. 6.2. New York: John Wiley & Sons; 2021.

36. Prince SA, Cardilli L, Reed JL, Saunders TJ, Kite C, Douillette K, et al. A comparison of self-reported and device measured sedentary behaviour in adults: a systematic review and meta-analysis. Int J Behav Nutr Phys Act. 2020;17:31. https://doi.org/10.1186/s12966-020-00938-3

37. Shephard RJ. Limits to the measurement of habitual physical activity by questionnaires. Br J Sports Med. 2003;37:197–206. http://dx.doi.org/10.1136/bjsm.37.3.197

38. Ruiz DC, Goransson A. Professional Android Wearables. New York: John Wiley & Sons; 2015.

39. Liu JY-W, Kor PPK, Chan CPY, Kwan RYC, Sze-Ki D. The effectiveness of a wearable activity tracker (WAT)-based intervention to improve physical activity levels in sedentary older adults: A systematic review and meta-analysis. Arch Gerontol Geriatr. 2020;91. https://doi.org/10.1016/j.archger.2020.104211

40. Sucharew H. Methods for Research Evidence Synthesis: The Scoping Review Approach. J Hosp Med. 2019;14:416. https://doi.org/10.12788/jhm.3248

41. Borenstein, M M, Hedges L, Higgins J, Rothstein H. Introduction to meta-analysis. 2nd ed. New York: John Wiley & Sons; 2009.

42. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. https://doi.org/10.1136/bmj.n71

43. World Health Organization. The health of youth. Geneva: World Health Organization; 1989.

44. Bromberg S. Consumer wristband activity monitors as a simple and inexpensive tool for remote heart failure monitoring. Toronto: University of Toronto; 2015.

45. Gaudet J, Gallant F, Bélanger M. A bit of fit: Minimalist intervention in adolescents based on a physical activity tracker. JMIR MHealth UHealth. 2017;5. https://doi.org/10.2196/mhealth.7647

46. Horne PJ, Hardman CA, Lowe CF, Rowlands AV. Increasing children's physical activity: A peer modelling, rewards and pedometer-based intervention. Eur J Clin Nutr. 2009;63:191–8. https://doi.org/10.1038/sj.ejcn.1602915

47. Routen A, Upton D, Edwards M, Peters D. The effect of pedometer step goal, feedback and self-monitoring interventions on accelerometer-measured physical activity in children. Grad J Sport Exerc Phys Educ Res. 2014;2:37–53.

48. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings

tables. J Clin Epidemiol. 2011;64:383–94. https://doi.org/10.1016/j.jclinepi.2010.04.026

49. Mayorga-Vega D, Bocanegra-Parrilla R, Ornelas M, Viciana J. Criterion-Related Validity of the Distance- and Time-Based Walk/Run Field Tests for Estimating Cardiorespiratory Fitness: A Systematic Review and Meta-Analysis. PloS One. 2016;11:e0151671. https://doi.org/10.1371/journal.pone.0151671

50. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;315:629. https://doi.org/10.1136/bmj.315.7109.629

51. Orwin RG. A fail-safe N for Effect Size in Meta-Analysis. J Educ Stat. 1983;8:157–9. https://doi.org/10.2307/1164923

52. Cohen J. A power primer. Psychol Bull. 1992;112:155–9. https://doi.org/10.1037/0033-2909.112.1.155

53. Duval S, Tweedie R. Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics. 2000;56:455–63. https://doi.org/10.1111/j.0006-341X.2000.00455.x

54. Remmert JE, Woodworth A, Chau L, Schumacher LM, Butryn ML, Schneider M. Pilot Trial of an Acceptance-Based Behavioral Intervention to Promote Physical Activity Among Adolescents. J Sch Nurs. 2019;35:449–61. https://doi.org/10.1177/1059840518786782

55. Leinonen A-M, Pyky R, Ahola R, Kangas M, Siirtola P, Luoto T, et al. Feasibility of gamified mobile service aimed at physical activation in young men: Population-based randomized controlled study (mopo). JMIR MHealth UHealth. 2017;5. https://doi.org/10.2196/mhealth.6675

56. Baldursdottir B, Taehtinen RE, Sigfusdottir ID, Krettek A, Valdimarsdottir HB. Impact of a physical activity intervention on adolescents' subjective sleep quality: a pilot study. Glob Health Promot. 2017;24:14–22. https://doi.org/10.1177/1757975915626112

57. Ermetici F, Zelaschi RF, Briganti S, Dozio E, Gaeta M, Ambrogi F, et al. Association between a school-based intervention and adiposity outcomes in

adolescents: The Italian "eAT" project. Obesity. 2016;24:687–95. https://doi.org/10.1002/oby.21365

58. Bronikowski M, Bronikowska M, Maciaszek J, Glapa A. Maybe it is not a goal that matters: A report from a physical activity intervention in youth. J Sports Med Phys Fitness. 2018;58:348–55. https://doi.org/10.23736/S0022-4707.16.06611-1

59. Kantanista A, Bronikowski M, Laudanska-Krzeminska I, Osinski W. Effects of pedometer-based walking intervention on psychological and biological variables in adolescent girls: does achieving goals make a difference? Med Sport. 2014;67:261–72.

60. Shore SM. The H.Y.P.P.E. initiative: A school-based physical activity promotion program. Ann Arbor: Temple University; 2010.

61. Rothstein H, Sutton A, Borenstein M. Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments. New York: John Wiley & Sons; 2006.

62. Aksornsri A. A pedometer-physical activity program to change self-efficacy and physical activity in Thai school-Age children. Ann Arbor: Saint Louis University; 2019.

63. Baldursdóttir B. Physical activity and well-being among adolescents. Gothenburg: University of Gothenburg; 2016.

64. Corepal R, Best P, O'Neill R, Kee F, Badham J, Dunne L, et al. A feasibility study of "The StepSmart Challenge" to promote physical activity in adolescents. Pilot Feasibility Stud. 2019;5. https://doi.org/10.1186/s40814-019-0523-5

65. Corr M, McMullen J, Morgan P, Barnes A, Murtagh E. Supporting Our Lifelong Engagement: Mothers and Teens Exercising (SOLE MATES); a feasibility trial. Women Health. 2019;1-18. https://doi.org/10.1080/03630242.2019.1688446

66. Corr M, Murtagh E. 'No one ever asked us': a feasibility study assessing the co-creation of a physical activity programme with adolescent girls. Glob Health Promot. 2020;27:34–43. https://doi.org/10.1177/1757975919853784

67. Duck AA, Hall KC, Klamm M, Temple M, Robinson JC. Physical activity and fitness: The feasibility and preliminary effectiveness of wearable activity tracker technology incorporating altruistic motivation in youth. J Spec Pediatr Nurs. 2021;26:e12313. https://doi.org/10.1111/jspn.12313

68. Duncan M, Birch S, Woodfield L. Efficacy of an integrated school curriculum pedometer intervention to enhance physical activity and to reduce weight status in children. Eur Phys Educ Rev. 2012;18:396–407. https://doi.org/10.1177/1356336X12450799

69. Evans EW, Abrantes AM, Chen E, Jelalian E. Using Novel Technology within a School-Based Setting to Increase Physical Activity: A Pilot Study in School-Age Children from a Low-Income, Urban Community. BioMed Res Int. 2017;2017. https://doi.org/10.1155/2017/4271483

70. Finkelstein EA, Tan YT, Malhotra R, Lee CF, Goh SS, Saw SM. A cluster randomized controlled trial of an incentive-based outdoor physical activity program. J Pediatr. 2013;163:167-72. https://doi.org/10.1016/j.jpeds.2013.01.009

71. Galy O, Yacef K, Caillaud C. Improving pacific adolescents' physical activity toward international recommendations: Exploratory study of a digital education app coupled with activity trackers. JMIR MHealth UHealth. 2019;7: e14854. https://doi.org/10.2196/14854

72. Grao-Cruces A, Ruiz-López R, Moral-García J-E, Ruiz-Ariza A, Martínez-López EJ. Effects of a steps/day programme with evaluation in physical education on body mass index in schoolchildren 11-12 years of age. Kinesiology. 2016;48:132–41. https://doi.org/10.26582/k.48.1.2

73. Groffik D, Frömel K, Pelclová J. Pedometers as a method for modification of physical activity in students. J Hum Kinet. 2008;20:131–8. https://doi.org/10.2478/v10078-008-0025-7

74. Hardman CA, Horne PJ, Lowe CF. A home-based intervention to increase physical activity in girls: the Fit "n" Fun Dudes program. J Exerc Sci Fit. 2009;7:1–8. https://doi.org/10.1016/S1728-869X(09)60001-0

75. Hardman CA, Horne PJ, Fergus Lowe C. Effects of rewards, peer-modelling and pedometer targets on children's physical activity: a school-based intervention study. Psychol Health. 2011;26:3–21. https://doi.org/10.1080/08870440903318119

76. Jauho A-M, Pyky R, Ahola R, Kangas M, Virtanen P, Korpelainen R, et al. Effect of wrist-worn activity monitor feedback on physical activity behavior: A

randomized controlled trial in Finnish young men. Prev Med Rep. 2015;2:628–34. https://doi.org/10.1016/j.pmedr.2015.07.005

77. Kantanista A, Bronikowski M, Laudanska-Krzeminska I, Krol-Zielinska M, Osinski W. Positive effect of pedometer-based walking intervention on body image and physical activity enjoyment in adolescent girls. Biomed Hum Kinet. 2017;9:34–42. https://doi.org/10.1515/bhk-2017-0006

78. Kerner C, Burrows A, McGrane B. Health wearables in adolescents: implications for body satisfaction, motivation and physical activity. Int J Health Promot Educ. 2019;57:191–202. https://doi.org/10.1080/14635240.2019.1581641

79. Larsen B, Benitez T, Cano M, Dunsiger SS, Marcus BH, Mendoza-Vasconez A, et al. Web-Based Physical Activity Intervention for Latina Adolescents: Feasibility, Acceptability, and Potential Efficacy of the Niñas Saludables Study. J Med Internet Res. 2018;20:e170. https://doi.org/10.2196/jmir.9206

80. Linck DT. Female teens step it up with the Fitbit Zip: A randomized controlled pilot study. Augusta: Augusta University; 2018.

81. Lubans D, Morgan P. Evaluation of an extra-curricular school sport programme promoting lifestyle and lifetime activity for adolescents. J Sports Sci. 2008;26:519–29. https://doi.org/10.1080/02640410701624549

82. Lubans DR, Morgan PJ, Callister R, Collins CE. Effects of integrating pedometers, parental materials, and E-mail support within an extracurricular school sport intervention. J Adolesc Health. 2009;44:176–83. https://doi.org/10.1016/j.jadohealth.2008.06.020

83. Macias-Cervantes MH, Malacara JM, Garay-Sevilla ME, Díaz-Cisneros FJ. Effect of recreational physical activity on insulin levels in Mexican/Hispanic children. Eur J Pediatr. 2009;168:1195–202. https://doi.org/10.1007/s00431-008-0907-7

84. Morris JL, Daly-Smith A, Defeyter MA, McKenna J, Zwolinsky S, Lloyd S, et al. A pedometer-based physically active learning intervention: The importance of using preintervention physical activity categories to assess effectiveness. Pediatr Exerc Sci. 2019;31:356–62. https://doi.org/10.1123/pes.2018-0128

85. Newton RLJ, Marker AM, Allen HR, Machtmes R, Han H, Johnson WD, et al. Parent-targeted mobile phone intervention to increase physical activity in sedentary

children: randomized pilot trial. JMIR MHealth UHealth. 2014;2:e48. https://doi.org/10.2196/mhealth.3420

86. Schofield L, Mummery WK, Schofield G. Effects of a controlled pedometer-intervention trial for low-active adolescent girls. Med Sci Sports Exerc. 2005;37:1414–20. https://doi.org/10.1249/01.mss.0000174889.89600.e3

87. Shapiro JR, Bauer S, Hamer RM, Kordy H, Ward D, Bulik CM. Use of text messaging for monitoring sugar-sweetened beverages, physical activity, and screen time in children: a pilot study. J Nutr Educ Behav. 2008;40:385–91. https://doi.org/10.1016/j.jneb.2007.09.014

88. Shimon JM, Petlichkoff LM. Impact of pedometer use and self-regulation strategies on junior high school physical education students' daily step counts. J Phys Act Health. 2009;6:178–84. https://doi.org/10.1123/jpah.6.2.178

89. Shore SM, Sachs ML, DuCette JP, Libonati JR. Step-Count Promotion Through a School-Based Intervention. Clin Nurs Res. 2014;23:402–20. https://doi.org/10.1177/1054773813485240

90. Smith JJ, Morgan PJ, Plotnikoff RC, Dally KA, Salmon J, Okely AD, et al. Smart-phone obesity prevention trial for adolescent boys in low-income communities: the ATLAS RCT. Pediatrics. 2014;134:e723-31. https://doi.org/10.1542/peds.2014-1012

91. Thompson D, Cantu D, Ramirez B, Cullen KW, Baranowski T, Mendoza J, et al. Texting to increase adolescent physical activity: Feasibility assessment. Am J Health Behav. 2016;40:472–83. https://doi.org/10.5993/AJHB.40.4.9

92. Wang S-H. The effects of goal setting on female middle school students' physical activity levels and motivation toward exercise. Ann Arbor: The Florida State University; 2004.

93. Zizzi S, Vitullo E, Rye J, O'hara-Tompkins N, Abildso C, Fisher B, et al. Impact of a three-week pedometer intervention on high school students' daily step counts and perceptions of physical activity. Am J Health Educ. 2006;37:35–40. https://doi.org/10.1080/19325037.2006.10598875

94. Lau PWC, Lau EY, Wong DP, Ransdell L. A Systematic review of information and communication technology-based interventions for promoting physical activity

behavior change in children and adolescents. J Med Internet Res. 2011;13. https://doi.org/10.2196/jmir.1533

95. Carlin A, Murphy MH, Gallagher AM. Current influences and approaches to promote future physical activity in 11–13 year olds: a focus group study. BMC Public Health. 2015;15:1270. https://doi.org/10.1186/s12889-015-2601-9

96. Telama R, Yang X, Leskinen E, Kankaanpää A, Hirvensalo M, Tammelin T, et al. Tracking of physical activity from early childhood through youth into adulthood. Med Sci Sports Exerc. 2014;46. https://doi.org/10.1249/MSS.0000000000000181

97. Tudor-Locke C, Craig CL, Thyfault JP, Spence JC. A step-defined sedentary lifestyle index: < 5000 steps/day. Appl Physiol Nutr Metab. 2013;38:100–14. https://doi.org/10.1139/apnm-2012-0235

98. Katapally TR, Muhajarine N. Towards Uniform Accelerometry Analysis: A Standardization Methodology to Minimize Measurement Bias Due to Systematic Accelerometer Wear-Time Variation. J Sports Sci Med. 2014;13:379.

99. Mayorga-Vega D, Martinez-Baena A, Viciana J. Does school physical education really contribute to accelerometer-measured daily physical activity and non sedentary behaviour in high school students? J Sports Sci. 2018;36:1913–22. https://doi.org/10.1080/02640414.2018.1425967

100. Migueles JH, Cadenas-Sanchez C, Ekelund U, Delisle Nyström C, Mora-Gonzalez J, Löf M, et al. Accelerometer Data Collection and Processing Criteria to Assess Physical Activity and Other Outcomes: A Systematic Review and Practical Considerations. Sports Med. 2017;47:1821–45. https://doi.org/10.1007/s40279-017-0716-0

101. Owen KB, Smith J, Lubans D, Ng J, Lonsdale C. Self-determined motivation and physical activity in children and adolescents: a systematic review and meta-analysis. Prev Med. 2014;67:270–9. https://doi.org/10.1016/j.ypmed.2014.07.033

102. Mayorga-Vega D, Viciana J. Las clases de educación física solo mejoran la capacidad cardiorrespiratoria de los alumnos con menor condición física: un estudio de intervención controlado. Nutr Hosp. 2015;32:330–5. https://dx.doi.org/10.3305/nh.2015.32.1.8919

103. Guijarro-Romero S, Mayorga-Vega D, Casado-Robles C, Viciana J. An intermittent physical fitness teaching unit only improves cardiorespiratory fitness levels of students with an unhealthy physical fitness profile. Retos. 2020;38:8-15. https://doi.org/10.47197/retos.v38i38.73605

104. Love RE, Adams J, van Sluijs EMF. Equity effects of children's physical activity interventions: a systematic scoping review. Int J Behav Nutr Phys Act. 2017;14:134. https://doi.org/10.1186/s12966-017-0586-8

105. Riddoch CJ, Bo Andersen L, Wedderkopp N, Harro M, Klasson-Heggebø L, Sardinha LB, et al. Physical activity levels and patterns of 9- and 15-yr-old European children. Med Sci Sports Exerc. 2004;36:86–92. https://doi.org/10.1249/01.MSS.0000106174.43932.92

106. Sallis JF, Prochaska JJ, Taylor WC. A review of correlates of physical activity of children and adolescents. Med Sci Sports Exerc. 2000;32:963–75. https://doi.org/10.1097/00005768-200005000-00014

107. Bandura A. Health promotion by Social Cognitive Means. Health Educ Behav. 2004;31:143–64. https://doi.org/10.1177/1090198104263660

108. Ryan RM, Deci EL. Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. Contemp Educ Psychol. 2020;61:101860. https://doi.org/10.1016/j.cedpsych.2020.101860

109. Hagger MS, Chatzisarantis NLD. The Trans-Contextual Model of Autonomous Motivation in Education: Conceptual and Empirical Issues and Meta-Analysis. Rev Educ Res. 2016;86:360–407. https://doi.org/10.3102/0034654315585005

110. Rhodes RE, Janssen I, Bredin SSD, Warburton DER, Bauman A. Physical activity: Health impact, prevalence, correlates and interventions. Psychol Health. 2017;32:942–75. https://doi.org/10.1080/08870446.2017.1325486

111. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med. 2013;46:81–95. https://doi.org/10.1007/s12160-013-9486-6

112. Bort-Roig J, Gilson ND, Puig-Ribera A, Contreras RS, Trost SG. Measuring and Influencing Physical Activity with Smartphone Technology: A Systematic Review. Sports Med. 2014;44:671–86. https://doi.org/10.1007/s40279-014-0142-5

113. Cairney J, Dudley D, Kwan M, Bulten R, Kriellaars D. Physical Literacy, Physical Activity and Health: Toward an Evidence-Informed Conceptual Model. Sports Med. 2019;49:371–83. https://doi.org/10.1007/s40279-019-01063-3

114. Bassett DR, Toth LP, LaMunion SR, Crouter SE. Step Counting: A Review of Measurement Considerations and Health-Related Applications. Sports Med. 2017;47:1303–15. https://doi.org/10.1007/s40279-016-0663-1

115. Richardson S, Mackinnon D. Becoming your own device: Self-tracking challenges in the workplace. Can J Sociol. 2018;43:265–90. https://doi.org/10.29173/cjs28974

116. Fairclough SJ, Noonan R, Rowlands AV, Van Hees V, Knowles Z, Boddy LM. Wear Compliance and Activity in Children Wearing Wrist- and Hip-Mounted Accelerometers. Med Sci Sports Exerc. 2016;48. https://doi.org/10.1249/MSS.0000000000000771

117. Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol. 2011;64:401–6. https://doi.org/10.1016/j.jclinepi.2010.07.015

SECTION 2

VALIDITY OF INSTRUMENTS FOR ASSESSING FACTORS INFLUENCING STUDENTS' PHYSICAL ACTIVITY HABITS

PAPERS VIII - IX

Paper VIII

CONOCIMIENTO DEL ENTORNO PARA LA PRÁCTICA DE ACTIVIDAD FÍSICA EN ESCOLARES (CEPAF): DESARROLLO Y VALIDACIÓN DE UNA PRUEBA ESCRITA OBJETIVA DE ELECCIÓN MÚLTIPLE

Casado-Robles, C., Mayorga-Vega, D., Guijarro-Romero, S., & Viciana, J.

Journal of Sport and Health Research

2021, *13*(2), 223-244.

Paper VIII

ABSTRACT

Objective: The objective of the present study was to develop and validate an *ad hoc* written test to evaluate the students' knowledge about the environment close to the educational center for the practice of physical activity (CEPAF).

Methods: 256 students (50% female) from the second and third-grade of Compulsory Secondary Education (mean age = 14.05 ± 0.89 years) participated in the study. The validation of the test followed six phases: (1) Construction and definition of the structure and purpose of the test; (2) Content validation by experts; (3) Pilot evaluation; (4) Refinement and application of the final instrument; (5) Evaluation of test-retest reliability; and (6) Evaluation of discriminant validity.

Results: Finally, the test consisted of 30 objective multiple-choice questions. The overall difficulty obtained was moderate (mean = 11.10 points out of 30). All questions obtained adequate indices of difficulty and discrimination, and met the qualitative criteria established by the experts. The test-retest reliability was adequate (ICC = 0.72). The test presents adequate discriminant validity, since the students who experienced the intervention program improved the test score with respect to the control group ($p < 0.001$).

Conclusions: The results show that it is a valid and reliable measuring instrument to gather information about the knowledge of the immediate environment for the schoolchildren's practice of physical activity. The present study represents an important contribution to the scientific literature and with valuable practical implications for the planning of Physical Education subject.

KEYWORDS

Physical Education. Sports practice. Adolescents. Compulsory Secondary Education. Evaluation

Introducción

La práctica regular de actividad física (AF) se considera un comportamiento muy importante para el desarrollo de beneficios en la salud de los adolescentes (Guthold, Stevens, Riley y Bull, 2019). Por ejemplo, la práctica regular de AF en los jóvenes contribuye al desarrollo de un sistema cardiovascular sano, beneficios psicológicos y beneficios sociales (Guthold et al., 2019; World Health Organization, WHO, 2014). Por este motivo, se han establecido recomendaciones diarias de AF, consistentes en un mínimo de 60 minutos de AF moderada-vigorosa en los adolescentes (WHO, 2010). Desafortunadamente, son alarmantes los bajos niveles de AF a nivel mundial que presentan los adolescentes, donde el 81% no cumple con las recomendaciones diarias (Guthold et al., 2019).

Por ello, la promoción de práctica de AF en los jóvenes es una prioridad de salud pública (WHO, 2018). Dentro de los posibles escenarios para su promoción, el contexto escolar, y especialmente la asignatura de Educación Física (EF), son considerados clave (WHO, 2018). Por un lado, al igual que en la mayoría de los países de nuestro entorno (European Commission, 2013), en el marco curricular de la asignatura de EF en España, se establece como uno de sus fines el desarrollo y consolidación de hábitos saludables de AF regular en la etapa de Educación Secundaria Obligatoria (ESO), así como la ocupación activa del tiempo libre de los adolescentes (Ministerio de Educación, Cultura y Deporte, 2015). En esta línea, la institución *Society of Health and Physical Educators* (SHAPE America, 2014) destaca que la asignatura de EF juega un papel importante en mejorar la "alfabetización física" de los estudiantes y sugieren a los profesores de EF su promoción mejorando la motivación de los estudiantes, la competencia física y el conocimiento necesario para mantener un estilo de vida activo.

Actualmente, el sistema educativo está enfocado hacia el aprendizaje por competencias (Ministerio de Educación, Cultura y Deporte, 2013), siendo uno de sus pilares fundamentales el principio de transferibilidad del aula a la vida (Miklos, 1999). Específicamente, para la consolidación de hábitos de salud y ocupación activa del tiempo libre, la EF tiene como objetivo el desarrollo de la competencia motriz. Para alcanzar dicha competencia, se debe reducir progresivamente la dependencia del alumno hacia el profesor, para conseguir que el alumno sea autónomo usando el entorno que le rodea (Ministerio de Educación, Cultura y Deporte, 2015). Además, en el reciente modelo de planificación de la EF en base a tres ejes, propuesto por Viciana y Mayorga-Vega (2018), uno de los ejes de planificación está referido a proporcionar esa autonomía progresiva al alumnado. Además, la "alfabetización física" se ve facilitada dentro de un plan de estudios en el que los estudiantes puedan establecer fuertes conexiones entre el conocimiento adquirido y las experiencias de la vida real (Ennis, 2011). De hecho, la adquisición e integración de conocimientos es uno de los objetivos fundamentales de la asignatura de EF, más allá de la mera práctica de AF durante las clases (Ministerio de Educación, Cultura y Deporte, 2015).

Por todas las razones anteriormente expuestas, el conocimiento que los estudiantes tienen sobre su entorno próximo para la práctica de AF debería ser un ámbito importante en la evaluación de la asignatura de EF. Este conocimiento podría influir en la cantidad de práctica de AF en su tiempo libre y, por tanto, en el cumplimiento de las recomendaciones diarias. La forma más común de evaluar la adquisición de aspectos cognitivos es mediante una prueba de conocimiento escrita (Baumgartner, Jackson, Mahar y Rowe, 2015). Estudios previos han construido y validado pruebas de evaluación del conocimiento relacionados con la asignatura de EF como, por ejemplo, sobre conocimiento del balance energético (Chen, Zhu y Kang, 2017), conocimiento técnico, táctico y reglamentario de los deportes (He, Ward y Wang, 2018), o conocimiento de la aptitud física (Chen, Chen, Sun y Zhu, 2013). Desafortunadamente, y a pesar de la importancia social de este asunto reflejada en los currículos actuales, de lo que conocemos, no hay estudios previos que desarrollen y validen pruebas para evaluar el conocimiento que poseen los escolares de su entorno para la práctica de AF. Consecuentemente, el objetivo del presente estudio fue construir y someter a un proceso de validación una prueba escrita *ad hoc* para crear una herramienta válida y fiable para la evaluación del conocimiento del entorno para la práctica de AF (CEPAF) en estudiantes de ESO.

Material y Métodos

Participantes

El protocolo del estudio fue primero aprobado por el Comité de Ética para Estudios Humanos de la Universidad de Granada. Después, se contactó con el director y profesor de EF de un centro público de ESO seleccionado por conveniencia. Se les informó de las características principales del proyecto y se les solicitó permiso para llevar a cabo el estudio. Tras obtener la aprobación del centro, los estudiantes y sus tutores legales fueron informados sobre las características del proyecto. Antes de participar en el estudio, se obtuvieron los consentimientos y asentimientos informados por escrito firmados sus tutores legales y los adolescentes, respectivamente. El reclutamiento se realizó entre noviembre y diciembre de 2018 y la toma de datos entre enero y marzo de 2019.

Todos los estudiantes matriculados en el segundo y tercer curso de ESO (13-15 años) del centro seleccionado fueron invitados a participar en el estudio (cursos donde se obtuvo el permiso del centro). El centro educativo estaba ubicado en el área metropolitana de Granada. Según los informes del centro, todas las familias de los estudiantes tenían un nivel socioeconómico medio. Se consideraron los siguientes criterios de inclusión: (a) estar matriculado en los cursos en los que se obtuvo el permiso del centro; (b) no padecer ningún trastorno de salud que les impida participar en AF; (c) presentar el correspondiente asentimiento informado por escrito firmado por los propios adolescentes; y (d) presentar el correspondiente consentimiento informado por escrito firmado por sus tutores legales. Se consideró el siguiente criterio de exclusión: (a) no haber contestado a todas las preguntas de la prueba.

La Figura 1 muestra el diagrama de flujo de los participantes incluidos en cada estudio. Todos los estudiantes invitados en cada estudio, aceptaron voluntariamente a participar y cumplieron con los criterios de inclusión. Sin embargo, entre el 2,67-3,43% de los estudiantes fueron excluidos por no cumplir satisfactoriamente con el criterio de exclusión. La muestra utilizada para realizar el estudio piloto y los índices de dificultad y discriminación del instrumento definitivo cumplimentaron la prueba una vez. Posteriormente, para estudiar la fiabilidad del instrumento definitivo, se utilizó una submuestra (submuestra fiabilidad) que repitieron la prueba 6 semanas más tarde. Para estudiar la validez discriminante del instrumento definitivo, seis clases pre-establecidas por el centro educativo fueron distribuidas de forma aleatoria y balanceadas por curso, al grupo control o intervención. La Tabla 1 muestra las características generales de los participantes incluidos.

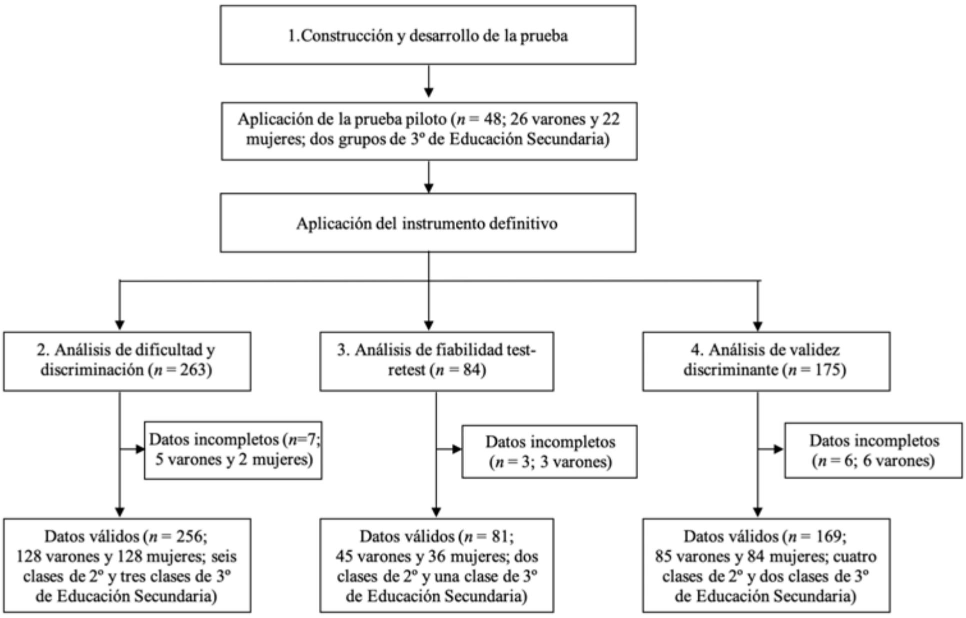


Figura 1. Diagrama de flujo de los participantes incluidos en cada estudio.

Tabla 1. Características generales de los estudiantes analizados[a]

	Piloto	Dificultad/ discriminación	Fiabilidad	Validez discriminante
N	48	256	81	169
Edad (años)	14,47 (0,65)	14,05 (0,89)	13,99 (0,77)	14,01 (0,79)
Edad (rango)	14-16	13-16	13-16	13-16
Curso (2º/3º)	0/48	163/93	50/31	108/61
Género (Mujer/Varón)	22/26	128/128	36/45	84/85

Nota. [a] Los datos están reportados como frecuencia o media (desviación estándar).

El cálculo del tamaño de la muestra a priori se estimó para cada estudio de la siguiente manera: (a) Para el estudio piloto de la prueba se siguió la recomendación de al menos 25-50 participantes (Fontes, García-Gallego, Garriga, Pérez-Llantada y Sarriá, 2007); (b) Para que los resultados del análisis de dificultad y discriminación de las preguntas definitivas sean fiables y válidos, se siguió la recomendación de una muestra mayor a 100 participantes (Baumgartner et al., 2015); (c) Para conseguir un coeficiente de correlación intraclase (CCI) superior a 0,6 en la fiabilidad test-retest, considerando α = 0,05 y 1 - β = 0,80, se estimó un mínimo de 45 participantes (Walter, Eliasziw y Donner, 1998); y (d) Respecto al estudio de la validez discriminante, el tamaño mínimo de muestra (128 participantes) se estimó con la función *F* del ANOVA de un factor mediante el programa G*Power versión 3.1.9.4 para Windows. Los parámetros se establecieron como: *f* = 0,25, α = 0,05, 1 - β = 0,80 y número de grupos = 2.

Desarrollo de la prueba

El procedimiento de construcción de la prueba se basó en los criterios sugeridos por la literatura previa (Baumgartner et al., 2015; Eignor, 2013), que se describen a continuación.

Paso 1. Construcción de la prueba

Para la recogida de información se diseñó una prueba escrita *ad hoc* objetiva de elección múltiple titulada "Prueba sobre el conocimiento del entorno próximo para la práctica de AF (CEPAF)", con el objetivo de evaluar el conocimiento que posee el alumnado de ESO sobre su entorno próximo para realizar AF. Para determinar la estructura de la prueba se realizó: (a) Una revisión bibliográfica sobre el término "conocimiento", ya que al ser un concepto multidimensional se debía decidir una clasificación de sus dimensiones acorde al objetivo; y (b) Un panel de expertos formado por doctores e investigadores en EF analizó para qué tipo de actividades utilizan los adolescentes el entorno y cómo pueden utilizarlo para realizar AF, con el objetivo de decidir qué contenidos del currículo de EF eran susceptibles de ser incluidos.

Respecto a la estructura del conocimiento a evaluar, se consideró la clasificación propuesta por Zack (1999) como la más adecuada según el objetivo, dividiendo al conocimiento en: (a) Declarativo, referido a la descripción de un fenómeno y el conocimiento explícito de conceptos (saber el qué de algo); (b) Procedimental, referido a cómo ocurre o se realiza algo mediante la ejecución de procedimientos o técnicas para la consecución de un fin (saber cómo hacer algo); (c) Causal, referido a la relación causa-efecto de las acciones (saber por qué ocurre algo). En segundo lugar respecto a los contenidos del área de EF en ESO se dividió la prueba en dos grandes bloques: (a) Condición física y motriz, enfocado al conocimiento sobre actividades y ejercicios que pueden utilizar para el desarrollo de la condición física en su entorno próximo desde un enfoque saludable; y (b) Juegos y deportes, referido al conocimiento sobre las posibilidades de práctica físico-deportiva de colaboración, oposición y colaboración-oposición, tanto tradicionales como alternativas, que pueden realizar en su entorno próximo. La prueba final se conformó por 30 preguntas, con cinco preguntas para cada dimensión del conocimiento (declarativo, procedimental y causal) y bloque de contenidos (condición física y juegos y deportes).

Tras determinar el objetivo de la prueba y su división en cada bloque de conocimiento, se realizó una revisión bibliográfica sobre el tema para el desarrollo de las preguntas. En primer lugar, se decidió la naturaleza de la prueba siguiendo las instrucciones de Baumgartner et al. (2015), optándose por una prueba de tipo objetiva y de elección múltiple con cuatro opciones de respuesta (Raymond, Stevens y Bucak, 2019) y una única opción correcta para todas las preguntas. Al finalizar esta fase, se desarrolló una amplia batería de 60 preguntas iniciales. Además, se desarrolló el encabezado de la prueba indicando una pequeña introducción de presentación e instrucciones para completar la prueba correctamente, así como un agradecimiento por su participación. En dichas instrucciones, se remarcaba que los estudiantes debían contestar todas las preguntas y que no había ninguna penalización por respuestas erróneas.

Paso 2. Validación de contenido por expertos

Tras el desarrollo de las preguntas, se utilizó un panel de expertos en EF para evaluar la validez de contenido del instrumento, es decir, el grado en que un instrumento cumple con el objetivo de su construcción y mide lo que realmente quiere medir para poder considerar la prueba válida (Eignor, 2013). El procedimiento seguido fue el juicio de expertos, siguiendo el método *Delphi*. Esta técnica se basa en una opinión de expertos cualificados y que pueden emitir juicios y valoraciones sobre el tema. Siguiendo las indicaciones de Escobar-Pérez y Cuervo-Martínez (2008) y Corral (2009), se siguieron estos pasos:

(1) Tras la revisión bibliográfica y el desarrollo de una batería de 60 preguntas, se determinaron los expertos para validar el contenido del instrumento. Se seleccionaron tres docentes e investigadores con vinculación científica y experiencia académica en el área de EF para garantizar la idoneidad y contribución de este análisis (Corral, 2009).

(2) Cada experto fue informado detalladamente sobre el objetivo del instrumento, su estructura y los indicadores sobre los que realizar su análisis. Luego, calificaron las preguntas como adecuadas/inadecuadas, basando su evaluación en cuatro categorías: (a) adecuación y complejidad (la pregunta está adecuadamente formulada para los destinatarios a evaluar); (b) coherencia o pertinencia con los objetivos de la investigación (recoge información relevante a la investigación); (c) claridad en la redacción (se comprenden fácilmente y la relación sintáctica y

semántica es adecuada); y (d) existencia de sesgo en la formulación (si sugieren o no una de las respuestas). Además, se comprobó que no hubiese errores gramaticales y que la introducción e instrucciones fueran claras. Posteriormente, cada experto emitió su juicio individual y sin contacto con el resto de expertos para evitar sesgos.

(3) Tras obtener el informe de los expertos, el investigador responsable de elaborar la prueba revisó las preguntas basándose en los juicios emitidos y generó un informe con todas las respuestas para analizar los datos y tomar decisiones. Aquellas preguntas con 100% de coincidencia favorable (adecuadas)/desfavorables (inadecuadas) entre los jueces quedaron incluidas/excluidas. Posteriormente el investigador responsable se reunió con los expertos para debatir sobre las opiniones de otros pares y volver a analizar las preguntas con la retroalimentación de otros participantes. Esto se realizó durante tres rondas, para revisar, reformular y/o sustituir aquellas preguntas en las que no había una coincidencia total entre expertos en los indicadores evaluados. Finalmente, tras realizar las modificaciones pertinentes, basándose en la responsabilidad compartida de los participantes, se editó de forma definitiva la batería de 60 preguntas.

Paso 3. Pilotaje

Se diseñaron dos pruebas piloto independientes y equivalentes a la prueba final (30 preguntas) para que los participantes pudieran responderlas en el tiempo establecido y no saturarlos con un elevado número de preguntas (Argimón y Jiménez, 2000). Ambas pruebas se conformaron atendiendo de forma equitativa a todas las dimensiones establecidas. La intención del pilotaje fue valorar la comprensión de las preguntas y respuestas en una muestra con las mismas características y nivel educativo que la población objeto de estudio, así como calcular la dificultad de dichas preguntas.

La prueba piloto se administró en las mismas condiciones que se iba a realizar posteriormente la prueba definitiva (Fontes et al., 2007). Durante la administración de la prueba, además se recogieron de forma cualitativa las dudas que surgieron para poder corregir posibles errores antes de conformar la prueba definitiva. El alumnado empleó una media de 22 minutos y 48 segundos (13-38 minutos) en finalizar la prueba. Finalmente, los participantes respondieron a un cuestionario para valorar la prueba de 10 preguntas con una escala tipo Likert de cinco puntos que iba desde 1 "Absolutamente en desacuerdo" hasta 5 "Absolutamente de acuerdo" (Tabla 2).

Tabla 2. Resultados de la escala de valoración de la prueba piloto

	Prueba A	Prueba B	Media
	Media (DE)	Media (DE)	Media (DE)
1. He entendido perfectamente las instrucciones para cumplimentar la prueba	4,1 (1,1)	4,1 (1,0)	4,1 (1,0)
2. He entendido perfectamente todas las palabras utilizadas en la prueba	4,2 (1,0)	4,0 (0,9)	4,1 (0,9)
3. He comprendido perfectamente las preguntas de la prueba	4,4 (0,9)	4,0 (0,9)	4,2 (0,9)
4. He entendido cómo debía contestar a las distintas preguntas	4,0 (1,2)	3,9 (1,2)	4,0 (1,2)
5. He sabido utilizar el sistema de respuesta según las respuestas "a/b/c/d" con sólo una respuesta correcta	4,6 (0,9)	4,0 (1,4)	4,3 (1,2)
6. El tema que trata la prueba me ha resultado interesante	3,9 (1,2)	3,8 (1,1)	3,8 (1,1)
7. Todas las preguntas me han parecido importantes para el tema que trata la prueba	4,1 (1,2)	3,5 (1,3)	3,8 (1,3)
8. La presentación de la prueba me ha parecido clara, me ha resultado fácil leer las preguntas	4,2 (1,1)	3,7 (1,2)	4,0 (1,2)
9. La prueba no me ha resultado liosa, cansada o pesada	3,6 (1,2)	3,3 (1,2)	3,4 (1,2)
10. En conjunto la prueba no presenta ningún problema	4,3 (1,1)	3,8 (1,0)	4,0 (1,1)
Global	4,1 (1,1)	3,8 (1,1)	4,0 (1,1)

Nota. DE = Desviación estándar.

Seguidamente se calcularon los índices cuantitativos de validez relativos a las preguntas (i.e., dificultad y discriminación de cada pregunta) en una hoja de Microsoft Office Excel 2016 (Microsoft® Corporation) (Baumgartner et al., 2015). No obstante, los índices relativos a la prueba global (i.e., dificultad y fiabilidad de la prueba completa) no se calcularon porque se aplicaron dos pruebas piloto independientes para finalmente conformar una única final con las preguntas con mejores índices. Los índices calculados fueron los propuestos por Baumgartner et al. (2015), comparándose los resultados para el 25% de participantes que obtuvo la mejor puntuación (grupo alto) y peor puntuación (grupo bajo). Específicamente, los índices calculados fueron: (a) *Dificultad de la pregunta*: Porcentaje de personas que aciertan cada pregunta. La pregunta es más fácil cuanto más se acerque al 100% de aciertos. Fórmula: [(Respuestas correctas grupo alto + respuestas correctas grupo bajo)/(numero estudiantes en grupo alto + numero estudiantes en grupo bajo)] x 100; (b) *Índice de discriminación*: Cómo discrimina esa pregunta entre el grupo que mejor y peor hace la prueba. Es un valor numérico entre -1 y +1. Las preguntas que mejor discriminan son aquellas que presenten una discriminación positiva entre 0 y 1. Si la discriminación es 0 o menor, esa pregunta debería ser revisada. Fórmula: [(Numero respuestas correctas en el grupo alto - numero respuestas correctas en el grupo bajo)/Número de estudiantes en cada grupo].

Finalmente, se seleccionaron las 30 preguntas que presentaban mejores índices y se comprobó que cumplieran con los estándares para evaluar las pruebas de respuesta múltiple propuestos por Baumgartner et al. (2015). Específicamente, para considerar la prueba adecuada tenía que cumplirse: (a) Menos del 5% de preguntas deben tener un índice de dificultad mayor al 90% y menos del 5% un índice de dificultad menor al 10%; (b) Más del 25% de las preguntas deben tener índices de discriminación mayores a 0,4; (c) Más del 25% de las preguntas deben tener índices de discriminación entre 0,21 y 0,39; (d) Más del 15% de las preguntas deben tener índices entre 0 y 0,20; (e) Menos del 5% de las preguntas deben tener índices de dificultad negativos; (f) En cada pregunta, todas las respuestas posibles (i.e., A, B, C y D) deben haber sido escogida por al menos el 5% de los participantes.

Por lo tanto, se eliminaron 30 preguntas por los siguientes motivos: (a) cuatro preguntas por índices de dificultad muy elevados o muy bajos; (b) ocho preguntas por índices de discriminación iguales a cero o negativos; (c) 10 preguntas porque alguna respuesta no fue escogida por ningún participante; (d) tres preguntas por producirse solapamiento de información con otras que presentaban índices mejores y que podría suponer que se acertara a raíz de otra pregunta; y (e) cinco preguntas para mantener una distribución equitativa en cada uno de los bloques de conocimiento-contenido. Por este motivo, en algunos casos, a pesar de que los resultados cuantitativos de las preguntas fueran relativamente peores que en otras, se escogieron para cubrir equitativamente todas las dimensiones. Los resultados obtenidos con las 30 preguntas seleccionadas estuvieron acorde a todos los criterios expuestos en el párrafo anterior, excepto en cuatro preguntas.

Paso 4. Depuración de la prueba y aplicación del instrumento definitivo

Tras la prueba piloto, se estudiaron las cuatro preguntas que no cumplían con los estándares cuantitativos, pero sí cumplían con el análisis cualitativo de los expertos. Aunque tres preguntas obtuvieron un índice de discriminación negativo durante la prueba piloto, tras una exhaustiva evaluación, se optó por mantenerlas dado que los expertos consideraron que tenían una validez de contenido adecuada. Además, en una pregunta había una respuesta que no fue escogida por ningún participante. En este caso, los expertos consideraron que podría deberse a que era una respuesta demasiado obvia de ser incorrecta y por ello se modificó para incrementar su verosimilitud en la prueba definitiva.

Por último, se tuvieron en cuenta las anotaciones cualitativas de la prueba piloto, con el objetivo de incluir aclaraciones y mejorar la claridad de aquellas preguntas que habían suscitado dudas en los estudiantes para facilitar su contestación (e.g., definición del juego del frontón en la pregunta "¿Dónde puedo practicar el frontón (o pelota-mano que consiste en golpear una bola frente a la pared) en el pueblo de Churriana de la Vega?"). Tras incluir estas pequeñas modificaciones (seis preguntas), se realizó una última reunión con los expertos para evaluar la validez de contenido de las preguntas modificadas según los pasos explicados previamente y se comprobó que eran adecuadas para ser incluidas en la prueba final.

La recogida de datos se llevó a cabo en las mismas condiciones que se había realizado la prueba piloto. La evaluación se realizó durante las clases de EF de los participantes, en su aula habitual, en silencio y garantizando el anonimato de sus respuestas. En primer lugar, el investigador principal realizó una breve introducción y explicó las instrucciones para cumplimentar correctamente el cuestionario. Posteriormente, los estudiantes tuvieron 40 minutos para completar el mismo.

Paso 5. Submuestra fiabilidad

La fiabilidad de la prueba fue calculada posteriormente con una submuestra mediante el método test-retest para establecer la estabilidad temporal o consistencia entre las puntuaciones de la prueba aplicada en dos momentos diferentes sobre el mismo grupo (Eignor 2013; Gomez, Vance y Stavropoulos, 2018). Se utilizó el CCI y su intervalo de confianza al 95% (95% IC) con un modelo mixto de dos factores de tipo acuerdo absoluto y basado en las medidas promedio (Koo y Li, 2016). Los valores de CCI < 0,50 son indicativos de una fiabilidad pobre, 0,50-0,75 moderada, 0,75-0,90 fiabilidad buena, y > 0,90 fiabilidad excelente (Koo y Li, 2016). Además de la fiabilidad de la prueba global, se calculó la fiabilidad de cada una de las dimensiones independientes (Eignor, 2013). Los análisis estadísticos se realizaron utilizando el programa SPSS versión 23.0 para Mac (IBM® SPSS® Statistics).

Paso 6. Submuestra intervención

Con la submuestra intervención se aplicó un programa de intervención para mejorar el conocimiento del entorno para la práctica de AF, para comprobar la validez discriminante de la prueba. La validez discriminante es un indicador de la capacidad para distinguir entre diferentes subgrupos, en este caso, entre la submuestra de intervención tras finalizar el programa y un grupo control.

El programa de intervención consistió en una unidad didáctica alternada para la práctica de AF *indoor* y *outdoor* (Viciana y Mayorga-Vega, 2016). La unidad didáctica constaba de cuatro sesiones *indoor* y cuatro *outdoor*, siguiendo la dinámica de impartir una sesión dentro del centro escolar (*indoor*) seguida de otra utilizando el entorno próximo (*outdoor*) durante todo el programa, con el objetivo de establecer una conexión en el aprendizaje de tareas para practicar AF en ambos espacios y dar a conocer el entorno próximo al alumnado. Los contenidos a trabajar fueron condición física y deportes. Las sesiones *indoor* se desarrollaron utilizando los materiales tradicionales del aula de EF (e.g., bancos suecos o balones medicinales) y espacios habituales dentro del centro educativo (gimnasio y pista polideportiva). Por su parte, las sesiones *outdoor* se desarrollaron aprovechando los espacios y recursos materiales que ofrecía el entorno próximo al centro educativo. En concreto, dos espacios contiguos al centro, un parque periurbano y el polideportivo municipal. Cada sesión duró 60 minutos y consistió en 10 minutos de calentamiento, 40-45 minutos de parte principal y 5 minutos de vuelta a la calma. La parte principal de las sesiones estuvo dividida en dos mitades de aproximadamente 20-25 minutos cada una. Durante la primera mitad, los estudiantes realizaron ejercicios de condición física (e.g., entrenamiento interválico, crossfit o relevos) y, durante la segunda mitad, trabajaron deportes de equipo (e.g., baloncesto, voleibol o ultimate).

Se utilizó el análisis de la varianza (ANOVA) de un factor (grupo control, grupo experimental) sobre los valores de la prueba CEPAF (global y dimensiones). El tamaño del efecto se estimó mediante la eta al cuadrado parcial (η^2_p) y la ∂ de Cohen (Field, 2017). Los valores del tamaño del efecto de $\partial < 0{,}20$ se consideran trivial, 0,20-0,49 pequeño, 0,50-0,79 moderado y $\geq$ 0,80 grande (Cohen, 1992). Los análisis estadísticos se realizaron utilizando el programa SPSS versión 23.0 para Mac (IBM® SPSS® Statistics). El nivel de significación estadística se estableció en $p < 0{,}05$.

Resultados

A continuación, se muestran los resultados obtenidos tras la aplicación de la prueba definitiva.

Análisis de las preguntas

En la Tabla 3 se observan los resultados obtenidos respecto a los índices referentes a las preguntas. Según los resultados obtenidos, la prueba cumplía con todos los criterios cuantitativos expuestos, habiéndose solventado en las cuatro preguntas que en la prueba piloto no se cumplieron. Además, la prueba definitiva presentaba una distribución adecuada de preguntas fáciles, moderadas y difíciles.

Tabla 3. Criterios cuantitativos de las preguntas incluidas en la prueba definitiva (*n* = 256)

	% preguntas (número)	Criterio correcto
Índice de dificultad mayor al 90%	0,0 (0 preguntas)	< 5%
Índice de dificultad menor al 10%	0,0 (0 preguntas)	< 5%
Índice de discriminación < 0,40	30,0 (9 preguntas)	> 25%
Índice de discriminación 0,21-0,39	50,0 (15 preguntas)	> 25%
Índice de discriminación 0-0,20	16,7 (5 preguntas)	> 15%
Índice de discriminación negativo	3,3 (1 pregunta)	< 5%
Alguna respuesta no escogida por al menos el 5% de participantes	0,0 (0 preguntas)	0%

Análisis de la prueba global

Respecto a los índices referentes a la prueba global, se calcularon: (a) La dificultad mediante la puntuación media del grupo en la prueba; (b) La variabilidad mediante la desviación estándar. Cuanto mayor es la desviación estándar, la prueba tiene mayor sensibilidad y discrimina mejor entre diferentes niveles de habilidad. La prueba completa obtuvo una puntuación media de 11,10 ± 3,46 puntos sobre 30. La puntuación en la dimensión de conocimiento declarativo fue de 3,58 ± 1,65 puntos sobre 10. En el conocimiento procedimental fue de 3,29 ± 1,69 puntos sobre 10. Por último, el conocimiento causal obtuvo 4,23 ± 1,86 puntos sobre 10.

Análisis de la fiabilidad

La Tabla 4 muestra los resultados de fiabilidad obtenidos, siendo moderada para la prueba global, y las dimensiones del conocimiento procedimental y causal, mientras que el declarativo obtuvo una fiabilidad pobre.

Tabla 4. Fiabilidad test-retest de la prueba definitiva (*n* = 81)

Variable	Puntuación 1 (DE)	Puntuación 2 (DE)	CCI (95% IC)
Global	10,83 (3,30)	11,40 (3,97)	0,72 (0,57-0,82)
Declarativo	3,40 (1,39)	3,63 (1,59)	0,26 (0,00-0,52)
Procedimental	3,51 (1,82)	3,78 (1,72)	0,51 (0,24-0,68)
Causal	3,93 (1,92)	3,99 (1,80)	0,53 (0,27-0,70)

Nota. DE = Desviación estándar; CCI = Coeficiente de correlación intraclase; 95% IC = 95% intervalo de confianza.

Análisis de la validez discriminante

La Tabla 5 muestra el efecto de la intervención sobre las puntuaciones en la prueba CEPAF. Los resultados del ANOVA de un factor mostraron que los estudiantes del grupo intervención obtuvieron una mejor puntuación de forma estadísticamente significativa respecto al grupo control tanto para la prueba global como para todas las dimensiones de forma independiente ($p < 0,001$).

TABLA 5. Efecto del programa sobre el conocimiento del entorno para la práctica de actividad física (*n* = 169)

Variable	Grupo	Post-intervención	ANOVA			TE
		M (DE)	F	p	η^2_p	∂
Global	Intervención (*n* = 88)	17,51 (3,97)	102,156	< 0,001	0,38	1,54
	Control (*n* = 81)	11,40 (3,90)				
Declarativo	Intervención (*n* = 88)	5,49 (1,72)	52,817	< 0,001	0,24	1,17
	Control (*n* = 81)	3,63 (1,59)				
Procedimental	Intervención (*n* = 88)	6,44 (1,58)	110,291	< 0,001	0,40	1,55
	Control (*n* = 81)	3,78 (1,72)				
Causal	Intervención (*n* = 88)	5,58 (1,93)	30,542	< 0,001	0,16	0,88
	Control (*n* = 81)	3,99 (1,80)				

TE = Tamaño del efecto; M = Media; DE = Desviación estándar; ∂ = Tamaño del efecto ∂ de Cohen.

Discusión

El objetivo del estudio fue construir y someter a un proceso de validación una prueba escrita *ad hoc* para crear una herramienta válida y fiable para la evaluación del conocimiento del entorno para la práctica de AF en estudiantes de ESO. La prueba CEPAF se ha conformado con una estructura de 30 preguntas dividida en tres dimensiones del conocimiento y se ha construido siguiendo los criterios sugeridos por la literatura previa (Baumgartner et al., 2015; Eignor, 2013). Los resultados demuestran que es un instrumento de medida válido, cumpliendo con los criterios de calidad cualitativos establecidos por los expertos y con los cuantitativos expuestos por Baumgartner et al. (2015). Además, los resultados demuestran que la prueba presenta una fiabilidad moderada, excepto para la dimensión de conocimiento declarativo cuya fiabilidad es pobre.

En línea con estudios previos sobre validación de pruebas escritas (e.g., Almonacid-Fierro, Feu y Vizuete-Carrizosa, 2018), las aportaciones de los jueces fueron aprovechadas para revisar las preguntas y conseguir una amplia batería de preguntas adecuadas, coherentes y claras en su redacción. La mayoría de correcciones estuvieron centradas en la redacción de las preguntas, modificaciones en las fotografías, sugerencia de variantes en las respuestas y perfeccionar las respuestas para asegurar que perteneciesen a esa dimensión del conocimiento. Por otro lado, gracias a la eliminación de las dificultades detectadas en la prueba piloto, el 100% de los participantes acabaron la prueba y el tiempo empleado estuvo dentro de los índices establecidos por la literatura científica (Argimón y Jiménez, 2000).

Por otro lado, una prueba donde las preguntas sean muy fáciles o muy difíciles, no será una prueba adecuada porque no será sensible al efecto de aprendizaje (Zhu et al., 2011). Por ello, el análisis de dificultad de las preguntas cobró una gran importancia en esta investigación. Consecuentemente, la dificultad global de la prueba debía ser moderada, conteniendo variedad de preguntas de diferentes niveles de dificultad que permita diferenciar entre niveles de habilidad (Zhu et al., 2011). De manera similar, estudios previos (e.g., Longmuir, Woodruff, Boyer, Lloyd y Trembaly, 2018) también han obtenido una distribución de preguntas fáciles, moderadas y difíciles en las pruebas de conocimiento desarrolladas. Respecto a la dificultad global de la prueba construida, ha sido adecuada, puesto que, sin realizar ningún programa de intervención, los estudiantes obtuvieron una media de

11,10 sobre 30 puntos (i.e., 37% de acierto) y cuando se lleva a cabo un programa de intervención, la media global de la prueba es apta (17,51 puntos sobre 30 preguntas; porcentaje de aciertos del 58,38%). Desde la perspectiva educativa, a pesar de ser el conocimiento del entorno para la práctica de AF es un objetivo requerido por el currículo educativo (Ministerio de Educación, Cultura y Deporte, 2015), resultan llamativos el bajo nivel obtenido por los estudiantes sin intervención. Este bajo conocimiento podría influir en la cantidad de práctica de AF que realizan los escolares en su tiempo libre y, por tanto, explicaría, al menos en parte, las altas tasas de inactividad física que presentan (Guthold et al., 2019).

En línea con lo anterior, dado que el porcentaje de aciertos medio post-intervención fue mayor en el grupo de intervención que en el grupo control (tamaño del efecto grande tanto en la prueba global, como en cada una de las dimensiones por separado), se confirmó la validez de la prueba para discriminar correctamente entre diferentes grupos de nivel. Por ello, se hace evidente la necesidad de incluir programas de formación específicos en estos contenidos para elevar el conocimiento de los escolares sobre su entorno para practicar AF. Aunque no se han encontrado estudios previos sobre el mismo cocimiento, en línea con los resultados del presente estudio, estudios previos encontraron que los escolares mejoraron el conocimiento (evaluado con un a prueba escrita) sobre diferentes contenidos del currículo de EF después de aplicar un programa de intervención (e.g. Chen, Zhu, Androzzic y Ho Nam, 2018; Ward et al., 2017).

Respecto a la fiabilidad test-retest, la prueba aporta resultados mayores para la fiabilidad de la prueba completa que para cada dimensión de forma independiente. Esto puede deberse a dos factores: (a) La fiabilidad se ve afectada por la longitud de la prueba y baja conforme disminuye el número de preguntas (Lacy y Williams, 2018). Por tanto, al disminuir la longitud de 30 preguntas del total a tan solo 10 preguntas de cada dimensión, la fiabilidad es más baja. Sin embargo, el número de preguntas que conforma cada dimensión no pudo ser mayor ya que se podría saturar a los alumnos con un mayor número de preguntas, afectando así a la correcta cumplimentación de la prueba (Argimón y Jiménez, 2000); y (b) La fiabilidad es más alta cuanto mayor es la dispersión de las puntuaciones y viceversa. Por consiguiente, la pobre fiabilidad obtenida en la dimensión de conocimiento declarativo puede deberse a una menor dispersión de sus puntuaciones (DE = 1,39; RIQ = 1), frente a las dimensiones de conocimiento procedimental y causal (DE = 1,82 y 1,92; RIQ = 3,00 y 3,25, respectivamente). Esta menor dispersión se observa también al comparar la dificultad de las preguntas entre dimensiones. Las puntuaciones máximas obtenidas en la dimensión de conocimiento declarativo únicamente llegan a seis puntos sobre 10, frente al conocimiento procedimental y causal que obtuvieron puntuaciones máximas de ocho puntos sobre 10. Realizando un análisis exhaustivo de la dificultad de cada una de las preguntas, se observa que cuatro de las 10 preguntas que componen el conocimiento declarativo, se encuentran entre las cinco preguntas con mayor índice de dificultad de la prueba completa. A pesar de que todas las preguntas cumplían con los criterios establecidos por Baumgartner et al. (2015) y ninguna presentaba índices de dificultad mayores al 90%, han podido resultar de una complejidad muy elevada para los estudiantes, acertándola pocos y reduciendo en última instancia la dispersión de las puntuaciones en la dimensión de conocimiento declarativo.

Fortalezas y limitaciones

Respecto a las fortalezas del estudio, a pesar de ser un aspecto demandado por todos los currículos educativos en España y a nivel internacional (e.g., Ministerio de Educación, Cultura y Deporte, 2015; Department for Education, 2013; SHAPE America, 2014), de lo que conocemos, es la primera prueba que se ha desarrollado para evaluar el conocimiento de los estudiantes sobre el entorno para realizar AF. Además, la construcción y validación de la prueba ha sido realizada siguiendo rigurosamente los estándares de construcción de pruebas educativas propuestos por la literatura científica (Baumgartner et al. 2015; Eignor, 2013). Además, a diferencia de estudios similares previos sobre la evaluación de conocimiento relacionado con la asignatura de EF (e.g., He et al., 2018), la prueba CEPAF tiene en cuenta la división de la prueba en los tres tipos de conocimiento (declarativo, procedimental y causal). Respecto a las limitaciones, debido a la naturaleza del estudio, la prueba está enfocada a un entorno concreto, lo que no permite generalizar los resultados a otros contextos. No obstante, estudios futuros podrían diseñar y evaluar una prueba de conocimiento con contenido similar, pero utilizando espacios y entornos generales. Es decir, no centrar las preguntas en entornos específicos del lugar sino en objetos o lugares que haya en todas las ciudades (e.g., cuestas, rampas o bancos sin especificar su localización).

Implicaciones prácticas

La validación llevada a cabo en el presente estudio representa una contribución importante para la literatura científica y con valiosas repercusiones prácticas para la asignatura de EF. Por un lado, dado el gran problema actual de inactividad física entre los estudiantes de ESO (Guthold et al., 2019), disponer de una herramienta válida para evaluar el conocimiento del entorno para la práctica de AF permite realizar estudios futuros que investigación su papel predictor sobre los niveles objetivos de AF de los escolares, así como mejorar la efectividad de diferentes programas de intervención para dicha promoción. Además, dado que la prueba está centrada en un entorno concreto, el presente estudio podría servir de preámbulo para desarrollar adaptaciones de la prueba para otros contextos concretos o incluso que pudiera aplicarse en todas las zonas geográficas. Por otro lado, como indica en el currículum de EF (Ministerio de Educación, Cultura y Deporte, 2015), es muy útil para comprobar si desde la asignatura de EF se está otorgando al alumno de las herramientas necesarias para ser activo de forma autónoma fuera del colegio.

Conclusiones

Los resultados del presente estudio demuestran que la prueba CEPAF es un instrumento de medida válido y fiable para obtener información sobre el conocimiento que poseen los escolares del entorno próximo para la práctica de actividad física. Por lo tanto, el estudio representa una contribución importante para la literatura científica y con valiosas repercusiones prácticas para la asignatura de Educación Física.

Agradecimientos

Los autores del presente estudio agradecen la colaboración de los estudiantes que han participado en el mismo, así como a los miembros responsables del centro donde se llevó a cabo el estudio, especialmente al director y profesor de educación física. Carolina Casado-Robles y Santiago Guijarro-Romero reciben una ayuda del Ministerio de Universidades de España (FPU16/03314 y FPU15/02387, respectivamente).

Referencias bibliográficas

Argimón, J., y Jiménez, J. (2000). *Métodos de investigación clínica y epidemiológica*. Madrid: Harcourt.

Baumgartner, T.A., Jackson, A.S., Mahar, M.T., & Rowe, D.A. (2015). *Measurement for evaluation in kinesiology*. Burlington, MA: Jones & Bartlett Publishers.

Chen, S., Chen, A., Sun, H., & Zhu, X. (2013). Physical activity and fitness knowledge learning in physical education: Seeking a common ground. *European Physical Education Review*, *19*(2), 256-270. doi: 10.1177/1356336X13486058

Chen, S., Zhu, X., & Kang, M. (2017). Development and validation of an energy-balance knowledge test for fourth- and fifth-grade students. *Journal of Sports Sciences, 35*(10), 1004-1011. doi: 10.1080/02640414.2016.1208837

Chen, S., Zhu, X., Androzzi, J., & Nam, Y.H. (2018). Evaluation of a concept-based physical education unit for energy balance education. *Journal of Sport and Health Science*, *7*(3), 353-362. doi: 10.1016/j.jshs.2016.06.011

Cohen, J. (1992). A power primer. *Psychological bulletin, 112*(1), 155.

European Commission/EACEA/Eurydice. (2013). *Physical Education and sport at school in Europe Eurydice Report*. Luxembourg: Publications Office of the European Union.

Corral, Y. (2009). Validez y confiabilidad de los instrumentos para la recolección de datos. *Revista ciencias de la educación, 33*, 228-247.

Department for Education. (2013). *National curriculum in England: physical education programmes of study.* Disponible en: https://www.gov.uk/government/publications/national-curriculum-in-england-physical-education-programmes-of-study [21 Noviembre 2019].

Eignor, D.R. (2013). *The standards for educational and psychological testing*. Washington, DC: American Psychological Association.

Ennis, C. (2011). Physical education curriculum priorities: Evidence for education and skillfulness. *Quest 63*, 5–18. doi: 10.1080/00336297.2011.10483659

Escobar-Pérez, J., & Cuervo-Martínez, Á. (2008). Validez de contenido y juicio de expertos: una aproximación a su utilización. *Avances en medición, 6*(1), 27-36.

Field, A. (2017). *Discovering statistics using IBM SPSS Statistics (5th ed.)*. London: SAGE Publications.

Almonacid-Fierro, A., Feu, S., y Vizuete-Carrizosa, M. (2018). Validación de un cuestionario para medir el Conocimiento Didáctico del Contenido en el profesorado de Educación Física. *Retos, 34*, 132-137

Fontes, G.S., García-Gallego, C., Garriga-Trillo, A.J., Pérez-Llantada, M.C., y Sarriá, E. (2007). *Diseños de investigación en psicología*. Madrid: UNED

Gomez, R., Vance, A., & Stavropoulos, V. (2018). Test-retest measurement invariance of clinic referred children's ADHD symptoms. *Journal of Psychopathology and Behavioral Assessment, 40*(2), 194–205. doi: 10.1007/s10862-017-9636-4

Guthold, R., Stevens, G.A., Riley, L.M., & Bull, F.C. (2019). Global trends in insufficient physical activity among adolescents: a pooled analysis of 298 population-based surveys with 1.6 million participants. *The Lancet Child & Adolescent Health*. doi: 10.1016/S2352-4642(19)30323-2

He, Y., Ward, P., & Wang, X. (2018). Validation of a common content knowledge test for soccer. *Journal of Teaching in Physical Education, 37*(4), 407-412. doi: 10.1123/jtpe.2017-0204

Koo, T.K., & Li, M.Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of chiropractic medicine, 15*(2), 155-163. doi: 10.1016/j.jcm.2016.02.012

Lacy, A.C., & Williams, S.M. (2018). *Measurement and evaluation in physical education and exercise science*. London: Routledge.

Longmuir, P.E., Woodruff, S.J., Boyer, C., Lloyd, M., & Tremblay, M.S. (2018). Physical Literacy Knowledge Questionnaire: feasibility, validity, and reliability for Canadian children aged 8 to 12 years. *BMC Public Health, 18*(2), 1035

Miklos, T. (1999). *Educación y capacitación basada en competencias*. Mexico: Limusa.

Ministerio de Educación, Cultura y Deporte. (2013). Ley Orgánica 2/2006, de 3 de mayo, de Educación. *Boletín Oficial del Estado, 106,* 17158-17207.

Ministerio de Educación, Cultura y Deporte. (2015). Real Decreto 1105/2014, de 26 de diciembre, por el que se establece el currículo básico de la Educación Secundaria Obligatoria y del Bachillerato. *Boletín Oficial del Estado, 3,* 169–256.

Raymond, M.R., Stevens, C., & Bucak, S.D. (2019). The optimal number of options for multiple-choice questions on high-stakes tests: application of a revised index for detecting nonfunctional distractors. *Advances in Health Sciences Education*, *24*(1), 141-150. doi: 10.1007/s10459-018-9855-9

Shape America (2014). *National standards & grade-level outcomes for K-12 physical education.* Leeds: Human Kinetics.

Viciana, J., & Mayorga-Vega, D. (2016). Innovative teaching units applied to physical education–changing the curriculum management for authentic outcomes. *Kinesiology, 48*(1), 142-152.

Viciana, J., & Mayorga-Vega, D. (2018). The three-axes model of planning in physical education. *Retos, 33,* 313-319.

Walter, S. D., Eliasziw, M., & Donner, A. (1998). Sample size and optimal designs for reliability studies. *Statistics in medicine, 17*(1), 101-110.

Ward, J.K., Hastie, P.A., Wadsworth, D.D., Foote, S., Brock, S.J., & Hollett, N. (2017). A Sport education fitness season's impact on students' fitness levels, knowledge, and in-class physical activity. *Research quarterly for exercise and sport, 88*(3), 346-351. doi: 10.1080/02701367.2017.1321100

World Health Organization. (2010). *Global recommendation on physical activity for health*. Geneva: WHO

World Health Organization. (2014). *Global Status Report on Noncommunicable Diseases 2014.* Geneva: WHO.

World Health Organization. (2018). *Global Action Plan on Physical Activity 2018–2030: More Active People for a Healthier World.* Geneva: WHO.

Zack, M.H. (1999). Developing a knowledge strategy. *California management review*, *41*(3), 125-145.

Zhu, W., Rink, J., Placek, J.H., Graber, K.C., Fox, C., Fisette, J.L., ... & Raynes, D. (2011). PE Metrics: Background, testing theory, and methods. *Measurement in Physical Education and Exercise Science*, *15*(2), 87-99.

Anexo 1. Instrumento definitivo aplicado

PRUEBA SOBRE EL CONOCIMIENTO DE TU ENTORNO PARA LA PRÁCTICA DE ACTIVIDAD FÍSICA

*En la siguiente prueba se va a evaluar el conocimiento sobre tu entorno para la práctica de actividad física. Lee tranquilamente la pregunta y todas las 4 posibles respuestas antes de elegir una respuesta.

*Para cada pregunta hay 4 posibles respuestas y sólo una de ellas es válida. Por lo que hay que elegir solo una de las opciones.

* Por favor, para cada pregunta indica EN LA PLANILLA DE RESPUESTAS la opción que has elegido, marcando claramente con una "X" la letra **"A, B, C o D".** En caso de que quieras cambiarla, táchala completamente y rodea con un círculo la letra que consideres válida. **NO ESCRIBAS EN LAS HOJAS DE LAS PREGUNTAS, SÓLO EN LA PLANILLA.**

*No penalizan los errores, por lo que siempre tienes que elegir una de las respuestas (la que se piense que es correcta).

1. **¿Qué músculos podrías trabajar en cualquier escalera como éstas de la plaza del pueblo?**

a) Bíceps y pectorales
b) Cuádriceps y pectorales
c) Dorsales y bíceps
d) Gemelos y dorsales

2. **En un campo de fútbol 7, además de poder jugar a fútbol, ¿qué otras actividades deportivas puedes practicar?**
a) Béisbol, ultimate y balonmano
b) Acrosport, hockey y tenis
c) Tenis, rugby y baloncesto
d) Pádel, béisbol y hockey

3. **¿Cómo podrías trabajar la fuerza de los gemelos en este espacio que hay frente al instituto?**

a) De pie en el suelo, con el peso en las punteras de los pies, y con las manos en el macetero para equilibrarnos, realizaríamos elevaciones y bajadas de talones.
b) Encima del escalón con los dos pies y con sólo el peso en las punteras de los pies, realizaríamos elevaciones y bajadas de talones todo lo que podamos sin caernos y sin llegar a tocar el suelo con los talones.
c) Realizando repetidos y rápidos saltos pequeños al escalón con los dos pies simultáneamente.
d) Todas las respuestas anteriores son correctas

4. **¿Cómo podrías aprovechar los árboles del parque de la Paz para hacer deporte?**
a) Colocando una cinta o cuerda de un árbol a otro para practicar salto de altura
b) Utilizándolos como elementos de trepa para hacer competiciones por equipos, donde el primero que llegue arriba del todo gana
c) Colocando una cinta o cuerda de un árbol a otro para jugar a voleibol
d) Todas las respuestas son correctas

5. **¿Se puede trabajar la condición física (fuerza, resistencia y flexibilidad) en los parques y zonas públicas del pueblo, igual que podrías hacer en el gimnasio?**

   a) No, no puedo trabajar la condición física en los parques y zonas públicas del pueblo porque necesitaría máquinas específicas (banco sueco, espalderas, etc.) y materiales (balón medicinal, etc.) que sólo están en el gimnasio.
   b) Sí, podría trabajar la condición física en los parques y zonas públicas del pueblo, pero sólo cuando haya alcanzado un nivel avanzado de condición física porque todos los ejercicios que pueda realizar al aire libre son más exigentes.
   c) Sí, podría trabajar la condición física en los parques y zonas públicas porque hay elementos urbanos que se parecen a los elementos del gimnasio y podemos hacer los mismos ejercicios en ellos.
   d) Sí, podría trabajar la resistencia aeróbica corriendo por los parques y zonas públicas del pueblo, pero no podría trabajar la fuerza y la resistencia porque no hay máquinas ni materiales específicos para ello.

6. **¿Podríamos practicar salto de altura utilizando sólo los elementos existentes en el parque de la Paz?**

   a) No, porque nos faltarían medidas de seguridad para realizar la caída después del salto.
   b) Sí, porque podemos colocar una cinta de un árbol a otro para marcar la distancia a la que hay que saltar.
   c) No, porque no tenemos espacio suficiente para realizar la carrera de aproximación antes de saltar.
   d) Sí, porque tenemos espacio suficiente para realizar la carrera de aproximación antes de saltar.

7. **Para evitar que nos lesionemos la espalda, ¿cuál es la mejor forma para estirar los músculos isquiosurales (músculo de la pierna, situado detrás del muslo)?**

a) 
b) 
c) 

d) Todas las respuestas anteriores son correctas, todos los ejercicios están aconsejados para su realización.

8. **¿Dónde puedo practicar el frontón (o pelota-mano que consiste en golpear una bola frente a la pared) en el pueblo de Churriana de la Vega?**

   a) En la zona de la pista de baloncesto del polideportivo municipal.
   b) En la valla exterior del instituto, junto a la puerta de entrada.
   c) En el parque de la Paz.
   d) En la pista de tenis del polideportivo municipal.

9. En la siguiente fotografía, vemos las barras que rodean el campo de fútbol del polideportivo municipal, que tiene una altura aproximada de 70 centímetros. ¿Podría trabajar los músculos del tren superior (parte superior de nuestro cuerpo, desde la cintura hasta arriba) con esta barra?

a) Sí, podría usarla para trabajar los pectorales y abdominales.
b) Sí, podría usarla para trabajar los dorsales y abdominales.
c) No, tiene una altura demasiado baja para trabajar el tren superior, necesitaría una barra alta de aproximadamente 2 metros.
d) Sí, podría usarla para trabajar los pectorales y dorsales.

10. ¿Cómo podrías jugar a béisbol (deporte competitivo donde se enfrentan dos equipos y se golpea la bola con un bate) en el pueblo de Churriana de la Vega?

a) En el parque de la Paz, utilizando la plaza redonda con maceteros alrededor como terreno de juego y que los maceteros sean las bases por donde haya que pasar corriendo para conseguir una carrera.
b) En el campo de fútbol del polideportivo municipal, utilizando el área de la portería como zona de golpeo y marcando las bases con cualquier objeto (por ejemplo, una mochila) en las líneas de banda del campo.
c) En el parque de la Paz, utilizando el espacio amplio del final de hormigón cómo pista y los elementos que hay alrededor (por ejemplo, los árboles y las farolas) como bases por donde haya que pasar corriendo para conseguir una carrera.
d) Todas las respuestas anteriores son correctas.

11. ¿Podemos utilizar los siguientes elementos del pueblo (banco y muro del parque) para trabajar los músculos de las piernas realizando saltos?

1) 

2) 

a) Sí, podemos utilizar los dos elementos porque tienen aproximadamente la misma altitud sobre la que podemos saltar con facilidad.
b) Sólo podríamos utilizar el banco (fotografía 1) porque ofrece una altitud adecuada y una cara plana sobre la que poder saltar de forma segura.
c) No, porque los saltos deberían realizarse con elementos muy seguros, específicamente diseñados para ello, como los cajones de gimnasio, y estos dos elementos no son seguros para su práctica.
d) Sólo podríamos utilizar el muro (fotografía 2) porque es de hormigón y sabemos que está fijo al suelo y es seguro para saltar

12. A continuación, observamos dos fotografías del pueblo de Churriana de la Vega, la primera (1) es la plaza del "Maestro Antonio Agustín", y la segunda (2) es el polideportivo municipal. ¿En cuál de estos recintos puedes jugar al baloncesto?

1)  2) 

a) Sólo en el polideportivo municipal, porque necesitamos una pista oficial con canastas para poder jugar.
b) Sólo en la plaza del pueblo, porque la pista de baloncesto del polideportivo municipal está reservada específicamente para los que están apuntados a la actividad en el ayuntamiento.
c) En ambos lugares se puede jugar, porque en el polideportivo tenemos una pista oficial para hacerlo y en la plaza del pueblo podemos formar canastas y delimitar el campo con los elementos presentes.
d) En ninguno de los dos lugares, porque el área de deportes del pueblo no ofrece el baloncesto como actividad deportiva y en la plaza del pueblo hay canastas.

13. ¿Podrías trabajar la fuerza del tren superior (parte superior de nuestro cuerpo, desde la cintura hasta arriba) utilizando estos maceteros que encontramos en las calles del pueblo?

a) Sí, utilizándolos como apoyo para las manos para hacer el pino.
b) Sí, realizando saltos sobre el banco con los pies juntos.
c) No, es un elemento urbano inadecuado para realizar ejercicio físico en él.
d) Sí, utilizándolos como apoyo para las manos para hacer flexiones de pecho y fondos de tríceps.

14. ¿Existe un carril bici que una tu localidad con Granada capital?

a) No, sólo hay posibilidad de ir por carretera.
b) Sí, existe un carril bici bien señalizado que discurre principalmente al lado de la carretera y con protecciones de barrera en los puntos más peligrosos
c) Sí, existe un carril bici pero que desaparece en algunos tramos y hay que continuar por la carretera hasta el siguiente tramo de carril bici.
d) Sí, existe un carril bici pero está en malas condiciones, por lo que es más seguro realizar el recorrido por la carretera

15. Estás realizando encogimientos en el suelo para trabajar los músculos abdominales, pero ya te resulta muy fácil y quieres incrementar el nivel de dificultad. ¿Cómo podrías aumentar el nivel de dificultad realizando el mismo ejercicio utilizando esta rampa?

a) Realizando el mismo ejercicio, pero tumbándonos con la cabeza en la parte más baja de la rampa y las piernas en la parte más alta.
b) Realizando el mismo ejercicio, pero tumbándonos con la cabeza en la parte más alta de la rampa y las piernas en la parte más baja.
c) Cambiando la posición de las piernas y colocarlas extendidas en lugar de flexionadas y tumbándonos con la cabeza en la parte más alta de la rampa y las piernas en la parte más baja.
d) Cambiando la posición de las piernas y colocándolas encima de la barandilla con el cuerpo en la rampa.

16. ¿Cómo podrías jugar a balonmano en el pueblo de Churriana de la Vega?

a) En el parque de la Paz, utilizando la zona de arena con árboles finos como terreno de juego y que el hueco entre dos árboles sean las porterías
b) En el parque de la Paz, utilizando el espacio amplio del final de hormigón cómo pista y que el hueco de muro para entrar al jardín sea una portería y en el lado contrario la portería sea el espacio que ocupan los tres bloques de hormigón
c) En el polideportivo municipal, en la pista polideportiva con las reglas oficiales.
d) Todas las respuestas anteriores son correctas.

17. ¿Dónde y por qué es más aconsejable realizar el siguiente ejercicio para trabajar los tríceps?

a) En el banco de la calle es más aconsejable porque lo estamos haciendo al aire libre.
b) Ambos son igual de aconsejables porque se puede realizar el mismo ejercicio en ellos.
c) En el banco de gimnasio es más aconsejable porque está específicamente diseñado para realizar ese tipo de ejercicios.
d) Ninguno de ellos es aconsejable, porque este ejercicio es perjudicial para nuestra espalda

18. **¿Podríamos jugar a algún juego tradicional con los amigos, aprovechando el espacio de la plaza Maestro Antonio Agustín?**

a) Sí, por ejemplo, podríamos jugar al "Balón prisionero", porque tan sólo necesitamos delimitar el campo y se puede hacer con las marcas de las baldosas del suelo o con los pilares de la izquierda.
b) Sí, podríamos jugar a "Atrapa la Bandera", porque toda la plaza sería la zona de juego y la delimitación del espacio de cada equipo y el de la zona segura de la bandera se puede hacer con referencias como las farolas, palmeras o marcas del suelo.
c) Sí, podríamos practicar "Judo" porque no necesitamos nada de material y es un deporte fácil y seguro para poder practicarlo en cualquier lugar de la calle.
d) Las respuestas "a y b" son correctas.

19. **¿En cuál de los siguientes lugares puedo estirar los gemelos de forma correcta?**

a) 
b) 
c) 

d) En todos los lugares de las fotografías anteriores se podrían estirar los gemelos.

20. **¿Qué actividad deportiva no ofrece la concejalía de deportes en Churriana de la Vega?**

a) Ciclismo
b) Lucha olímpica
c) Bádminton
d) Tenis

21. **Estás realizando flexiones de brazos con los pies en el suelo, pero ya te resulta muy fácil y quieres incrementar el nivel de intensidad del ejercicio. ¿Podrías utilizar alguno de los elementos de la siguiente fotografía para realizar un ejercicio de mayor dificultad para trabajar los músculos pectorales?**

a) Sí, podría aumentar la dificultad usando el muro para apoyar las piernas y hacer flexiones de brazos.
b) Sí, podría aumentar la dificultad usando el muro para apoyar las manos y hacer flexiones de brazos.
c) Sí, podría aumentar la dificultad usando la rama de los árboles para colgarme y hacer dominadas.
d) Las respuestas "a y b" son correctas para aumentar la dificultad en el ejercicio de flexiones de brazos.

22. ¿Cómo podría jugar al ultimate (juego con un frisbee donde se enfrentan dos equipos y para conseguir punto hay que capturar el frisbee en una zona específica) en esta zona del pueblo (Plaza Maestro Antonio Agustín)?

a) Podría usar las baldosas del suelo para delimitar las zonas de puntuación y el campo central.
b) Podría usar los pilares como zonas de puntería donde habría que dar con el frisbee en la zona más alta para conseguir punto.
c) Podría usar los espacios entre dos árboles o dos bancos que hay repartidos en la plaza como porterías donde hay que colar el frisbee para conseguir punto.
d) Ninguna respuesta anterior es correcta, no podemos jugar al ultimate en esta plaza del pueblo.

23. En las siguientes fotografías se están realizando ejercicios para fortalecer el tren superior (parte superior de nuestro cuerpo, desde la cintura hasta arriba). ¿En cuál de ellas se están trabajando los músculos dorsales de la de la espalda?

1)  2) 

a) Los músculos dorsales se están trabajando en el ejercicio de la segunda fotografía (2) porque para trabajar los músculos dorsales necesitamos realizar un empuje contra el suelo como si quisiéramos hundirlo.
b) En ambos ejercicios se están trabajando los músculos dorsales porque se pueden trabajar tanto en acciones de empujar (ejercicio de la fotografía 2) como en acciones de tirar (ejercicio de la fotografía 1)
c) Los músculos dorsales se están trabajando en el ejercicio de la primera fotografía (1) porque para trabajar los músculos dorsales necesitamos tirar de la barra hacia nosotros como si quisiéramos acercárnosla para subir.
d) En ningún ejercicio se están trabajando los músculos dorsales, porque necesitaríamos una barra más alta donde colgarnos completamente de la barra sin tocar el suelo para poder trabajar los dorsales.

24. ¿Podrías practicar diferentes modalidades de atletismo en los espacios urbanos del pueblo?

a) Sí, por ejemplo, podría realizar carreras de velocidad porque en el parque de la Paz hay un espacio recto y seguro de más de 100 metros.
b) Sí, podría trabajar carreras de medio fondo, por ejemplo, 1.500 metros, porque es la distancia que hay corriendo al lado del carril bici desde el Polideportivo Municipal hasta el río Dílar.
c) Las respuestas "a y b" son correctas, porque hay suficiente distancia y un espacio seguro para hacerlo.
d) Las respuestas "a y b" son incorrectas, para trabajar estas modalidades de atletismo tendría que hacerlo en una pista específica que mida las distancias exactas.

25. **¿Cuáles de los siguientes ejercicios podrías hacer en el espacio urbano del pueblo para trabajar los músculos abdominales?**

a)  b)  c)  d) 

26. **Relaciona los siguientes espacios del pueblo con la actividad deportiva que puedes hacer en ellos:**

a) En el parque de la Paz puedo jugar a la petanca y en el parque de la Ermita puedo realizar ejercicios con las máquinas biosaludables.
b) Al lado del instituto puedo realizar ejercicios con las máquinas biosaludables y en el parque de la Ermita puedo hacer escalada en rocódromo.
c) En la plaza del Ayuntamiento puedo realizar ejercicios con las máquinas biosaludables, y en el Polideportivo municipal puedo hacer escalada en rocódromo.
d) En el parque de la Paz puedo jugar a la petanca y en el Polideportivo municipal puedo realizar ejercicios con las máquinas biosaludables.

27. **¿Cómo podrías usar el tronco de un árbol para estirar los gemelos?**

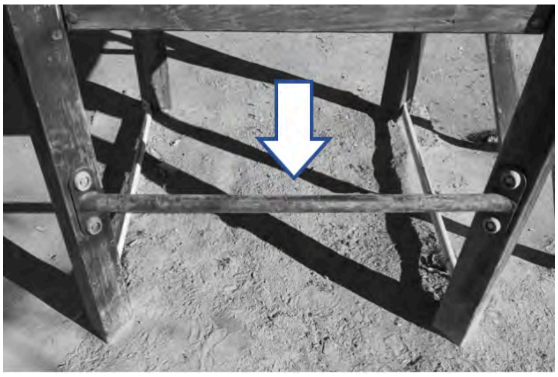
a)  b)  c) 

d) Todas las respuestas son correctas.

28. **¿Se puede practicar baloncesto en el parque de la Paz?**

a) No, no se puede practicar porque no hay canastas ni pista polideportiva.
b) Sí, podemos jugar con las reglas oficiales porque dispone de canastas.
c) Sí, podemos jugar, aunque con reglas adaptadas, por ejemplo, utilizando un aro atado a un árbol como canasta o dando con el balón en un sitio concreto usado como diana.
d) Ninguna respuesta anterior es correcta

29. En la siguiente fotografía, vemos la barra baja de los columpios del parque de la Paz, que tiene una altura aproximada de 20-30 centímetros. ¿Podrías trabajar los músculos pectorales con esta barra?



a) No, porque el cuerpo estaría demasiado horizontal con respecto al suelo y así trabajaríamos los músculos dorsales.
b) Sí, porque puedo tumbarme debajo y subir el pecho hasta que toque la barra aguantando en esa posición.
c) No, necesitaríamos una barra más alta porque necesitamos colgarnos completamente de la barra sin tocar el suelo para poder hacer el ejercicio.
d) Sí, porque puedo usarlas como apoyo para las manos al realizar un ejercicio de flexiones de pecho.

30. ¿Podríamos jugar a voleibol en algún lugar del pueblo de Churriana de la Vega?

a) No, porque el área de deportes del pueblo no ofrece el voleibol como actividad deportiva.
b) Sí, por ejemplo, en el parque de la Paz o la plaza Maestro Antonio Agustín porque con los elementos presentes en estos lugares, podemos colocar una cinta creando una red y delimitar el campo.
c) Sí, por ejemplo, el polideportivo municipal o en el carril bici que hay al lado del polideportivo porque podemos formar una red y delimitar el campo con cualquier elemento que tengamos.
d) Sí, pero sólo podríamos jugar en el instituto y en el polideportivo municipal, porque en los demás espacios urbanos del pueblo no hay redes y campos delimitados para jugar a voleibol.

ADAPTATION OF THE SPORT SATISFACTION INSTRUMENT TO THE PHYSICAL EDUCATION LESSON IN SPANISH SECONDARY STUDENTS: PSYCHOMETRIC PROPERTIES AND FACTOR INVARIANCE

Casado-Robles, C., Viciana, J., Guijarro-Romero, S., & Mayorga-Vega, D.

Draft

Paper IX

ABSTRACT

Objective: The present study aimed to adapt and analyze the psychometric properties of the Sport Satisfaction Instrument adapted to Physical Education (PE) in Spanish Secondary students to assess their enjoyment towards: (a) PE subject in general (i.e., during several lessons; SSI-PE-subject); and (b) only the PE class that just finished (i.e., one particular lesson; SSI-PE-lesson).

Methods: A sample of 402 Spanish Secondary students (200 females), aged 11-17 years old, participated in the first phase of the study by filling the Spanish version of the previously validated SSI-PE-subject. For the second phase, a subsample of 304 students (153 females) completed the adapted SSI-PE-lesson. Firstly, the SSI-PE was adapted to measure the students' intrinsic enjoyment at the end of one specific PE lesson, assessing the content validity of the items by a panel of experts. Subsequently, confirmatory factor analyses and the internal consistency analysis of the SSI-PE-subject and SSI-PE-lesson questionnaires were performed. Finally, factorial invariance analyses by gender were conducted.

Results: Results showed that the model with four items (i.e., items one, five, six and eight of the original version) and a single factor (i.e., enjoyment) obtains the best indices of fit for both versions (SSI-PE-subject: CFI = 1.000; RMSEA < 0.001; GFI = 0.999; SSI-PE-lesson: CFI = 1.000; RMSEA < 0.001; GFI = 0.997). The internal consistency was excellent in both models (Ω = 0.920). Factorial invariance was also verified across gender.

Conclusions: In conclusion, the proposed four-item and single-factor model of the SSI-PE-subject and SSI-PE-lesson questionnaires show adequate psychometric properties among Spanish students.

KEYWORDS

Enjoyment. Content validity. Factor analysis. Reliability. Adolescents.

Introduction

The enjoyment of physical activity (PA) is a relevant regulatory process driving students' intrinsic motivation (Ryan & Deci, 2020). Moreover, in the school setting, without students' motivation learning is not possible (Rehman & Haider, 2013), so pupils' enjoyment in carrying out educational tasks is crucial in learning achievement. Furthermore, enjoyment is considered a positive predictor of students' intention to engage in PA, as well as to reduce sedentary behavior (Bai et al., 2018; García et al., 2016; Pulido et al., 2014; Sánchez-Oliva et al., 2014). For these reasons, the students' enjoyment should be considered a key factor when designing intervention programs to promote healthy PA habits. Specifically, Physical Education (PE) subject has been considered an important setting to carry out programs for promoting students' acquisition of lifelong healthy PA habits (World Health Organization, 2018). Therefore, teachers should consider the relationship between the enjoyment of PE and students' PA behaviors when planning and implementing PE programs to ensure that PE lessons are enjoyable for them (Bai et al., 2018; Barr-Anderson et al., 2007). Consequently, it is essential to measure students' enjoyment in PE lessons using valid and reliable scales.

In this sense, several studies have been focused on the adaptation and validation of "The Intrinsic Satisfaction Classroom scale" from the academic to the PE context. This scale originally emerged to measure students' satisfaction with school learning, asking them how satisfied versus bored or frustrated are they with learning in school, using an 11-item scale divided into two subscales: satisfaction/enjoyment and boredom (Nicholls et al., 1985, 1989). Later, Duda & Nicholls (1992) developed the Sport Satisfaction Instrument (SSI), carrying out a scale adaptation to the sport context. It was an eight-item scale divided into the same two subscales: satisfaction/enjoyment (five items) and boredom (three items). After that, Balaguer et al. (1997) translated the SSI into the Spanish context with adolescents, supporting the same structure as Duda & Nicholls (1992). Some years later, this scale was also studied in Spain by Cervelló et al. (1999), Castillo et al. (2002) and Álvarez et al. (2009) with adolescents, and Granero-Gallegos et al. (2014) with female athletes. However, these four studies proposed a 7-item scale (five items for satisfaction/enjoyment and only two items for boredom) by eliminating the second (i.e., "I often daydream instead of thinking about what I am doing"; Álvarez et al., 2009; Castillo et al., 2002; Granero-Gallegos et al., 2014) or the third item (i.e., "When playing sports, I am usually bored"; Cervelló et al., 1999) from the scale to improve boredom dimension reliability up to acceptable values and guarantee convergent validity of the model. Besides, Ruiz-Juan et al. (2010) also proposed a 7-item scale (four items for satisfaction/fun and three items for boredom) to measure intrinsic satisfaction in young elite paddlers, but in this case, the deleted item was number six (i.e., "When I do sport I usually find time flies") belonging to the enjoyment dimension to improve its reliability.

Finally, focusing on the PE setting, Baena-Extremera et al. (2012) adapted the SSI questionnaire by Balaguer et al. (1997) to measuring intrinsic satisfaction toward PE lessons in Spanish adolescents, and later Baena-Extremera & Granero-Gallegos (2015) also adapted that scale to the bilingual PE context. Those studies proposed both a 7-item model like Castillo et al. (2002) and the original 8-item model, obtaining slightly better results for the 8-item model but suggesting taken the 7-item model future research in which the model may achieve a good fit. Therefore, due to the fact that previous studies focused on the adaptation and validation of the scale removed some items (e.g., two or six) to improve its results (e.g., Castillo et al., 2002; or Ruiz-Juan et al., 2010), future studies should test different structures in order to achieve the best configuration depending on the context to be applied.

Throughout the several validations carried out about the scale in different contexts, the construct to be measured has been defined as the degree of intrinsic satisfaction, interest, enjoyment, or fun that a person experience in a certain activity, using all these terms interchangeably as synonyms (e.g., Baena-Extremera et al., 2012; Castillo et al., 2002). Similarly, empirical studies that apply these scales also use these terms interchangeably (e.g., Baena-Extremera & Granero-Gallegos, 2013; Baños et al., 2019). Besides that, these definitions also include the term boredom as the opposite feeling of enjoyment or fun while doing an activity, that is, students get bored practicing PA when they do not have fun (Baena-Extremera et al., 2012; Castillo et al., 2002; Duda & Nicholls, 1992). Although there is no agreed-upon definition of boredom (Vodanovich, 2003), this relationship between enjoyment and boredom is commonly established in the thesaurus and scientific research as two antagonistic poles of the same continuum, commonly defining boredom as a feeling of unpleasant, dissatisfaction, disinterest, or lack of enjoyment (Hamilton et al., 1984;

O'Brien, 2014; Vogel-Walcutt et al., 2012). Therefore, the psychometric characteristics of the SSI adapted to PE (SSI-PE-subject) could be studied using a single dimension that encompasses both satisfaction and boredom subscales, understanding them as opposite poles of the same concept (Cervelló et al., 1999).

Furthermore, these previously validated scales both in the sport and the school contexts were referred to an extended period of time, that is, to the students' satisfaction in general during a particular period of time involved in PA practice or PE lessons. However, none of them were focused on evaluating the "acute effect" of one specific PE lesson that students just completed. Therefore, adapting the scale to a specific PE lesson is pending, and it could help teachers check what kind of activities, teaching methodologies, or contents are more enjoyable for students to improve their motivation. Consequently, the present study aimed to: (a) adapt the Spanish version of SSI-PE-subject only focused on the PE lesson they just finished (i.e., SSI-PE-lesson); (b) evaluate the content validity of the items constituting the questionnaires with experts judges; (c) test several structures of the adapted SSI-PE-lesson and the original SSI-PE-subject through confirmatory factor analyses and the analysis of its internal consistency in a sample of Secondary students (aged 11-17 years old); and (d) verify the factorial invariance of both versions of the scale by gender.

Methods

Participants

A total of 410 students (204 females and 206 males), aged 11-17 years old, and belonging to Secondary Education level, were invited to participate in the present study. To assess the adapted version of the questionnaire (i.e., SSI-PE-lesson), a subsample of 304 students (153 females and 151 males) was used. The following inclusion criteria were considered: (a) participating in the normal PE classes; (c) being free of any disability that would make them unable to fill the questionnaire; d) presenting the corresponding signed written consent by their legal guardians, and e) presenting the corresponding signed written assent by the adolescents. The exclusion criterion was not filling out all the items of the questionnaires. Table 1 shows the general characteristics of the included participants.

Table 1. General characteristics of the sample

	Total (N = 402)		Subsample (N = 304)	
	Females (n = 200)	Males (n = 202)	Females (n = 153)	Males (n = 151)
Age (years)	13.60 (1.03)	13.83 (1.08)	13.46 (1.06)	13.76 (1.14)
Body mass (kg)	55.28 (11.45)	58.70 (13.76)	54.76 (11.26)	57.97 (13.49)
Body height (cm)	159.51 (6.31)	166.11 (9.50)	159.66 (6.31)	165.45 (9.94)
Body mass index (kg/m^2)	21.65 (3.92)	21.11 (3.89)	21.42 (3.93)	21.02 (3.90)

Note. Data are reported as mean (standard deviation), except for gender that percentage is reported instead.

Measures

Enjoyment and boredom in Physical Education subject

The Spanish version of the SSI-PE-subject developed by Baena-Extremera et al. (2012) was used to assess the adolescents' intrinsic enjoyment and boredom towards the PE subject. This questionnaire was originally composed of eight items belonging to two dimensions related to enjoyment (five items; e.g., "I usually have fun in PE classes") and to boredom (three items; e.g., "In PE classes, I usually get bored"). Following the stem "Indicate your degree of disagreement or agreement with the following statements, referring to your PE classes…", in the original version, participants respond to each item on a five-point Likert-type scale, ranging from 1 ("Strongly disagree") to 5 ("Strongly agree"). However, in the present study, in order to facilitate the evaluation of the items by the adolescents a 10-point Likert-type scale was used, ranging from 1 ("Strongly disagree") to 10 ("Strongly agree"), since is the scale that Spanish school-aged children receive in their scholar marks (e.g., Casado-Robles et al., 2020).

Enjoyment and boredom in the Physical Education lesson

The above-mentioned Spanish version of the SSI-PE-subject (Baena-Extremera et al., 2012) was adapted to measure the students' intrinsic enjoyment and boredom toward a specific PE lesson (i.e., SSI-PE-lesson). A 10-point Likert-type scale, ranging from 1 ("Strongly disagree") to 10 ("Strongly agree") was also used.

Procedure

Adaptation process

Before carrying out the data collection, the Spanish version of the SSI-PE-subject (Baena-Extremera et al., 2012) was adapted to measure the students' intrinsic enjoyment and boredom towards one specific PE lesson (i.e., SSI-PE-lesson). The original questionnaire was referred to the PE subject in general. Therefore, in this version, the items were slightly modified to refer to the class that had just finished. Firstly, the main researcher changed the items from the original SSI-PE-subject. After that, the list of these modified items was given to two experienced researchers who were experts in the PE area and were familiar with the construct being evaluated to assess the comprehension of the items.

Four items were modified regarding experts' suggestions: (a) the verbal tense was modified in two items (i.e., item 2: "*he soñado*" was changed for "*soñaba*"; and item 4: "*he deseado*" was changed for "*deseaba*"); and (b) the order of the sentence was modified in two items (i.e., item 5: "*La clase de Educación Física de hoy la he encontrado interesante*" was changed for "*He encontrado la clase de Educación Física de hoy interesante*"; and item 7: "*En la clase de Educación Física de hoy he participado activamente*" was changed for "*He participado activamente en la clase de Educación Física de hoy*"). Moreover, the stem which preceded the eight-item scale was also modified by "*Indica tu grado de desacuerdo o acuerdo con las siguientes afirmaciones, referidas a la clase de Educación Física de hoy...*" / "Indicate your degree of disagreement or agreement with the following statements, referring to today's PE class…". Supplementary File 1 shows the final Spanish version of the SSI-PE-lesson.

Content validity

After that, the content validity of each item of both questionnaires was assessed by a panel of five experts using the Delphi method to come to a consensus on which items adequately reflect students' enjoyment and boredom (Boateng et al., 2018). The acceptance or rejection of an item was based on majority opinion. An agreed item denoted that the item was appropriate, accurate, and interpretable.

Administration of the questionnaires

Data collection was carried out in February 2019. The PE teachers of two state high school centers from Granada and Membrilla (Spain) chosen by convenience were contacted. They were informed about the project, and permission to conduct the study was requested. After the approval of the schools was obtained, students and their legal guardians were fully informed about the features of the project. Students' signed written informed assents and their legal guardians' signed written informed consents were obtained before participating in the study.

Evaluations were carried out during PE lessons. Firstly, participants had to self-report demographic characteristics (i.e., gender, age, and grade). Furthermore, students' body mass and body height were assessed at the beginning of the study following the International Standards for Anthropometric Assessment (Stewart et al., 2011). After that, all the 410 students filled out the original SSI-PE-subject developed by Baena-Extremera et al. (2012) to measure intrinsic enjoyment and boredom in the PE subject. The following week, a subsample of 304 students filled out the adapted SSI-PE-lesson during the cool-down of one PE lesson. The students filled out the questionnaires in an ordinary classroom in quiet and comfortable conditions, and they were asked for their maximum sincerity. At the beginning of the evaluation session, the main researcher explained how to fill out the questionnaire correctly. Furthermore, the researcher was present during the whole evaluation session to clarify any questions that might arise. The questionnaire was completed in 10-15 minutes approximately.

Statistical analysis

Descriptive statistics (mean ± standard deviation) of each item were reported. Skewness and kurtosis of each item were first examined to determine if they met the normality assumption. Additionally, multivariate normality was calculated using the Mardia's multivariate coefficient. Adequate values of normality were considered as follows: skewness < |2|, kurtosis < |7| (Byrne, 2013; Kline, 2011), and Mardia's multivariate coefficient ≤ 70 (Rodríguez & Ruiz, 2008). Then, confirmatory factor analyses of the SSI-PE-subject and SSI-PE-lesson questionnaires were performed. Table 2 showed the eleven measurement models compared for both questionnaires and indicating: (a) the number of factors (i.e., the original two-factor structure or only one-factor); (b) the number of items (i.e., the original eight-item structure or less); and (c) the original items which were removed in each model (e.g., the SSI-PE-subject-1.4 model was a four-item scale following a one-factor structure without the original items 2, 3, 4 and 7).

Table 2. Measurement models compared

Model	Number of factors	Number of items	Original removed items
SSI-PE-subject-2.8	2	8	-
SSI-PE-subject-2.7	2	7	2
SSI-PE-subject-2.6	2	6	2, 7
SSI-PE-subject-2.6b	2	6	1, 2
SSI-PE-subject-2.5	2	5	1, 2, 7
SSI-PE-subject-1.8	1	8	-
SSI-PE-subject-1.7	1	7	2
SSI-PE-subject-1.6	1	6	2, 4
SSI-PE-subject-1.5	1	5	2, 4, 7
SSI-PE-subject-1.5b	1	5	2, 3, 4
SSI-PE-subject-1.4	1	4	2, 3, 4, 7

Note. SSI-PE-subject = Sport Satisfaction Instrument adapted to Physical Education subject

The variances of the error terms were specified as free parameters; in the latent variables (factors), one of the structural coefficients associated with one was set, so that its scale was equal to that of one of the observable variables (items). The maximum likelihood estimation method was used. This method is very robust for non-normality cases, especially if the sample is wide enough and the values of asymmetry and kurtosis are not extreme, as they were in the present study (Byrne, 2013; Kline, 2011). To assess the model's goodness of fit, the chi-squared coefficient (χ^2), the goodness-of-fit index (GFI), and the root mean square error of approximation (RMSEA) were calculated as measures of absolute fit. The adjusted goodness-of-fit index (AGFI), the Tucker-Lewis index (TLI), and the comparative fit index (CFI) were used as incremental fit indices. Chi-squared ratio divided by degrees of freedom (CMIN/DF) and the Akaike information criterion (AIC) were used as measures of parsimonious fit (Byrne, 2013; Gelabert et al., 2011). Goodness of fit was considered as follows: GFI and CFI ≥ 0.95, and RMSEA ≤ 0.08 were considered acceptable, and GFI and CFI ≥ 0.95, and RMSEA ≤ 0.05 were considered optimums (Hu & Bentler, 1999). According to Thompson's (2004) recommendations, the results were interpreted not only corroborating the theoretical model's goodness of fit but also that of various alternative models, employing the maximum likelihood estimation and computing goodness-of-fit indices for those models in order to select the best one. Finally, the internal consistency/reliability of each model of the SSI-PE-subject and SSI-PE-lesson was calculated using Cronbach's alpha (α) and Omega coefficient (Ω) (Revelle & Zinbarg, 2009). Reliability values 0.70–0.79 were considered acceptable, 0.80–0.89 good, and 0.90–1.00 excellent (Nunnally & Bernstein, 1995).

Afterwards, factorial invariance analysis was conducted for females and males. According to Abalo et al.'s (2006) recommendations, the best model obtained in the confirmatory factor analyses was used (i.e., SSI-PE-subject-1.4 and SSI-PE-lesson-1.4). To assess the model's goodness of fit, the χ^2, GFI, normed fit index (NFI), CFI, RMSEA, and AIC were calculated. Then, the difference in mean between the males and females was estimated using the sample females as a reference point, fixing that sample's means to zero and freely estimating the means for the sample of males. The constraints to regression coefficients and intercepts needed for the mean comparisons were applied automatically. All statistical analyses were performed using the SPSS version 25.0 and AMOS version 21.0 for Windows (IBM® SPSS® Statistics; Arbuckle, 2012).

Results

Content validity

The panel of experts raised doubt about the content validity of three items to measure students' enjoyment or boredom in both questionnaires. Specifically, experts considered item 2 (i.e., "*En las clases de Educación Física a menudo sueño despierto en vez de pensar en lo que hago realmente*" / "In PE classes, I often daydream instead of thinking about what I'm really doing"), item 4 (i.e., "*En Educación Física deseo que la clase termine rápidamente*" / "In PE, I usually wish the class would end quickly"), and item 7 (i.e., "*Normalmente participo activamente en las clases de Educación Física*" / "I usually get involved in PE classes") as ambiguous to measure students' boredom (item 2 and 4) and enjoyment (item 7). Therefore, they were considered inappropriate because they did not have adequate content validity.

Firstly, regarding the elimination of the second item of the scale, the expression "to daydream" is defined as "the activity of thinking about pleasant things that you would like to do or have happened to you, instead of thinking about what is happening now" (Cambridge University press online dictionary, 2021). Following this definition, experts highlighted that it is a general phenomenon in people daily lives which generally reflect an individual's current concerns or things that would like to happen (Klinger, 1987; Singer, 2003), and which does not necessarily have to be directly related to boredom during PE lessons.

Regarding item 4, experts suggested that the fact that students wish the class to end quickly may be due to other reasons beyond students' boredom, such as the school schedule. For instance, if the PE lesson is established before recess or just before a student's favorite subject, it can lead students to wish the PE class would end quickly to enjoy their free time or hobbies, but not because it was bored for them. Moreover, regarding item 7 belonging to the enjoyment dimension, the experts also considered it ambiguous because there may be many other reasons why students get involved in PE classes beyond just having fun during lessons. For example, students' perceived competence, how they are evaluated by the teacher (i.e., marks), or students' preferences regarding curricula contents (Bevans et al., 2010; Johnson et al., 2017).

Descriptive analyses, skewness, and kurtosis

Table 3 shows the results of descriptive analyses (mean and standard deviation), skewness, and kurtosis for each of the eight items of both questionnaires (i.e., SSI-PE-subject and SSI-PE-lesson). With reference to the SSI-PE-lesson questionnaire, all skewness and kurtosis indicated that the distribution was normal except for the item 2 which slightly exceeds the recommended value (i.e., skewness = 2.054). Regarding the SSI-PE-subject, all skewness and kurtosis also indicated that the distribution was normal except for the item 4 (i.e., skewness = 2.34). Furthermore, according to the result of Mardia's multivariate coefficient, multivariate normality could be inferred in both questionnaires.

Table 3. Descriptive statistics, skewness, and kurtosis of all the items of the questionnaires.

Item	SSI-PE-lesson (N = 304; 153 females / 151 males)				SSI-PE-subject (N = 402; 200 females / 202 males)			
	M	SD	Skewness	Kurtosis	M	SD	Skewness	Kurtosis
Item 1	8.59	1.73	-1.24	0.68	8.59	1.74	-1.54	2.63
Item 2	2.91	2.35	2.05	3.70	3.04	2.48	1.12	0.16
Item 3	2.21	1.92	1.57	1.54	2.18	1.89	1.94	3.48
Item 4	1.95	1.89	1.52	1.11	2.05	2.02	2.34	4.85
Item 5	8.19	1.94	-0.96	-0.16	8.15	2.01	-1.30	1.54
Item 6	8.44	2.14	-0.74	-0.87	8.43	2.18	-1.55	1.75
Item 7	8.15	1.98	-1.43	1.71	8.14	2.07	-1.31	1.40
Item 8	8.74	1.77	-1.26	0.60	8.70	1.83	-1.83	3.45
Mardia's multivariate coefficient				60.48				70.02

Note. SSI-PE-subject = Sport Satisfaction Instrument adapted to Physical Education subject; SSI-PE-lesson = Sport Satisfaction Instrument adapted to the Physical Education lesson; M = mean, SD = standard deviation

Confirmatory factor analysis and internal consistency for the Sport Satisfaction Instrument adapted to the Physical Education lesson

From the 11 models compared, the overall results corresponding to confirmatory factor analysis showed that all except the model 1.8 could be adequate (see Table 4). However, analyzing more in depth to obtain the most optimal model, the models 2.7, 2.6b, 1.6, 1.5, 1.5b, and 1.4 seem preferable because they presented non-significant chi-squared coefficients. From these 6 models, although the difference with the others is minimal and all of them met the recommended values, the models 1.5 and 1.4 presented the highest GFI and CFI, and the smallest RMSEA, CMIN/DF and AIC values being the most optimal models (see Table 4).

Furthermore, according to the values shown in Table 5, the two-factor models presented high intercorrelations between factors (i.e., factor correlations in 2.7 and 2.6b models = 0.90) and the correlation square was greater than the average variance extracted, showing that models 2.7 and 2.6b did not present adequate discriminant validity. Regarding the one-factor models, the 1.5 and 1.4 showed again the most optimal results (see Table 5), since all the items showed a high loading on their predicted dimension, and also the reliability in both models was excellent (i.e., Ω = 0.922 and 0.920, respectively).

Table 4. Absolute, incremental, and parsimonious fit indices for the generated models of the SSI-PE-lesson questionnaire (*N* = 304)

	Absolute indices			Incremental indices			Parsimonious indices	
Model	χ^2	GFI	RMSEA	AGFI	TLI	CFI	CMIN/DF	AIC
SSI-PE-lesson-2.8	40.726*	0.968	0.061	0.939	0.978	0.985	2.143	74.729
SSI-PE-lesson-2.7	21.675	0.980	0.047	0.957	0.990	0.994	1.667	51.675
SSI-PE-lesson-2.6	16.105*	0.983	0.058	0.955	0.988	0.994	2.013	42.105
SSI-PE-lesson-2.6b	12.892	0.986	0.045	0.964	0.991	0.995	1.611	38.892
SSI-PE-lesson-2.5	7.650	0.990	0.055	0.964	0.990	0.996	1.912	29.650
SSI-PE-lesson-1.8	64.640*	0.947	0.086	0.904	0.957	0.970	3.232	96.640
SSI-PE-lesson-1.7	35.086*	0.970	0.071	0.939	0.978	0.985	2.506	63.086
SSI-PE-lesson-1.6	10.755	0.988	0.025	0.972	0.998	0.999	1.195	34.755
SSI-PE-lesson-1.5	4.577	0.994	0.000	0.981	1.000	1.000	0.915	24.577
SSI-PE-lesson-1.5b	7.929	0.989	0.044	0.968	0.994	0.997	1.586	27.929
SSI-PE-lesson-1.4	1.905	0.997	0.000	0.984	1.000	1.000	0.952	17.905

Note. SSI-PE-lesson = Sport Satisfaction Instrument adapted to the Physical Education lesson; χ^2 = chi-squared coefficient; GFI = goodness-of-fit index; RMSEA = root mean square error of approximation; AGFI = adjusted goodness-of-fit index; TLI = Tucker-Lewis index; CFI = comparative fit index; CMIN/DF = chi-squared fit index over degrees of freedom; AIC = Akaike information criterion. * $p < 0.05$.

Table 5. Standardized confirmatory factor analyses solutions and internal consistency of the SSI-PE-lesson questionnai

Model	Factor	Factor loadings								Facto corre
		Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	
SSI-PE-lesson-2.8	F1	.88	*	*	*	.87	.76	.61	.93	.88
	F2	*	.30	.84	.72	*	*	*	*	
SSI-PE-lesson-2.7	F1	.88	*	*	*	.87	.76	.61	.93	.90
	F2	*	-	.83	.71	*	*	*	*	
SSI-PE-lesson-2.6	F1	.89	*	*	*	.87	.77	-	.92	.90
	F2	*	-	.83	.71	*	*	*	*	
SSI-PE-lesson-2.6b	F1	-	*	*	*	.86	.77	.61	.93	.90
	F2	*	-	.81	.72	*	*	*	*	
SSI-PE-lesson-2.5	F1	-	*	*	*	.87	.77	-	.92	.90
	F2	*	-	.81	.72	*	*	*	*	
SSI-PE-lesson-1.8	F1	.88	.21	.75	.65	.86	.76	.61	.93	*
SSI-PE-lesson-1.7	F1	.88	-	.75	.65	.86	.76	.61	.93	*
SSI-PE-lesson-1.6	F1	.89	-	.74	-	.87	.76	.61	.92	*
SSI-PE-lesson-1.5	F1	.89	-	.74	-	.87	.76	-	.92	*
SSI-PE-lesson-1.5b	F1	.88	-	-	-	.87	.76	.61	.93	*
SSI-PE-lesson-1.4	F1	.88	-	-	-	.87	.77	-	.92	*

Note. SSI-PE-lesson = Sport Satisfaction Instrument adapted to the Physical Education lesson; F1 = Factor related Boredom. AVE = Average Variance Extracted. * Item that does not belong to the evaluated factor; - Item removed from

Confirmatory factor analysis and internal consistency for the Sport Satisfaction Instrument adapted to Physical Education subject

From the 11 models compared, the overall results corresponding to confirmatory factor analysis showed that all, except the 1.8 and 1.7 models, could be acceptable (see Table 6). However, analyzing more in-depth to obtain the most optimal model, the models 2.5, 1.5b, and 1.4 could be preferable because they presented non-significant chi-squared coefficients. From these three models, the models 1.5b and 1.4 presented the highest GFI and CFI, and the smallest RMSEA, CMIN/DF, and AIC values standing out as the most optimal models, although the difference with the 2.5 model is minimal and all of them met the recommended values (see Table 6).

Furthermore, according to the values shown in Table 7, the two-factor model (i.e., 2.5) presented moderate intercorrelations between factors and all the items showed a high loading on their predicted dimensions. However, reliability in boredom dimension was slightly below an acceptable value (i.e., Ω = 0.677). Regarding the one-factor models, the 1.5b and 1.4 ones showed again the most optimal results (see Table 7), since all the items showed a good loading on their predicted dimension, and reliability in both models was excellent (i.e., Ω = 0.908 and 0.920, respectively).

Table 6. Absolute, incremental, and parsimonious fit indices for the generated models of the SSI-PE-subject questionnaire (*N* = 402)

	Absolute indices			Incremental indices			Parsimonious indices	
Model	χ^2	GFI	RMSEA	AGFI	TLI	CFI	CMIN/DF	AIC
SSI-PE-subject-2.8	57.031*	0.967	0.071	0.937	0.946	0.963	3.002	91.031
SSI-PE-subject-2.7	43.368*	0.969	0.076	0.934	0.951	0.970	3.336	73.368
SSI-PE-subject-2.6	23.383*	0.980	0.069	0.948	0.962	0.980	2.923	49.383
SSI-PE-subject-2.6b	19.062*	0.985	0.059	0.960	0.973	0.986	2.383	45.524
SSI-PE-subject-2.5	8.029	0.992	0.060	0.970	0.981	0.993	2.007	30.029
SSI-PE-subject-1.8	125.094*	0.992	0.114	0.860	0.857	0.898	6.255	157.094
SSI-PE-subject-1.7	93.935*	0.933	0.119	0.867	0.880	0.920	6.710	121.935
SSI-PE-subject-1.6	34.995*	0.971	0.085	0.932	0.949	0.969	3.888	58.995
SSI-PE-subject-1.5	17.758*	0.982	0.080	0.946	0.958	0.979	3.552	37.758
SSI-PE-subject-1.5b	4.877	0.995	0.000	0.985	1.000	1.000	0.975	24.877
SSI-PE-subject-1.4	0.782	0.999	0.000	0.995	1.000	1.000	0.391	16.782

Note. SSI-PE-subject = Sport Satisfaction Instrument adapted to Physical Education subject; χ^2 = chi-squared coefficient; GFI = goodness-of-fit index; RMSEA = root mean square error of approximation; AGFI = adjusted goodness-of-fit index; TLI = Tucker-Lewis index; CFI = comparative fit index; CMIN/DF = chi-squared fit index over degrees of freedom; AIC = Akaike information criterion. * $p < 0.05$.

Table 7. Standardized confirmatory factor analyses solutions and internal consistency for the SSI-PE-subject questionnai

Model	Factor	Factor loadings								Facto correl
		Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	
SSI-PE-subject-2.8	F1	.70	*	*	*	.70	.74	.71	.75	.66
	F2	*	.30	.74	.71	*	*	*	*	
SSI-PE-subject-2.7	F1	.70	*	*	*	.70	.74	.71	.75	.68
	F2	*	-	.73	.71	*	*	*	*	
SSI-PE-subject-2.6	F1	.73	*	*	*	.69	.73	-	.72	.73
	F2	*	-	.73	.70	*	*	*	*	
SSI-PE-subject-2.6b	F1	-	*	*	*	.69	.74	.73	.76	.63
	F2	*	-	.71	.72	*	*	*	*	
SSI-PE-subject-2.5	F1	-	*	*	*	.69	.75	-	.74	.67
	F2	*	-	.71	.72	*	*	*	*	
SSI-PE-subject-1.8	F1	.69	.16	.54	.52	.69	.73	.70	.74	*
SSI-PE-subject-1.7	F1	.71	-	.53	.52	.69	.73	.70	.74	*
SSI-PE-subject-1.6	F1	.70	-	.49	-	.70	.74	.72	.74	*
SSI-PE-subject-1.5	F1	.72	-	.53	-	.70	.73	-	.71	*
SSI-PE-subject-1.5b	F1	.68	-	-	-	.70	.75	.74	.75	*
SSI-PE-subject-1.4	F1	.69	-	-	-	.70	.75	-	.73	*

Note. SSI-PE-subject = Sport Satisfaction Instrument adapted to Physical Education subject; F1 = Factor related Enjoym
AVE = Average Variance Extracted. * Item that does not belong to the evaluated factor; - Item removed from this model.

Factorial invariance between females and males

Enjoyment and boredom in the Physical Education lesson

Following the SSI-PE-lesson-1.4 as the most suitable model, its factorial invariance between females and males was analyzed. The Chi-square value and the adjustment indices obtained (GFI = 0.991, CFI = 0.999, RMSEA = 0.032, and AIC = 37.259) allowing to accept the factorial invariance compliance in the model without restrictions (Table 8).

Furthermore, metric invariance was characterized by adding restrictions on factor loadings to the base model. The values shown in Table 8 allowed accepting this level of invariance. The goodness-of-fit index (GFI = 0.981) and the root mean square error of approximation (RMSEA = 0.047) also provided convergent information in the sense of metric invariance. Furthermore, the Akaike information criterion (AIC = 37.756) and the Bentler's comparative fit index (CFI = 0.995) did not undergo large variations with respect to the previous model. Specifically, using the criteria for the evaluation of nested models proposed by Cheung and Rensvold (2002), the difference between CFIs obtained (i.e., less than 0.01) allowed accepting the metric invariance model. Consequently, it can be concluded that factor loadings between females and males were equivalent.

Additionally, once the metric invariance between both samples had been demonstrated, the equivalence between intercepts (i.e., strong factorial invariance) was evaluated. The indices (Table 8) showed an optimal fit of this model, independently assessed or regarding its nesting with the measurement invariance model (CFI differences = 0.004; GFI = 0.981; RMSEA = 0.040). Consequently, accepting the strong invariance, the two models evaluated were equivalent with respect to the factorial coefficients and the intercepts. Furthermore, the factor obtained in the confirmatory factor analysis achieved internal consistency values above 0.85 in both samples (i.e., Ω = 0.904 in males, and Ω = 0.943 in females).

Table 8. Goodness of fit indices for each of the models tested in the factor invariance for the Sport Satisfaction Instrument adapted to the Physical Education lesson (*N* = 304)

Model	Adjusted indices						
	χ^2	gl	GFI	NFI	CFI	RMSEA	AIC
Model without restrictions	5.259	4	.991	.995	.999	.032	37.259
Metric invariance	11.756	7	.981	.988	.995	.047	37.756
Strong factorial invariance	11.843	8	.981	.988	.996	.040	35.843

Note. χ^2 = chi-squared coefficient; gl = grades of freedom; GFI = goodness-of-fit index; NFI = normed fit index; CFI = comparative fit index; RMSEA = root mean square error of approximation; AIC = Akaike information criterion.

Enjoyment and boredom in Physical Education subject

Choosing the SSI-PE-subject-1.4 as the most suitable model, its factorial invariance between gender was analyzed. The Chi-square value and the adjustment indices obtained (GFI = 0.998, CFI = 1.000, RMSEA = 0.000, and AIC = 33.650) allowing to accept the factorial invariance compliance in the model without restrictions (Table 9).

Furthermore, metric invariance was characterized by adding restrictions on factor loadings to the base model. The values shown in Table 9 allowed accepting this level of invariance. The goodness-of-fit index (GFI = 0.985) and the root mean square error of approximation (RMSEA = 0.000) also provided convergent information in the sense of metric invariance. Furthermore, the Akaike information criterion (AIC = 30.161) and the Bentler's comparative fit index (CFI = 1.000) did not undergo large variations with respect to the previous model. Specifically, there are no differences between CFIs obtained in both models, which allows accepting the metric invariance model (Cheung & Rensvold, 2002). Consequently, it can be concluded that factor loadings between females and males were equivalent.

Table 9. Goodness of fit indices for each of the models tested in the factor invariance for the Sport Satisfaction Instrument adapted to the Physical Education subject (*N* = 402)

Model	Adjusted indices						
	χ^2	gl	GFI	NFI	CFI	RMSEA	AIC
Model without restrictions	1.650	4	.998	.997	1.000	.000	33.650
Metric invariance	4.161	7	.985	992	1.000	.000	30.161
Strong factorial invariance	4.306	8	.995	.992	1.000	.000	28.306

Note. χ^2 = chi-squared coefficient; gl = grades of freedom; GFI = goodness-of-fit index; NFI = normed fit index; CFI = comparative fit index; RMSEA = root mean square error of approximation; AIC = Akaike information criterion.

Additionally, once the metric invariance between both samples had been demonstrated, the equivalence between intercepts (i.e., strong factorial invariance) was evaluated. The indices (Table 9) showed an optimal fit of this model, independently assessed or regarding its nesting with the measurement invariance model (CFI differences < 0.001; GFI = 0.995; RMSEA = 0.000). Consequently, accepting the strong invariance, the two models evaluated were equivalent with respect to the factorial coefficients and the intercepts. Furthermore, the factor obtained in the confirmatory factor analysis achieved internal consistency values above 0.78 in both samples (i.e., $\Omega = 0.782$ in males, and $\Omega = 0.839$ in females).

Reliability

The reliability of the factor obtained in the confirmatory factor analyses for the total sample was excellent in both the SSI-PE-subject-1.4 and SSI-PE-lesson-1.4 models (i.e., $\Omega = 0.920$). Moreover, the reliability of the factors obtained in the factorial invariance analysis was from good to excellent for both females and males in the SSI-PE-subject-1.4 (females $\Omega = 0.943$; males $\Omega = 0.904$) and SSI-PE-lesson-1.4 models (females $\Omega = 0.839$; males $\Omega = 0.782$).

Discussion

The first aim of this study was to adapt the Spanish version of SSI-PE-subject by Baena-Extremera et al. (2012) only focused on the PE class that had just finished (i.e., SSI-PE-lesson). Several previous studies have been focused on adapting the original version of the SSI into the Spanish language with adolescents and into PE context (e.g., Baena-Extremera et al., 2012; or Castillo, 2002). However, a new adaptation was required because they were referred to an extended period of time (i.e., students' satisfaction in general during PE lessons), but none of them focused on evaluating the "acute effect" of one particular PE lesson students had just completed. According to the modifications performed and described in the Method section, the adaptation was carried out (see Supplementary File 1), and its applicability to Spanish Secondary students was tested after the data collection. The adaptation of the SSI-PE-lesson allows the comparison between teaching methodologies or activities in intragroup research designs, for instance (i.e., where alternated independent variables could be applied). This allows us to know which are the more satisfactory interventions for students.

After that, the second aim was to evaluate the content validity of the items constituting the questionnaires with experts' judges, who advised removing items 2, 4, and 7 for presenting ambiguities to measure students' enjoyment and boredom, as previously explained in the Results section. Regarding the elimination of item two, García-Calvo et al. (2019) also highlighted that many students did not understand the meaning of this item well when they took the test, and teachers had to explain its meaning on numerous occasions. Therefore, it could lead to wrong answers from the students and errors in the measurement of the boredom dimension. Furthermore, previous research proposed its removal using a 7-item scale (Álvarez et al., 2009; Castillo et al., 2002; Granero-Gallegos et al., 2014) to improve boredom dimension reliability up to acceptable values and guarantee convergent validity of the model. With reference to items 4 and 7, experts suggested their removing due to may be many other reasons beyond students' boredom why they wish the class to end quickly (item 4), as well as why they get involved in PE classes beyond just having fun during lessons (item 7). Furthermore, it should be highlighted that in the original 8-item SSI-PE-subject questionnaire, the enjoyment and boredom factors were not balanced in the number of items (i.e., five items for enjoyment and three items for boredom) which also highlights the difficulty of establishing valid items that measure boredom meeting the standards of the fit indices. This is even more aggravated considering that items 2 and 4, which the experts have classified as inappropriate, belong to the boredom dimension, reducing it to only a single apparently valid item (i.e., item 3). It is also reflected in the poor reliability results of the boredom dimension in other previous 8-item and 7-item versions of the questionnaire (Álvarez et al., 2009; Castillo et al., 2002; Granero-Gallegos et al., 2014). Therefore, based on the reasons mentioned above, in the confirmatory factor analyses of the questionnaires, those models without items whose content validity was questionable (i.e., items 2, 4, and 7) were preferable.

The third aim was to test several structures of the adapted SSI-PE-lesson and the original SSI-PE-subject through confirmatory factor analyses and the analysis of its internal consistency. The present study's findings suggest that it is necessary to distinguish between the results obtained for the SSI-PE-subject and for the SSI-PE-lesson. Firstly, for the SSI-PE-lesson, the two-factor models (i.e., enjoyment and boredom dimensions) were considered inappropriate because they did not present adequate discriminant validity between factors. These results agree with those obtained by Cervelló et al. (1999) that found only the enjoyment factor applying the SSI with adolescent athletes. Furthermore, following common definitions of enjoyment and boredom in previous research (e.g., Baena-Extremera et al., 2012; Castillo et al., 2002; or Vogel-Walcutt et al., 2012), these results also agree with the theory that enjoyment and boredom are opposite poles of the same concept, and therefore, the SSI-PE-subject should be studied using a single dimension that encompasses both satisfaction and boredom subscales (Cervelló et al., 1999).

Therefore, from a quantitative point of view, all one-factor models (except model 1.8) could be adequate and preferable. However, both the models 1.5 and 1.4 presented slightly better results (SSI-PE-lesson-1.5: GFI = 0.994; RMSEA = 0.000; CFI = 1.000; and SSI-PE-lesson-1.4: GFI = 0.997; RMSEA = 0.000; CFI = 1.000) and could be defined as the most optimal models. Furthermore, standardized factor loadings according to the proposed structure for both 1.5 and 1.4 models were appropriate (all items loaded over 0.70), and excellent internal consistencies were observed (i.e., Ω = 0.922 and 0.920, respectively). Besides, these quantitative results agree with experts' judgments about removing items 2, 4 and 7 in both models, ensuring the content validity of the instrument. Secondly, regarding the SSI-PE-subject, from a quantitative point of view, both two-factors and one-factor models (except 1.8 and 1.7) could be appropriated according to the results of the analyses conducted. However, the models 2.5, 1.5b, and 1.4 presented better results (SSI-PE-subject-2.5: GFI = 0.992; RMSEA = 0.060; CFI = 0.993; SSI-PE-subject-1.5b: GFI = 0.995; RMSEA = 0.000; CFI = 1.000; and SSI-PE-subject-1.4: GFI = 0.999; RMSEA = 0.000; CFI = 1.000). Furthermore, standardized factor loadings according to the proposed structure 2.5, 1.5b and 1.4 models were appropriate (all items loaded over 0.68). However, regarding internal consistency, model 2.5 presented worse results, being the reliability in the boredom dimension below an acceptable value (i.e., Ω = 0.677), while models 1.5b and 1.4 obtained excellent internal consistencies (i.e., Ω = 0.908 and 0.920, respectively). Therefore, given the low reliability obtained in the boredom dimension for the 2.5 model, together with the fact that the present study considers enjoyment and boredom as opposite poles of the same concept, the models 1.5b and 1.4 were preferable. Considering the quantitative results with the experts' judgments, model 1.4 seems better because it removed the items with low content validity, while model 1.5b maintained the item *7* classified by the expert as inappropriate.

Considering the reasons mentioned above, researchers of the present study suggest that the most suitable models are those of only one factor, as well as those without the items that the experts considered ambiguous (i.e., items 2, 4, and 7). Consequently, the SSI-PE-lesson 1.5 and 1.4 models would be the most suitable for assessing enjoyment and boredom in one specific PE class, while model SSI-PE-subject-1.4 would be the most suitable for assessing enjoyment and boredom in PE subject. Besides, in order to converge on a common structure for both questionnaires, the model SSI-PE-lesson-1.4 and SSI-PE-subject-1.4 would be the selected models if researchers would like to use them together in the same context. In any case, the present study reported the results obtained with each of the tested models (i.e., the original 8-item model and 10 possible variants) to give the freedom to researchers who want to use it in future studies to select other suitable options that best suit their needs, aims, or ideas.

Finally, the last aim of this study was to verify the factorial invariance of both versions of the scale across gender. The factorial structure of the two scales using the selected 1.4 model was verified as equivalent according to the results obtained in both genders, as well as high structural reliability was obtained given the factor coefficients and intercepts. These findings suggest strong evidence of cross-validation of the measure and, therefore, of the stability of the structure.

Regarding future research, a feasible application could be to adapt the SSI-PE-lesson to a non-paper based format using new technologies (e.g., Additio or Socrative gamification apps) in order to obtain immediate feedback from the students on each part of the lesson (i.e., warm-up, main part and cool-down) using their smartphones or educative tablets to answer the questionnaire. Furthermore, the scale could also be adapted to be more attractive for younger students (e.g., using thumbs up or down, or happy or sad faces), improving their understanding and motivation. Besides that, the SSI-PE-subject and SSI-PE-lesson questionnaires only assess whether the students are having fun or not. However, future studies could design and validate other questionnaires that answer the question "Why is the student having fun?" in order to continue studying the conceptualization of the term enjoyment. The reasons, among others, why students have fun would depend on their age or gender, and that is why probably the enjoyment in PE class for Primary Education students comes from different causes than for Secondary Education students (Dismore & Bailey, 2011; Prochaska et al., 2003; Subramaniam & Silverman, 2007). For this reason, a first phase with discussion groups including students from different educational stages, analyzing their content and obtaining causes of enjoyment during PE lessons adapted to those ages could be a good strategy for future studies. Subsequently, these causes could be taken as dimensions of students' enjoyment in PE class developing a multidimensional questionnaire of enjoyment adapted to each educational stage. This questionnaire would be focused on identifying which elements of the PE lesson make the student enjoy or get bored during the lesson so that the teacher could use their answers to improve students' motivation towards PE, and ultimately, toward the PA practice.

Strengths and limitations

Regarding the strengths of the present study, as far as we know, it is the first study to adapt and validate the SSI-PE-subject questionnaire refers only to the PE class that just finished (i.e., SSI-PE-lesson) for Spanish adolescents. Additionally, to our knowledge, this is also the first study to assess the factorial invariance of the SSI-PE-subject and SSI-PE-lesson questionnaires by gender. However, the present study also has limitations. Due to human, time, and economic resource restrictions, a probability, and larger sample could not be examined. This fact could limit the generalizability of the obtained outcomes to the particular studied population and context. Therefore, future research must confirm the structure obtained in both questionnaires, which will allow for more robust evidence regarding the factorial structure of the questionnaire. Furthermore, the measurements were transversal applying the SSI-PE-lesson only during one specific PE lesson and possible changes between lessons were not considered in this study. Nevertheless, further applications of this questionnaire applying the construct validity analysis (e.g., between two different PE contents/activities) or in longitudinal studies should be considered.

Conclusions

As a result of the present findings, it was concluded that the SSI-PE-subject-1.4 and SSI-PE-lesson-1.4 (Supplementary File 1) are valid, reliable, and available to be applied in Spanish Secondary students' research. Additionally, the factorial invariance analyses support the equivalence of both 4-items questionnaires among males and females. Nevertheless, it is necessary to continue applying this questionnaire to larger samples, involving younger and older Spanish students, and conducting construct validity analyses, which will allow for more robust evidence regarding the factorial structure and the general validity of the questionnaires.

Acknowledgments

The authors want to thank you to Aliisa Hatten for the English revision of the manuscript. We also thank all the students who made this research possible. Carolina Casado-Robles is supported by a research grant from the Spanish Ministry of Universities [grant number: FPU16/03314].

References

Abalo, J., Lévy, J., Rial, A., & Varela, J. (2006). Factor invariance with multiple samples. In: Lévy, J. (Ed). *Modelización con estructuras de covarianzas en Ciencias Sociales* (pp. 259-278). Netbiblo.

Álvarez, M. S., Balaguer, I., Castillo, I., & Duda, J. L. (2009). Coach autonomy support and quality of sport engagement in young soccer players. *The Spanish journal of psychology, 12*(1), 138–148. https://doi.org/10.1017/s1138741600001554

Arbuckle, J. R. (2012). *AMOS users guide version 21.0.* Marketing Department, SPSS Incorporated.

Baena-Extremera, A., & Granero-Gallegos, A. (2013). Efecto de un programa de Educación de Aventura en la orientación al aprendizaje, satisfacción y autoconcepto en secundaria. *Revista Iberoamericana de Diagnóstico y Evaluación, 2*(36), 163-182.

Baena-Extremera, A., & Granero-Gallegos, A. (2015). Versión española del Sport Satisfaction Instrument (SSI) adaptado al aprendizaje de la Educación Física bilingüe en Inglés. *Porta Linguarum, 24*, 63-76. https://doi.org/10.30827/Digibug.53800

Baena-Extremera, A., Granero-Gallegos, A., Bracho-Amador, C., & Pérez-Quero, F. J. (2012). Spanish version of the sport satisfaction instrument (SSI) adapted to physical education. *Revista de Psicodidáctica, 17*(2), 377-396. https://doi.org/10.1387/Rev.Psicodidact.4037

Bai, Y., Allums-Featherston, K., Saint-Maurice, P. F., Welk, G. J., & Candelaria, N. (2018). Evaluation of Youth Enjoyment Toward Physical Activity and Sedentary Behavior. *Pediatric exercise science, 30*(2), 273–280. https://doi.org/10.1123/pes.2017-0101

Balaguer, I., Atienza, F. L., Castillo, I., Moreno, Y., & Duda, J. L. (1997). *Factorial structure of measures of satisfaction/interest in sport and classroom in the case of Spanish adolescents*. In Abstracts of 4thEuropean Conference of Psychological Assessment (p. 76). Lisbon: Portugal.

Baños, R., Marentes, M., Zamarripa, J., Baena-Extremera, A., Ortiz-Camacho, M., & Duarte-Félix, H. (2019). Influencia de la satisfacción, aburrimiento e importancia de la educación física extraescolar en adolescentes mexicanos. *Cuadernos de Psicología del Deporte,19*(3), 205-215. https://doi.org/10.6018/cpd.358461

Barr-Anderson, D. J., Young, D. R., Sallis, J. F., Neumark-Sztainer, D. R., Gittelsohn, J., Webber, L., Saunders, R., Cohen, S., & Jobe, J. B. (2007). Structured physical activity and psychosocial correlates in middle-school girls. *Preventive medicine*, *44*(5), 404–409. https://doi.org/10.1016/j.ypmed.2007.02.012

Bevans, K., Fitzpatrick, L. A., Sanchez, B., & Forrest, C. B. (2010). Individual and Instructional Determinants of Student Engagement in Physical Education. *Journal of teaching in physical education*, *29*(4), 399–416. https://doi.org/10.1123/jtpe.29.4.399

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. *Frontiers in public health*, *6*, 149. https://doi.org/10.3389/fpubh.2018.00149

Byrne, B. (2013). *Structural Equation Modeling with AMOS: Basic concepts, applications, and programming. 2th edition*. Taylor & Francis.

Cambridge University press online dictionary. (2021). Daydreaming. In *Cambridge University press online dictionary*. Retrieved 30 September 2021, from https://dictionary.cambridge.org/es/diccionario/ingles/daydreaming

Casado-Robles, C., Mayorga-Vega, D., Guijarro-Romero, S., & Viciana, J. (2020). Sport education-based irregular teaching unit and students' physical activity during school recess. *The Journal of Educational Research, 113*(4), 262-274. https://doi.org/10.1080/00220671.2020.1806014

Castillo, I., Balaguer, I., & Duda, J. L. (2002). Las perspectivas de meta de los adolescentes en el contexto deportivo. *Psicothema, 14*(2), 280-287.

Cervelló, E., Escartí, A., & Balagué, G. (1999). Relaciones entre la orientación de meta disposicional y la satisfacción con los resultados deportivos, las creencias sobre las causas de éxito en el deporte y la diversión con la práctica deportiva. *Revista de Psicología del Deporte, 8*(1), 7-21.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Dismore, H., & Bailey, R. (2011). Fun and enjoyment in physical education: young people's attitudes, *Research Papers in Education, 26*(4), 499-516. https://doi.org/10.1080/02671522.2010.484866

Duda, J. L., & Nicholls, J. G. (1992). Dimensions of achievement motivation in schoolwork and sport. *Journal of Educational Psychology, 84*(3), 290–299. https://doi.org/10.1037/0022-0663.84.3.290

García-Calvo, S., Banegas, D.L., & Salaberri, M. S. (2019). Satisfaction' study in Bilingual Physical Education after applying a schedule based on content and language integrated learning. *Sportis. Scientific Journal of School Sport, Physical Education and Psychomotricity,* 5(2), 305-322. https://doi.org/10.17979/sportis.2019.5.2.5235

Garcia, J. M., Sirard, J. R., Larsen, R., Bruening, M., Wall, M., & Neumark-Sztainer, D. (2016). Social and psychological factors associated with adolescent physical activity. *Journal of physical activity and health, 13*(9), 957-963. https://doi.org/10.1123/jpah.2015-0224

Gelabert, E., García-Esteve, L., Martín-Santos, R., Gutiérrez, F., Torres, A., & Subirà, S. (2011). Psychometric properties of the Spanish version of the Frost Multidimensional Perfectionism Scale in women. *Psicothema, 23*(1), 133–139.

Granero-Gallegos, A., Baena-Extremera, A., Gómez-López, M., & Abraldes, J. A. (2014). Psychometric properties of the "sport satisfaction instrument (SSI)" in female athletes: predictive model of sport commitment. *Psychological reports*, *115*(1), 148–164. https://doi.org/10.2466/08.06.PR0.115c14z1

Hamilton, J. A., Haier, R. J., & Buchsbaum, M. S. (1984). Intrinsic enjoyment and boredom coping scales: Validation with personality, evoked potential, and attention measures. *Personality and Individual Differences, 5*(2), 183–193. https://doi.org/10.1016/0191-8869(84)90050-3

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Johnson, C. e., Erwin, H. E., Kipp, L., & Beighle, A. (2017). Student Perceived Motivational Climate, Enjoyment, and Physical Activity in Middle School Physical Education. *Journal of Teaching in Physical Education, 36*(4), 398–408. https://doi.org/10.1123/jtpe.2016-0172

Kline, R. (2011). *Principles and practice of structural equation modeling*. Guilford Press.

Klinger, E. (1987). The power of daydreams. *Psychology Today, 21*, 37–44.

Nicholls, J. G., Cheung, P. C., Lauer, J., & Patashnick, M. (1989). Individual differences in academic motivation: Perceived ability, goals, beliefs, and values. *Learning and Individual Differences, 1*(1), 63–84. https://doi.org/10.1016/1041-6080(89)90010-1

Nicholls, J. G., Patashnick, M., & Nolen, S. B. (1985). Adolescents' theories of education. *Journal of Educational Psychology, 77*(6), 683–692. https://doi.org/10.1037/0022-0663.77.6.683

Nunnally, J., & Bernstein, I. (1995). *Psychometric Theory*. McGraw-Hill.

O'Brien, W. (2014). Boredom. *Analysis, 74*(2), 236–244. https://doi.org/10.1093/analys/anu041

Prochaska, J., Sallis, J., Slymen, D., & McKenzie, T. (2003). A longitudinal study of children's enjoyment of physical education. *Pediatric Exercise Science 15*(2), 170-178. https://doi.org/10.1123/pes.15.2.170

Pulido, J. J., Sánchez-Oliva, D., Amado, D., González-Ponce, I., & Sánchez-Miguel, P. A. (2014). Influence of motivational processes on enjoyment, boredom and intention to persist in young sportspersons. *South African journal for research in sport, physical education and recreation, 36*(3), 135-149.

Rehman, A., & Haider, K. (2013). The impact of motivation on learning of secondary school students in Karachi: An analytical study. *Educational Research International, 2*(2), 139-147.

Revelle, W., & Zinbarg, R. E. (2009). Coefficients Alpha, Beta, Omega, and the glb: Comments on Sijtsma. *Psychometrika 74*, 145-154. https://doi.org/10.1007/s11336-008-9102-z

Rodríguez, M., & Ruiz, M. (2008). The reduction of skewness and kurtosis of observed variables by data transformation: Effect on factor structure. *Psicológica 29*(2), 205–227.

Ruiz-Juan, F., Gómez-López, M., Pappous, A., Cárceles, F., & Allende, G. (2010) Dispositional goal orientation, beliefs about the causes of success and intrinsic satisfaction in young elite paddlers. *Journal of Human Kinetics, 26*, 123–136. https://doi.org/10.2478/v10078-010-0056-8

Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology, 61,* 101860. https://doi.org/10.1016/j.cedpsych.2020.101860

Sánchez-Oliva, D., Sánchez-Miguel, P. A., Leo, F. M., Kinnafick, F. E., & García-Calvo, T. (2014). Physical education lessons and physical activity intentions within Spanish secondary schools: A self-determination perspective. *Journal of Teaching in Physical Education*, *33*(2), 232-249. https://doi.org/10.1123/jtpe.2013-0043

Singer, J. L. (2003). Daydreaming, consciousness, and self-representations: Empirical approaches to theories of William James and Sigmund Freud. *Journal of Applied Psychoanalytic Studies, 5*(4), 461–483. https://doi.org/10.1023/A:1026075732578

Stewart, A., Marfell-Jones, M., Olds, T., & De Ridder, J. (2011). *International standards for anthropometric assessment*. International Society for the Advancement of Kinanthropometry.

Subramaniam, P., & Silverman, S. (2007). Middle school students' attitudes toward physical education. *Teaching and Teacher Education, 23*(5), 602–611. https://doi.org/10.1016/j.tate.2007.02.003

Thompson, B. (2004). *Exploratory and confirmatory factor analysis. Understanding concepts and applications*. American Psychological Association.

Vodanovich S. J. (2003). Psychometric measures of boredom: a review of the literature. *The Journal of psychology*, *137*(6), 569–595. https://doi.org/10.1080/00223980309600636

Vogel-Walcutt, J. J., Fiorella, L., Carper, T., & Schatz, S. (2012). The definition, assessment, and mitigation of state boredom within educational settings: A comprehensive review. *Educational Psychology Review, 24*(1), 89–111. https://doi.org/10.1007/s10648-011-9182-7

World Health Organization. (2018). *Promoting physical activity in the education sector*. World Health Organization

Supplementary File 1. Comparison of the previous SSI-PE and the current validated SSI-PE-lesson

Item	Baena-Extremera et al. (2012) SSI-PE	The adapted SSI-PE-lesson
Introductory stem	Indica tu grado de desacuerdo o acuerdo con las siguientes afirmaciones, referidas a tus clases de Educación Física…	Indica tu grado de desacuerdo o acuerdo con las siguientes afirmaciones, referidas a tu clase de Educación Física de hoy…
1	Normalmente me divierto en las clases de Educación Física	En la clase de Educación Física de hoy me he divertido
2	En las clases de Educación Física a menudo sueño despierto en vez de pensar en lo que hago realmente	En la clase de Educación Física de hoy, a menudo soñaba despierto en vez de pensar en lo que hacía realmente
3	En las clases de Educación Física, normalmente me aburro	En la clase de Educación Física de hoy me he aburrido
4	En Educación Física deseo que la clase termine rápidamente	En la clase de Educación Física de hoy, deseaba que la clase terminara rápidamente
5	Normalmente encuentro la Educación Física interesante	He encontrado interesante la clase de Educación Física de hoy
6	Cuando hago Educación Física parece que el tiempo vuela	En la clase de Educación Física de hoy parecía que el tiempo volaba
7	Normalmente participo activamente en las clases de Educación Física	He participado activamente en la clase de Educación Física de hoy
8	Normalmente me lo paso bien haciendo Educación Física	Hoy, me lo he pasado bien haciendo Educación Física

Note. SSI-PE = Sport Satisfaction Instrument adapted to Physical Education; SSI-PE-lesson = Sport Satisfaction Instrument adapted to the Physical Education subject.

SECTION 3

VALIDITY OF ACTIVITY TRACKERS FOR ESTIMATING STUDENTS' DAILY PHYSICAL ACTIVITY AND SEDENTARY BEHAVIOR LEVELS

PAPERS X - XI

Paper X

VALIDITY OF XIAOMI MI BAND 2, 3, 4 AND 5 WRISTBANDS FOR ASSESSING PHYSICAL ACTIVITY IN 12-TO-18-YEAR-OLD ADOLESCENTS

Casado-Robles, C., Mayorga-Vega, D., Guijarro-Romero, S., & Viciana, J.

Draft

Paper X

ABSTRACT

Objective: The purpose of the present study was to assess the convergent validity of four generations of Xiaomi Mi Band wristbands for the assessment of step count and PA levels using the ActiGraph wGT3X-BT as the reference measure among adolescents under free-living conditions.

Methods: A sample of 62 high-school students (34 females and 28 males), aged 12-18 years old (Mage = 14.1±1.6 years) wore a wGT3X-BT accelerometer (ActiGraph, LLC, Pensacola, USA) on their right hip (PA and steps reference measures) and four activity wristbands (Xiaomi Mi Band 2, 3, 4, and 5) on their non-dominant wrist during waking time of one day.

Results: The agreement between the daily PA levels (i.e., light, moderate-to-vigorous and total PA) measured by the accelerometer and the Xiaomi Mi Band wristbands were poor (ICC, 95% IC = 0.06-0.78, 0.00–0.92; MAPE = 50.1-150.6%). However, agreement between daily steps measured by the accelerometer and the Xiaomi Mi Band wristbands were between acceptable (MAPE = 12.2-13.6%) to excellent (ICC, 95% IC = 0.94-0.95, 0.90–0.97). Furthermore, Xiaomi Mi Band wristbands have good to excellent validity for correctly classifying adolescents as meeting or not meeting the recommended 10,000 steps per day (P = 0.89–0.95, k = 0.71–0.87) and the recommended 60 minutes of moderate-to-vigorous PA per day (P = 0.89–0.94, k = 0.69–0.83).

Conclusions: Xiaomi Mi Band wristbands presented good convergent validity for measuring adolescents' steps, and they accurately classified adolescents as meeting or not meeting the international recommendations under free-living conditions.

KEYWORDS

Consumer-wearable activity tracker. Wearable band. Accelerometer. Convergent validity. Step counts. Youth. Free-living

Introduction

The practice of moderate-to-vigorous physical activity (MVPA) is a protective behavior that provides numerous health benefits for adolescents, such as cardiometabolic health, cognitive outcomes, or healthy weight (Chaput et al., 2020; World Health Organization, WHO, 2020). Furthermore, daily total PA levels (i.e., the sum of minutes at all intensities or number of steps per day) is also associated with relevant health indicators and quality of life in adolescents (Poitras et al., 2016). For these reasons, the WHO (2020) recommends achieving, on average, at least 60 minutes daily of MVPA across the week for adolescents. Besides, Parra Saldías et al. (2018) translated these PA guidelines to a simple and easier-to-understand recommendation for adolescents of 10,000 steps per day. However, four out of five adolescents aged 11 to 17 do not meet these PA recommendations (Guthold et al., 2020). This is worrying because physical inactivity is considered a global pandemic, being a leading risk factor for global mortality (WHO, 2020). Therefore, the promotion of adequate PA levels has been considered a scientific research priority, and a global action plan on PA has been developed to reverse these current trends (Gillis et al., 2013; WHO, 2018).

However, to check the effectiveness of these global policies and monitor their progress (WHO, 2018), it is necessary to objectively measure adolescents' PA levels across time (Brooke et al., 2014; Corder et al., 2007; Metcalf et al., 2012). Among the large number of methods for the assessment of adolescents' PA levels, research-grade accelerometers (e.g., Actigraph accelerometers) have been highlighted as the most common and valid method for objectively assessing adolescents' PA levels during free-living conditions (Romanzini et al., 2014; Shephard & Tudor-Locke, 2016; Van Hecke et al., 2016). These research-grade accelerometers provide information about the intensity, frequency, and volume of PA (Dhurandhar et al., 2015). Nevertheless, these research-grade accelerometers are usually very expensive instruments, and in addition, they are unattractive and not interactive with the users (only the new models present a limited programmable display showing steps and kcals), which make them not very useful to promote adolescents' PA practice (Actigraph Corporation, 2021; Šimůnek et al., 2019). On the contrary, the new consumer-wearable activity trackers share elements of research-based devices and they are generally cheaper, more interactive, more user-friendly, and are increasingly being used in research not only for measuring but also for promoting PA levels (Gorzelitz et al., 2020; Henriksen et al., 2018).

These consumer-wearable activity trackers are electronic devices worn on the body as an accessory to monitor and record daily PA and fitness-related metrics and provide users real-time behavioral feedback (Ruiz & Goransson, 2015; Strath & Rowley, 2018). They usually integrate an accelerometer or pedometer to track physical movements automatically, and their outputs are generally based on step counts, the amount and intensity of PA, energy expenditure, periods of inactivity, sleep time, or heart rate (Franssen et al., 2020; Mercer et al., 2016). Moreover, these devices can be synchronized with their specific smartphone apps, PCs, or websites to obtain more detailed feedback over days and weeks (Alley et al., 2016). Besides the real-time feedback, these devices often include other features that also may be facilitators of users' positive behavior change such as personalized goal-setting (i.e., a goal based on daily steps or minutes of PA), self-efficacy, peer comparison, or social support (Brickwood et al., 2019; Patel et al., 2015; Rich & Miah, 2016). Therefore, these devices can be an ideal cost-effective option both to objectively measure adolescents' PA levels, as well as a motivational instrument to promote their PA practice through self-monitoring behavior technique (Michie et al., 2009, 2013; Strath & Rowley, 2018).

Furthermore, the widespread sales of consumer-wearable activity trackers in recent years reflect the increasing popularity of these devices among the general public (International Data Corporation, 2020). However, there are different kinds of wearables (e.g., wristbands, smartwatches, or pedometers), as well as a large number of ever-growing brands and models available in the market, which can vary considerably in terms of the features they include, their accuracy or their cost (Henriksen et al., 2018). Specifically, wristband devices are preferred by the general population, especially by younger users, due to their low weight, size, price, easy use, or smart design (Alley et al., 2016; Stamm & Hartanto, 2018). Regarding the wide range of brands and models, the Xiaomi Mi Band is one of the top-3 most worldwide used wristbands, reflected in the millions of units shipped internationally year-over-year, and which also presents a low price (Henriksen et al., 2018; International Data Corporation, 2020). Furthermore, the Xiaomi Mi Band wristbands presented the highest rating for users' willingness to wear and buy this device compared to other recognized brands like Apple, Samsung, Huawei or Fitbit (Jia et al., 2018). This makes the Xiaomi Mi Band an excellent cost-effective option to measure and promote adolescents' PA levels.

Nevertheless, although the use of wristbands to monitor PA levels is increasingly widespread, its validity has not been sufficiently studied, especially among children and adolescents (Evenson et al., 2015; Fuller et al., 2020; Gorzelitz et al., 2020). Furthermore, previous studies specifically assessing the validity of the Xiaomi Mi Band to measure PA are still more scarce. Specifically, the systematic reviews by Fuller et al. (2020) and Henriksen et al. (2018) highlighted that validity of Xiaomi Mi Bands are the least studied wristbands, including only two and one studies, respectively. However, none of them were carried out among adolescents, and they only considered the oldest generations (i.e., Mi Band, and Mi Band 2). In addition, many other systematic reviews about this topic did not include any study about the Xiaomi brand (e.g., Cosoli et al., 2020; Evenson et al., 2015; Gorzelitz et al., 2020). Besides that, more recent studies have assessed the validity of Xiaomi Mi Band wristbands (i.e., Mi Band 2, 3, and 4), but most of them were carried out in adult population and/or under laboratory setting (e.g., Topalidis et al., 2021; Pino-Ortega et al., 2021; DeGroote et al., 2020; Hartung et al., 2020; Stamm & Hartanto, 2018; Tam & Cheung, 2018). To our knowledge, only the studies by Hao et al. (2021), Campos-Meirinhos et al. (2019) and Yang et al. (2019) assessed the validity of Xiaomi Mi Band 2 wristbands to measure total PA among adolescents under free-living conditions. Therefore, assessing the validity of the newest generations of Xiaomi Mi Band wristbands for assessing PA among adolescents under free-living conditions is still lacking, which is necessary to assess the effectiveness of intervention programs designed to increase PA (DeGroote et al., 2020).

Consequently, the main purpose of the present study was to assess the convergent validity of four generations of Xiaomi Mi Band wristbands (i.e., Mi Band 2, 3, 4, and 5) for the assessment of step count and PA levels using the ActiGraph wGT3X-BT as the reference measure among adolescents aged 12-18 years under free-living conditions. The secondary aim of this study was to assess the convergent validity of Xiaomi Mi Band wristbands for correctly classifying adolescents as meeting or not meeting the international PA recommendations.

Method

Participants

The present study is reported according to the GRRAS guidelines (Kottner et al., 2011). The protocol of the present study conforms to the Declaration of Helsinki statements (64th WMA, Brazil, October 2013) and it was approved by the Ethical Committee for Human Studies at the University of Granada. The present agreement study followed a cross-sectional design.

The state high school chosen by convenience was located in the city of Motril (Granada, Spain). A total of 100 adolescents (i.e., from 7^{th} to 11^{th} grade, inviting 10 males and 10 females per grade) enrolled in the selected school were invited to participate in the present study. According to the school's reports, all the adolescents' families had a middle socioeconomic level. The inclusion criteria were: (a) being enrolled in the selected high school (i.e., in the 7^{th} to 11^{th} grade); (b) being free of any health disorder that would make them unable to engage in PA normally; (c) presenting the corresponding signed written informed assents of the adolescents, and (d) presenting the corresponding signed written informed consents of their legal guardians. The exclusion criterion was not having completed and valid data from the four wristbands and the accelerometer.

A priori sample size calculation was estimated with the Arifin´s web-based sample size calculator (Arifin, 2018). Based on steps values, parameters were set as follows: ICC, $\rho0$ = 0.70 (Nunnally, 1978); $\rho1$ = 0.85 (Campos-Meirinhos et al., 2019), α = 0.05, 1 – β = 0.80, k = 2, dropout = 23% (Howie & Straker, 2016). Kappa, k_0 = 0.40 (Cicchetti, 2001); k_1 = 0.80 (Mayorga-Vega et al., 2021), p = 0.25 (Guthold et al., 2020), α = 0.05, 1 – β = 0.80, k = 2, dropout = 23% (Howie & Straker, 2016). A final sample size of at least 53 adolescents (minimum initial sample size equal to 69) was estimated. In addition to exceed the minimum required sample size, obtaining a sample balanced by grade and gender was also aimed.

Measures

Demographic characteristics

Adolescents' age (in years), grade (7^{th} to 11^{th}grade), gender (males/females), and non-dominant hand (left/right) information was self-reported.

Anthropometric measures

Adolescents' anthropometric measurements were measured following the International Standards for Anthropometric Assessment (Stewart et al., 2011). Firstly, body mass (Seca, Ltd., Hamburg, Germany; accuracy = 0.1 kg) and height (Holtain Ltd., Pembs, United Kingdom; accuracy = 0.1 cm) were measured. Then, the body mass index was calculated as body mass divided by body height squared (kg/m^2). Finally, adolescents' body weight status was categorized by gender- and age-adjusted body mass index thresholds as overweight/obesity or non-overweight/obesity (Cole et al., 2000). Body mass index and body weight status scores have shown high evidence supporting validity among adolescents (Cole et al., 2000).

Habitual physical activity

Adolescents' habitual PA (days/week) was estimated by the adapted and validated Spanish version of the Physician-based Assessment and Counseling for Exercise questionnaire (PACE) for adolescents (Martínez-Gómez et al., 2009). It consists of two questions that assess how many days in the last week (i.e., "In the last 7 days, how many days did you do PA for 60 minutes or more?") and in a normal week (i.e., "In a normal week, how many days do you do PA for 60 minutes or more?") at least 60 minutes of PA are performed. The items were preceded by a brief explanation about what PA is and some examples, indicating that the time spent in school physical education should not be included. A 7-point Likert-type scale, ranging from 0 to 7 days was used. Then, the mean of two items was calculated. The Spanish version of the PACE questionnaire has shown adequate convergent validity (accelerometer) for assessing PA among adolescents (r = 0.43; Martínez-Gómez et al., 2009).

Wristbands

Adolescents' PA levels were estimated by the Xiaomi Mi Band 2, 3, 4, and 5 wristbands (Xiaomi, Pekin, China). According to the user manual, the wristbands were fit snugly on the top of adolescents' non-dominant wrist, close and above the wrist bone. Moreover, the four devices display were blinded not to show adolescents' PA, avoiding potential biases due to adolescents' reactivity. These wristbands are characterized to be small, light-weight, and relatively inexpensive considering their launch price in Spain (Mi Band 2: 4.30 x 1.57 x 1.05 cm, 7.0 g, and 25€; Mi Band 3: 4.69 x 1.79 x 1.20 cm, 20.0 g, and 30€; Mi Band 4: 4.70 x 1.81 x 1.08 cm, 22.1 g, and 35€; Mi Band 5: 4.69 x 1.81 x 1.24 cm, 11.9 g, and 35€). The models Mi Band 2 and 3 are based on tri-axial built-in accelerometers, while Mi Band 4 and 5 are based on both tri-axial built-in accelerometers and tri-axial built-in gyroscope. Furthermore, these wristbands have their own algorithmic equation to estimate the daily steps taken and the minutes engaged in each specific intensity-related PA. Wristbands data were synchronized via Bluetooth to its specific application to download and store data (i.e., Mi Fit version 5.3.2 for Android), and the features of interest were calculated from the data shown in this app at the end of the data collection.

Regarding the data scoring, steps (number) were registered as directly reported in the Mi Fit app. However, specific information regarding intensity-related PA cut points is not publicly available. Therefore, intensity-related PA scores (minutes) were calculated as follows: (a) "slow walking" was calculated by summing the total time spent on all the bouts of "slow walking", assuming that the measured variable corresponded to light PA as measured by the ActiGraph; (b) "brisk walking" was calculated by summing the total time spent on all the bouts of "fast walking", assuming that the measured variable corresponded to MVPA as measured by the ActiGraph; (c) "MVPA" was calculated by summing the total time spent on all the bouts of "moderate activity" and "vigorous activity", assuming that the measured variable corresponded to MVPA as measured by the ActiGraph; (d) "slow-brisk walking" was also calculated by summing the total time spent on all the bouts of "slow walking" and "fast walking", assuming that the measured variable corresponded to total PA as measured by the ActiGraph; and (e) "total PA" was registered as directly reported in the Mi Fit app as "active minutes", assuming that the measured variable corresponded to total PA as measured by the ActiGraph.

Accelerometer

Adolescents' reference standards of steps and PA were determined by wGT3X-BT accelerometers (ActiGraph, LLC, Pensacola, FL, USA), adjusted on the adolescents' right hip. This model is a small (1.05 x 3.03 x 4.06 cm) and light-weight (19 g) commercially available triaxial accelerometer. Accelerometer data were initialized, downloaded, and processed using the ActiLife Lifestyle Monitoring System Software version 6.13.3 (ActiGraph, LLC, Pensacola, FL, USA).

Accelerometers were initialized with a sample ration of 30 Hz (Evenson et al., 2008; Trost et al., 2011; Migueles et al., 2017). A 15-second epoch was used when downloading the data (Evenson et al., 2008; Migueles et al., 2017). Valid wear time was set as equal to or higher than 10 hours per day (Migueles et al., 2017), with non-wear periods set as 60 minutes or more of consecutive zero-count *epochs* with up to two minutes spike tolerance (Oliver et al., 2011).

Regarding the data scoring, steps (number) were assessed by the default settings of ActiLife for step count. Furthermore, the Evenson's cut-points were applied to categorize the time (minutes) engaged in each specific intensity-related PA (i.e., light PA = 101-2295 *counts*/min; MVPA ≥ 2,296 *counts*/min; and total PA ≥ 101 *counts*/min; Evenson et al., 2008). According to the cross-validation study performed by Trost et al. (2011), these cut-off points have demonstrated the best evidence supporting score validity for assessing intensity-related PA among adolescents. Finally, adolescents' steps and MVPA were dichotomized as achieving or not achieving the daily recommendation of at least 10,000 steps and 60 minutes of MVPA, respectively (Parra Saldías et al., 2018; WHO, 2020). ActiGraph wGT3X-BT accelerometer scores have shown high evidence supporting validity for assessing steps and intensity-related PA among adolescents (Lee et al., 2015; Romanzini et al., 2014; Trost et al., 2011).

Procedure

Firstly, the principal and the physical education teachers of the high school chosen by convenience were contacted. They were informed about the project, and permission to conduct the study was requested. After the approval of the school was obtained, all the students and their legal guardians were fully informed about the features of the project. Afterward, adolescents' demographic characteristics, anthropometric measurements, and self-reported habitual PA levels were recorded. Then, wristbands and accelerometers were adjusted on adolescents from Monday to Thursday, while Fridays were used to collect the activity trackers, download data, and charge batteries. Due to the limitations of material resources, waves of seven adolescents per day were carried out.

For each wave, adolescents were met from 8:00 to 8:30 a.m. in the school gym. According to the user manual, the four wristbands were simultaneously adjusted on the adolescents' wrists of the non-dominant hand, while the accelerometer was adjusted on the adolescents' right hip using an elastic waistband. In order to avoid that the relative position of the four wristbands on the wrist influenced the results, they were adjusted in random order varying across adolescents (Hartung et al., 2020). Moreover, adolescents were instructed to wear the activity trackers for one whole day until bedtime and only take them off when they took a bath/shower. Furthermore, adolescents were provided with a diary to record when they put on and took off the devices throughout the day. In order to not influence participants' PA patterns, they were instructed to maintain their habitual PA levels. Lastly, adolescents were also required to remove the wristbands and the accelerometer when they go to bed, leaving them in a plastic box inside their schoolbags. In the morning of the following day, the activity trackers were collected and adjusted to the next seven participants, following the same protocol.

Statistical analysis

Descriptive statistics for all the variables of the included participants were calculated. Firstly, all the statistical tests assumptions were check and met (e.g., histograms and Q-Q plots for normality). Furthermore, univariate (i.e., $z \pm 3.0$) and multivariate outliers (i.e., Mahalanobis distance) were removed. Afterward, the agreement between the PA scores (i.e., continuous variables) measured by the wristbands and the accelerometers were calculated as follows: a) Limits of Agreement (LOA) with its confident intervals (95% CI) (Bland & Altman, 1986); b) Mean Absolute Error (MAE) (Willmott & Matsuura, 2005); c) Mean Absolute Percentage Error (MAPE) (Johnston et al., 2021); and d) Intraclass Correlation Coefficient (ICC), and its 95% CI, by a two-way random-effects model with absolute agreement and single measurement [also known as ICC(2,1)] (Koo & Li, 2016). Additionally, LOA plots, which are the individual participant differences between the two scores plotted against the respective individual means, were performed (Bland & Altman, 1986). Heteroscedasticity was also examined objectively by calculating the Pearson´s correlation coefficient (*r*) between the absolute differences and the individual means (Atkinson & Nevill, 1998). Based on Cohen's (1992) benchmarks, a correlation coefficient > 0.50 was considered as indicative of heteroscedasticity. Finally, the agreement between the PA scores dichotomized as achieving or not achieving the daily recommendations of 10,000 steps (Mayorga-Vega et al., 2021) and 60 min of MVPA (i.e., categorical variables) (WHO, 2020) measured by the wristbands and the accelerometers were calculated as the proportion of agreement [*P* = number of agreements/ (number of agreements + disagreements)] and kappa coefficient (*k*) (Hernaez, 2015). Agreement values were interpreted as follows: MAPE, > 15.0% poor, 10.1-15.0% acceptable, 5.1-10.0% good, and 0.0-5.0% excellent (Johnston et al., 2021); ICC, 0.00-0.69 poor, 0.70-0.79 acceptable, 0.80-0.89 good, and 0.90-1.00 excellent (Nunnally, 1978); k, 0.00-0.39 poor, 0.40-0.59 acceptable, 0.60-0.74 good, and 0.75-1.00 excellent (Cicchetti, 2001). Based on statistical inference, each ICC value was interpreted according to its 95% IC, that means, there was 95% chance that the true ICC value landed on any point between the 95% IC range (Koo & Li, 2016). All statistical analyses were performed using the SPSS version 25.0 for Windows (IBM® SPSS® Statistics). The statistical significance level was set at $p < 0.05$.

Results

General characteristics

Figure 1 shows the flow diagram of the participants through the study. An initial sample of 70 adolescents agreed to participate in the study and met the inclusion criteria. However, since some adolescents met the exclusion criterion, the final sample consisted of 62 participants (i.e., non-compliance rate of 11.4%). Table 1 shows the general characteristics of the included participants.

Table 1. General characteristics of the analyzed participants

	Sample (n = 62)
Age (years)	14.1 (1.6)
Grade ($7^{st}/8^{nd}/9^{rd}/10^{th}/11^{th}$)	22.6/17.7/19.4/21.0/19.4
Gender (males/females)	45.2/54.8
Body mass (kg)	58.5 (12.7)
Body height (cm)	161.5 (8.3)
Body mass index (kg/m^2)	22.4 (4.5)
Overweight/obesity (no/yes)	66.1/33.9
Non-dominant hand (left/right)	88.7/11.3
Self-reported habitual PA (days/week)	2.2 (1.5)

Note. Data are reported as mean (standard deviation) or percentage.

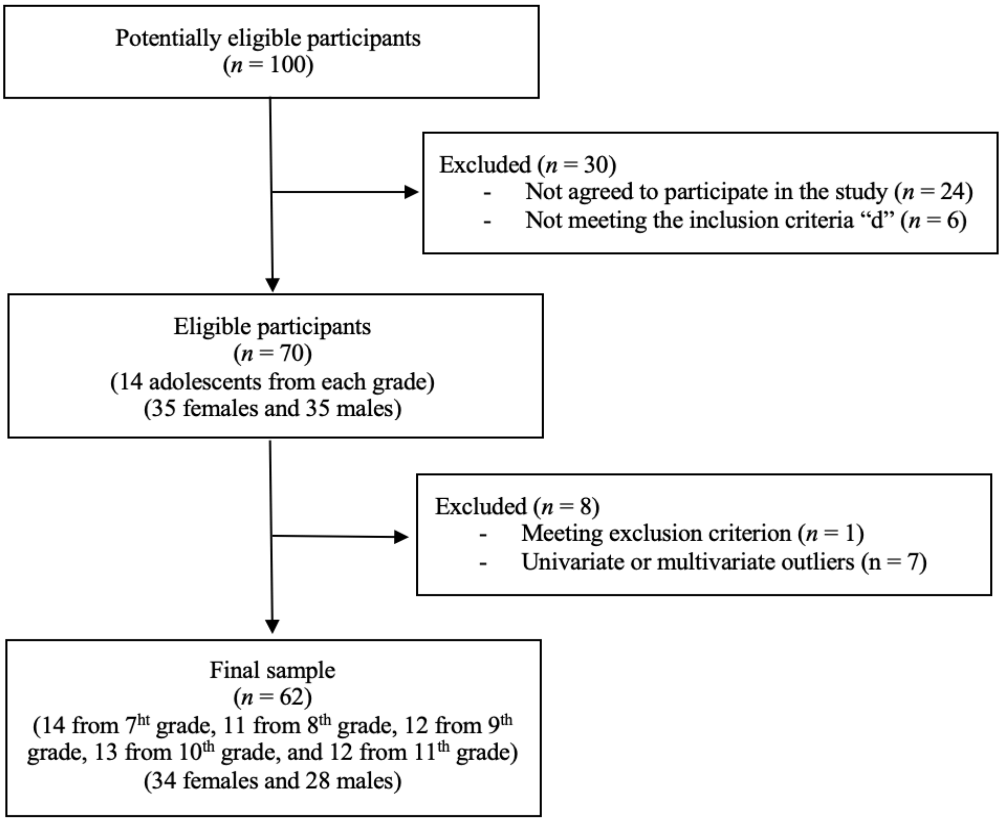


Figure 1. Flow chart of participants included in the present study.

Validity of the Xiaomi Mi Band wristbands during daily free-living conditions

Table 2 shows the validity of the wristbands for estimating daily PA in adolescents during free-living conditions. According to the 95% IC of the ICC values, differences between the studied generations of the Xiaomi Mi Band on the validity results within each PA score were not found. Based on the 95% IC values of the ICC, the validity results of the steps assessed by all the studied generations of the Xiaomi Mi Band were excellent. Similarly, based on the MAPE values, all studied generations of the Xiaomi Mi Band showed an acceptable validity. However, the validity results of all the studied generations of the Xiaomi Mi Band for assessing slow, brisk, and slow-brisk pace walking, total PA and MVPA was poor. Supplementary File 1 shows the LOA plots. Pearson´s correlation coefficients did not show heteroscedasticity with any wristbands on steps and brisk walking scores. However, heteroscedasticity was found for slow and slow-brisk pace walking, MVPA and total PA by all the generations of Xiaomi Mi Band wristbands (except for the Xiaomi Mi Band 4 and 5 with MVPA) (Supplementary File 2).

Table 2. Validity of the wearable bands for estimating daily physical activity during unstructured free-living conditions (n = 62)

Instrument	Mean (SD)	LOA (95% CI)	MAE	MAPE	ICC (95% CI)
Steps (n)					
ActiGraph wGT3X-BT	7,066.8 (3622.3)	-	-	-	-
Xiaomi Mi Band 2	7,287.0 (4017.2)	-220.2 (-2461.5, 2021.1)	791.8	12.3	0.95 (0.93, 0.97)
Xiaomi Mi Band 3	7,193.9 (4041.0)	-127.2 (-2574.8, 2320.4)	890.6	13.6	0.95 (0.91, 0.97)
Xiaomi Mi Band 4	7,544.0 (4124.2)	-477.2 (-2927.2, 1972.8)	916.9	13.2	0.94 (0.90, 0.97)
Xiaomi Mi Band 5	7,441.9 (4068.2)	-375.1 (-2634.2, 1884.0)	824.5	12.2	0.95 (0.92, 0.97)
Slow pace walking (min)[a]					
ActiGraph wGT3X-BT	169.1 (59.5)	-	-	-	-
Xiaomi Mi Band 2	40.1 (30.3)	129.1 (37.2, 221.0)	129.1	77.7	0.11 (0.00, 0.36)
Xiaomi Mi Band 3	40.6 (31.5)	128.6 (35.1, 222.1)	128.6	77.0	0.11 (0.00, 0.36)
Xiaomi Mi Band 4	45.1 (30.6)	124.0 (28.4, 219.6)	124.0	74.4	0.11 (0.00, 0.35)
Xiaomi Mi Band 5	43.1 (30.1)	126.0 (36.4, 215.6)	126.0	75.7	0.12 (0.00, 0.38)
Brisk pace walking (min)[b]					
ActiGraph wGT3X-BT	45.1 (28.4)	-	-	-	-
Xiaomi Mi Band 2	57.5 (34.7)	15.2 (-13.2, 43.6)	17.6	50.1	0.78 (0.18, 0.92)
Xiaomi Mi Band 3	71.1 (47.6)	17.4 (-12.8, 47.6)	19.9	54.9	0.74 (0.08, 0.90)
Xiaomi Mi Band 4	45.8 (29.0)	14.9 (-14.9, 44.7)	17.7	50.1	0.78 (0.22, 0.91)
Xiaomi Mi Band 5	76.7 (34.9)	13.8 (-16.6, 44.2)	17.5	51.6	0.78 (0.31, 0.91)
Slow-brisk pace walking (min)[c]					
ActiGraph wGT3X-BT	214.2 (71.1)	-	-	-	-
Xiaomi Mi Band 2	70.0 (46.1)	144.2 (51.1, 237.3)	144.2	69.6	0.18 (0.00, 0.50)
Xiaomi Mi Band 3	68.3 (44.7)	145.9 (48.5, 243.3)	145.9	69.8	0.16 (0.00, 0.48)
Xiaomi Mi Band 4	75.3 (47.2)	138.9 (41.3, 236.5)	138.9	66.7	0.18 (0.00, 0.51)
Xiaomi Mi Band 5	74.4 (47.4)	139.8 (45.1, 234.5)	139.8	67.2	0.19 (0.00, 0.52)
Moderate-to-vigorous physical activity (min)					
ActiGraph wGT3X-BT	45.1 (28,4)	-	-	-	-
Xiaomi Mi Band 2	29.9 (30.4)	-12.4 (-93.2, 68.4)	35.3	110.8	0.15 (0.00, 0.37)
Xiaomi Mi Band 3	27.7 (31.3)	-26.0 (-130.5, 78.5)	43.5	137.5	0.06 (0.00, 0.27)
Xiaomi Mi Band 4	30.1 (31.3)	-0.8 (-72.9, 71.3)	29.6	89.2	0.18 (0.00, 0.41)
Xiaomi Mi Band 5	31.3 (30.4)	-31.6 (-116.1, 52.9)	42.4	150.6	0.06 (0.00, 0.24)
Total physical activity (min)					
ActiGraph wGT3X-BT	214.2 (71.1)	-	-	-	-
Xiaomi Mi Band 2	86.9 (40.9)	127.3 (32.6, 222.0)	127.3	60.1	0.19 (0.00, 0.52)
Xiaomi Mi Band 3	86.1 (41.6)	128.1 (31.1, 225.1)	128.1	60.4	0.19 (0.00, 0.51)
Xiaomi Mi Band 4	89.8 (42.2)	124.4 (30.5, 218.3)	124.4	58.7	0.20 (0.00, 0.54)
Xiaomi Mi Band 5	88.0 (41.3)	126.2 (29.8, 222.6)	126.2	59.5	0.19 (0.00, 0.52)

Note. SD = Standard deviation; LOA = Limits of agreement; 95% CI = 95% confident interval; MAE = Mean absolute error; MAPE = Mean absolute percentage error; ICC = Intraclass correlation coefficient; [a] Compared with the accelerometer-measured light physical activity (min); [b] Compared with the accelerometer-measured moderate-to-vigorous physical activity (min); [c] Compared with the accelerometer-measured total physical activity (min).

Table 3 shows the validity of the wristbands for estimating the daily PA recommendations in adolescents during free-living conditions. A total of 24.2% and 29.0% of adolescents met the accelerometer-measured step- and MVPA-based recommendations, respectively. The validity results of the daily step-based recommendations assessed by the Xiaomi Mi Band wristbands ranged from good to excellent. Moreover, based on the brisk pace walking score of the Xiaomi Mi Band wristbands, the validity results of the daily MVPA-based recommendations also ranged from good to excellent. However, based on the MVPA score of the Xiaomi Mi Band wristbands, the validity results of the daily MVPA-based recommendations were poor.

Table 3. Validity of the wearable bands for estimating the daily physical activity recommendations during unstructured free-living conditions (*n* = 62)

		ActiGraph wGT3X-BT									
		10,000 steps				60 min of MVPA				60 min of MVPA		
Instrument		%TP	*P*	*k*		%TP	*P*	*k*		%TP	*P*	*k*
Xiaomi Mi Band 2	10,000 steps	25.8	0.92	0.79†	Brisk pace walking	22.6	0.94	0.83†	60 min of MVPA	38.7	0.58	0.07
Xiaomi Mi Band 3		25.8	0.95	0.87†		17.7	0.89	0.69†		53.2	0.56	0.15
Xiaomi Mi Band 4		29.0	0.89	0.71†		21.0	0.92	0.79†		37.1	0.53	-0.05
Xiaomi Mi Band 5		27.4	0.94	0.83†		21.0	0.92	0.79†		64.5	0.48	0.08

Note. MVPA = Moderate-to-vigorous physical activity; %TP= Wearable-based percentage of total positive cases according to the recommendation; *P* = Proportion of agreement; *k* = Kappa coefficient.

† $p < 0.001$

Discussion

The main purpose of the present study was to assess the convergent validity of four generations of Xiaomi Mi Band wristbands (i.e., Mi Band 2, 3, 4, and 5) for the assessment of step count and PA levels using the ActiGraph wGT3X-BT as the reference measure among adolescents aged 12-18 years under free-living conditions. Firstly, the findings of the present study showed that the Xiaomi Mi Band wristbands have between acceptable (MAPE = 12.2-13.6%) to excellent (ICC, 95% IC = 0.94-0.95, 0.90–0.97) agreement results for step count comparing with the wGT3X-BT accelerometer. These results are of great importance due to the main wristbands PA output is total daily steps, which is considered a simple, easier-to-understand, and credible indicator of daily PA (Parra-Saldías et al., 2018; Tudor-Locke et al., 2011). Furthermore, users highlighted step count as the most useful feature on their activity trackers (Maher et al., 2017), and its importance is also reflected in the fact that most consumer-wearable activity trackers-based interventions to promote PA in adolescents used steps per day as the most appropriate goal (Strath & Rowley, 2018).

Regarding previous research about the validity of Xiaomi Mi Band wristbands for measuring adolescents' steps count, the results of the present study agree with those carried out by Campos-Meirinhos et al. (2019) about the excellent validity of Xiaomi Mi Band 2 under free-living conditions (i.e., MAPE = 12.3% vs. 12.7%; ICC = 0.95 vs. 0.90). Besides that, the present results also agree with those obtained by Yang et al. (2019) study about Xiaomi Mi Band 2 wristband validity for measuring steps among children and adolescents (i.e., MAPE = 12.2% vs. 14.5%). However, comparing the results from the Bland-Altman plots, the present results seem greatly better than those by Yang et al. (2019) for Xiaomi Mi Band 2 (LOA; 95% CI = -220.2; -2,462–2,021 vs. -633.5; -6,981–5,714, respectively). Furthermore, the present study also seems slightly better results than those obtained by Hao et al. (2021) to assess the validity of Xiaomi Mi Band 2 for step count under simulated free-living conditions. Specifically, although both studies obtained that Xiaomi Mi Band 2 overestimate the step count compared to Actigraph accelerometers, the MAPE values in the present study are greatly lower (i.e., MAPE = 12.3% vs. 21.3%). These differences may be due to methodological decisions such as Actigraph accelerometer placement. Specifically, the study carried out by Hao et al. (2021) adjusted the accelerometer on the adolescents' non-dominant wrist, while the previous studies (Campos-Meirinhos et al., 2019; Yang et al., 2019) and the present one placed it on their hip. In this line, previous empirical studies have found that wrist- and hip-worn accelerometer steps outputs are not always comparable (Evenson et al., 2015; Tudor-Locke et al., 2015). However, as in the present study, hip-worn accelerometers are considered the reference placement for assessing PA through accelerometry (Migueles et al., 2017). Besides that, these differences may also be affected by the adolescents' PA patterns of each sample. Specifically, adolescents in the study by Hao et al. (2021) are more physically active than those in the present study (i.e., Mean Actigraph step count = 10,143.8 vs. 7,066.8, respectively), which could be translated into that the greater the number of steps per day, the greater the bias introduced by the Xiaomi Mi Band 2. Unfortunately, due to the lack of research on the validity of the other Xiaomi Mi Band generations (i.e., 3, 4, and 5) among adolescents to measure steps, the present results have also been compared with available literature in other populations and settings. Even so, no previous studies have been found about the validity of Xiaomi Mi Band 5 for measuring steps.

Regarding studies under free-living conditions among healthy adults, DeGroote et al. (2020) and Topalidis et al. (2021) carried out concurrent validity studies of Xiaomi Mi Band wristbands for measuring steps (i.e., Mi Band 2 and 3, respectively). Both studies showed a strong positive correlation of Xiaomi Mi Band wristbands and ActiGraph accelerometers for measuring steps (rho = 0.91 and 0.85, respectively). Furthermore, comparing the results from the Bland-Altman plots, the present results seem greatly better than those by DeGroote et al. (2020) for Xiaomi Mi Band 2 (LOA; 95% CI = -220.2; -2,462–2,021 vs. 1,011; -2,713–4,737, respectively), as well as those by Topalidis et al. (2021) for Xiaomi Mi Band 3 (LOA; 95% CI = -127.2; -2,575–2,320 vs. -4,050; -8,350–275, respectively).

Regarding studies carried out under structured conditions among adults, Stamm et al. (2018) and Tam & Cheung (2018) assessed the validity of Xiaomi Mi Band 2 walking on a treadmill at different predetermined speeds. Firstly, results from Stamm et al.'s (2018) study showed no differences between step count by the Xiaomi Mi Band 2 and the IMU sensor, obtaining the strongest agreement compared to the other studied wearables (i.e., Fitbit, Samsung, and Vidonn trackers). Regarding Tam & Cheung (2018), they also showed no differences between step count by the Xiaomi Mi Band 2 wristband and manual step count, also resulting as accurate as another high-cost wristband (i.e., Fitbit Charge HR). These results agree with those obtained by DeGroote et al. (2020) under free-living conditions highlighting the Xiaomi Mi Band wristbands as the best model of those studied from a wide price range. Besides, Hartung et al. (2020) assessed the concurrent validity of Xiaomi Mi Band 3 against observed manual steps under structured activity protocols among adults. They also showed excellent results under overground walking conditions (i.e., MAPE = 4.9%) but a considerable overestimation under the activities of daily life protocol (i.e., MAPE = 269.6%). However, the high overestimation of steps obtained during the daily life protocol may be due to the activities included in their structured protocol were focused only on arm movements (e.g., washing dishes or simulating eating). Therefore, the use of highly standardized, repetitive, and uniform movements probably does not exactly reflect actual free-living situations, causing measurement higher bias than in free-living conditions (Hartung et al., 2020).

Lastly, Pino-Ortega et al. (2021) carried out a similar study to the present one assessing the convergent validity of Xiaomi Mi Band 2, 3, and 4 against WIMU PRO inertial device, although among elderly adults walking continuously for five sets of four minutes. Their results also showed nearly perfect agreement with the criterion measure for the three generations of Xiaomi Mi Band as the present study (ICC, 95% CI = 0.99, 0.98–1.00 vs. 0.94-0.95, 0.90–0.97) and no statistical differences were found between devices. Therefore, considering the results obtained by these previous studies, the Xiaomi Mi Band 2, 3, and 4 seems accurate for step count under walking conditions, as in the present study. Nevertheless, although the results obtained by Hartung et al. (2020) and Pino-Ortega et al. (2021) under continuous walking conditions seem slightly better than those of the present study, the differences in settings are crucial. Specifically, it should be considered that while in the studies carried out in laboratory conditions participants were constrained to a predefined path with stable gait patterns, the present study was carried out under a greater variability of motor patterns, including a wide range of adolescents' daily life behaviors which probably increase the bias in measurement (Johnston et al., 2021). Therefore, due to the main goal of consumer-wearable activity trackers is to assess adolescents' daily PA levels or to use them as a motivating tool to increase adolescents' PA practice, the results obtained from free-living conditions are closer to reality, and therefore they are more meaningful and valuable (Duncan et al., 2018).

However, despite the good results obtained for the steps count, the findings of the present study showed poor agreement for light PA, MVPA and total PA minutes between Xiaomi Mi Band wristbands and the wGT3X-BT accelerometer. Regarding previous research about the validity of Xiaomi Mi Band devices for measuring adolescents' PA levels, only the study by Yang et al. (2019) was found assessing the validity of Xiaomi Mi Band 2. They also showed unacceptable validity as the present study for measuring MVPA (MAPE = 61.2% vs. 110.8%; LOA, 95%CI = 42.6, -56.1–

141.3; vs. -12.4, -93.2–68.4), MVPA based on the brisk pace walking score (MAPE = 61.2% vs. 50.1%; LOA, 95%CI = 42.6, -56.1–141.3; vs. 15.2, -13.2–43.6), and light PA levels (MAPE = 29.1% vs. 77.7%). However, regarding minutes in total PA, results by Yang et al. (2019) are greatly better than those of the present study (based on the slow-brisk pace walking score) if MAPE values are considered (MAPE = 8.0% vs. 69.6%), but considering Bland-Altman plots the results are similar (LOA, 95%CI = 21.4, -129.1–171.9 vs. 144.2, 51.1–237.3). Furthermore, the study by DeGroote et al. (2020) also investigated the concurrent validity of Xiaomi Mi Band 2 for measuring MVPA minutes, obtaining similar unacceptable results as the present study for minutes involved in MVPA (ICC, 95%CI = 0.15, -0.08-0.39 vs. 0.15, 0.00-0.37; MAPE = 293.29% vs. 110.8%), although it was carried out with adult population. However, comparing the previous MVPA results by DeGroote et al. (2020) with the present study but based on the brisk pace walking score, the present ones seem slightly better, ranging from poor (MAPE = 50.1%) to acceptable (ICC, 95%CI = 0.78, 0.18-0.92) validity. Meanwhile, both previous and present studies (only if minutes in brisk pace walking is considered) showed that Xiaomi Mi Band wristbands overestimated the time spent on MVPA compared to the accelerometer. Unfortunately, no previous studies have been found about the validity of the other Xiaomi Mi Band generations (i.e., 3, 4, and 5) to measure PA levels in any population or setting to compare the present results.

Nevertheless, these results are also consistent with other previous studies that found the consumer-wearable activity trackers from different brands and models valid to measure step count but not to measure PA at different intensities, even with high-cost trackers (e.g., DeGroote et al., 2020; Evenson et al., 2015; Feehan et al., 2018; Fuller et al., 2020; Voss et al., 2017). However, these discrepancies between the wristbands and research-grade accelerometers may arise not only for measurement bias but also from the specific algorithmic equation used (i.e., the cut-points used for classifying PA intensity; Ferguson et al., 2015). Firstly, regarding the research-grade accelerometer as the reference measure, although the best available literature has been considered in the present study, it is still no solid evidence-based consensus about methodological issues (e.g., the best MVPA cut-point among adolescents or device placement; Migueles et al., 2017). Furthermore, unfortunately, specific information regarding intensity-related PA cut-points used by Xiaomi Mi Band is not publicly available. Therefore, it is not possible to obtain substantial results about Xiaomi Mi Band wristbands validity for measuring PA levels, and the present results should be interpreted with caution since it is only based on assumptions made by researchers about intensity-related PA cut-points used by the wristbands (DeGroote et al., 2020; Ferguson et al., 2015). In any case, brands should provide more transparent information on their algorithms to classify PA at different intensities to obtain more robust results in future research.

Lastly, the secondary aim of this study was to assess the convergent validity of Xiaomi Mi Band wristbands for correctly classifying adolescents as meeting or not meeting the international PA recommendations. The findings of the present study showed that the Xiaomi Mi Band wristbands have good to excellent validity for correctly classifying adolescents as meeting or not meeting the recommendation of 10,000 steps per day (P = 0.89–0.95, k = 0.71–0.87) and the recommendation of 60 minutes of MVPA per day (based on the brisk pace walking score; P = 0.89–0.94, k = 0.69–0.83). Therefore, although Xiaomi Mi Band wristbands present low validity for assessing intensity-related PA outputs, these results are promising for public health. Specifically, they allow knowing if adolescents are achieving the minimum PA levels recommended and, therefore, its consequent health benefits (WHO, 2020). Moreover, these results can help physical education teachers or policymakers to set goals within Xiaomi Mi Band-based PA promotion programs among adolescents, establishing a minimum of 10,000 daily steps and/or 60 minutes of brisk pace walking per day to ensure the accomplishment with the international recommendations.

Strengths and limitations

An important strength of the present study was, to our knowledge, being the first one to examine the validity of four generations of Xiaomi Mi Band wristbands (i.e., 2, 3, 4, and 5) among adolescents for measuring different PA variables (i.e., steps, time spent on PA at different intensities, and thresholds for correctly classifying them as meeting or not meeting the accelerometer-measured steps and MVPA recommendation). Thus, it allows addressing an important gap in the scientific literature to date, due to most previous studies have been carried out with adults and older people, and/or only using the oldest Xiaomi Mi Band generations (i.e., Xiaomi Mi Band 2 and 3), and/or only assessing their validity for measuring steps. Furthermore, the evaluation in free-living conditions better reflects the validity of the wristband for measuring actual PA behavior of adolescents during their daily life (Duncan et al., 2018). Therefore, it provides more useful information for monitoring adolescents' PA levels during health promotion programs, PA surveillance studies, or for the provision of feedback in behavior change programs (Duncan et al., 2018; Michie et al., 2009; Strath & Rowley, 2018).

However, the present study is not without limitations. Firstly, a non-probability and relatively small sample has been used, limiting the obtained outcomes' generalizability to the particular studied setting (i.e., adolescents with similar characteristics and PA patterns). However, due to the Covid-19 pandemic, human, and material resource restrictions, a probability and larger sample could not be examined. Secondly, although Actigraph accelerometers have been highlighted as the most common and valid method for objectively assessing adolescents' PA levels during free-living conditions (e.g., Lee et al., 2015; Romanzini et al., 2014; Trost et al., 2011), today there is no solid evidence-based consensus about many methodological issues (Migueles et al., 2017). Therefore, it may contribute to the variability of Xiaomi Mi Band wristbands convergent validity results, although the best current evidence-based decisions were adopted in the present study (Migueles et al., 2017). Lastly, the third limitation is related to the fact that Xiaomi brand did not provide transparent information about their algorithms to classify PA at different intensities (i.e., cut-off points), which may bias the present results. Therefore, it would be necessary to obtain more robust and reliable results their validity in further research.

Conclusions

In conclusion, the Xiaomi Mi Band 2, 3, 4, and 5 wristbands presented good convergent validity with the ActiGraph wGT3X-BT accelerometer for measuring adolescents' steps. Furthermore, although convergent validity for the measurement of PA at different intensities was unacceptable, they accurately classified adolescents as meeting or not meeting the recommendation of 10,000 steps and the recommendation of 60 minutes of MVPA per day. This highlights the potential of Xiaomi Mi Band wristbands for monitoring adolescents' PA levels and obtaining accurate information about compliance with international PA recommendations, offering a feasible and affordable for most people alternative to the research-grade accelerometers. Furthermore, they could be used during health promotion programs to provide accurate feedback to adolescents (especially for steps output), as well as to set specific goals based on 10,000 daily steps and/or 60 minutes of brisk pace walking per day to ensure their accomplishment with the international recommendations. However, further research investigating their validity and reliability for measuring steps and PA among different populations and contexts is needed, especially if the Xiaomi brand provided transparent information about their algorithms to classify PA at different intensities.

Acknowledgements

The authors want to thank you to all the students and teachers who made this research possible. We also thank Aliisa Hatten for the English revision of the manuscript. This work is funded by FEDER/*Junta de Andalucía-Consejería de Economía y Conocimiento* (Reference: B-SEJ-029-UGR18) and by the Vice-Rector's Office for Research and Transfer of the University of Granada. Carolina Casado-Robles is also supported by a research grant from the Spanish Ministry of Universities [grant number: FPU16/03314].

Supplementary Files

Scan or click here to see Supplementary Files [Escanea el código o pincha aquí para ver los Archivos Suplementarios]

Supplementary Files

References

ActiGraph Corporation. (2021). *Activity monitors comparisons*. http://actigraphcorp.com

Alley, S., Schoeppe, S., Guertler, D., Jennings, C., Duncan, M. J., & Vandelanotte, C. (2016). Interest and preferences for using advanced physical activity tracking devices: results of a national cross-sectional survey. *BMJ open*, *6*(7), e011243. https://doi.org/10.1136/bmjopen-2016-011243

Arifin, W. (2018). A web-based sample size calculator for reliability studies. *Education in Medicine Journal*, *10*(3), 67–76. https://doi.org/10.21315/eimj2018.10.3.8

Atkinson, G., & Nevill, A. M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports medicine*, *26*(4), 217–238. https://doi.org/10.2165/00007256-199826040-00002

Bland, J., & Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, *1*(8476), 307–310.

Brickwood, K. J., Watson, G., O'Brien, J., & Williams, A. D. (2019). Consumer-Based Wearable Activity Trackers Increase Physical Activity Participation: Systematic Review and Meta-Analysis. *JMIR mHealth and uHealth*, *7*(4), e11819. https://doi.org/10.2196/11819

Brooke, H. L., Corder, K., Atkin, A. J., & van Sluijs, E. M. (2014). A systematic literature review with meta-analyses of within- and between-day differences in objectively measured physical activity in school-aged children. *Sports medicine,* *44*(10), 1427–1438. https://doi.org/10.1007/s40279-014-0215-5

Campos-Meirinhos, E. J., Mayorga-Vega, D., Casado-Robles, C., Guijarro-Romero, S., & Viciana, J. (2019). Are activity wristbands valid to estimate moderate-to-vigorous physical activity in adolescents during free-living conditions? *Motricidade*, *15*(S1), 18.

Chaput, J. P., Willumsen, J., Bull, F., Chou, R., Ekelund, U., Firth, J., Jago, R., Ortega, F. B., & Katzmarzyk, P. (2020). 2020 WHO guidelines on physical activity and sedentary behaviour for children and adolescents aged 5–17 years: summary of the evidence. *International Journal of Behavioral Nutrition and Physical Activity 17,* 141. https://doi.org/10.1186/s12966-020-01037-z

Cicchetti, D. (2001). The precision of reliability and validity estimates re-visited: Distinguishing between clinical and statistical significance of sample size requirements. *Journal of Clinical and Experimental Neuropsychology, 23*(5), 695–700.

Cohen J. (1992). A power primer. *Psychological bulletin*, *112*(1), 155–159. https://doi.org/10.1037//0033-2909.112.1.155

Cole, T. J., Bellizzi, M. C., Flegal, K. M., & Dietz, W. H. (2000). Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ (Clinical research ed.)*, *320*(7244), 1240–1243. https://doi.org/10.1136/bmj.320.7244.1240

Corder, K., Brage, S., & Ekelund, U. (2007). Accelerometers and pedometers: methodology and clinical application. *Current opinion in clinical nutrition and metabolic care*, *10*(5), 597–603. https://doi.org/10.1097/MCO.0b013e328285d883

Cosoli, G., Spinsante, S., & Scalise, L. (2020). Wrist-worn and chest-strap wearable devices: Systematic review on accuracy and metrological characteristics. *Measurement*, *159*, 107789. https://doi.org/10.1016/j.measurement.2020.107789

Degroote, L., Hamerlinck, G., Poels, K., Maher, C., Crombez, G., De Bourdeaudhuij, I., Vandendriessche, A., Curtis, R. G., & DeSmet, A. (2020). Low-Cost Consumer-Based Trackers to Measure Physical Activity and Sleep Duration Among Adults in Free-Living Conditions: Validation Study. *JMIR mHealth and uHealth*, *8*(5), e16674. https://doi.org/10.2196/16674

Dhurandhar, N. V., Schoeller, D., Brown, A. W., Heymsfield, S. B., Thomas, D., Sørensen, T. I., Speakman, J. R., Jeansonne, M., Allison, D. B., & Energy Balance Measurement Working Group (2015). Energy balance measurement: when something is not better than nothing. *International journal of obesity*, *39*(7), 1109–1113. https://doi.org/10.1038/ijo.2014.199

Duncan, M. J., Wunderlich, K., Zhao, Y., & Faulkner, G. (2018). Walk this way: validity evidence of iphone health application step count in laboratory and free-living conditions. *Journal of sports sciences*, *36*(15), 1695–1704. https://doi.org/10.1080/02640414.2017.1409855

Evenson, K. R., Catellier, D. J., Gill, K., Ondrak, K. S., & McMurray, R. G. (2008). Calibration of two objective measures of physical activity for children. *Journal of sports sciences*, *26*(14), 1557–1565. https://doi.org/10.1080/02640410802334196

Evenson, K. R., Goto, M. M., & Furberg, R. D. (2015). Systematic review of the validity and reliability of consumer-wearable activity trackers. *The international journal of behavioral nutrition and physical activity*, *12*, 159. https://doi.org/10.1186/s12966-015-0314-1

Feehan, L. M., Geldman, J., Sayre, E. C., Park, C., Ezzat, A. M., Yoo, J. Y., Hamilton, C. B., & Li, L. C. (2018). Accuracy of Fitbit Devices: Systematic Review and Narrative Syntheses of Quantitative Data. *JMIR mHealth and uHealth*, *6*(8), e10527. https://doi.org/10.2196/10527

Ferguson, T., Rowlands, A. V., Olds, T., & Maher, C. (2015). The validity of consumer-level, activity monitors in healthy adults worn in free-living conditions: a cross-sectional study. *The international journal of behavioral nutrition and physical activity*, *12*, 42. https://doi.org/10.1186/s12966-015-0201-9

Franssen, W., Franssen, G., Spaas, J., Solmi, F., & Eijnde, B. O. (2020). Can consumer wearable activity tracker-based interventions improve physical activity and cardiometabolic health in patients with chronic diseases? A systematic review and meta-analysis of randomised controlled trials. *The international journal of behavioral nutrition and physical activity*, *17*(1), 57. https://doi.org/10.1186/s12966-020-00955-2

Fuller, D., Colwell, E., Low, J., Orychock, K., Tobin, M. A., Simango, B., Buote, R., Van Heerden, D., Luan, H., Cullen, K., Slade, L., & Taylor, N. (2020). Reliability and Validity of Commercially Available Wearable Devices for Measuring Steps, Energy Expenditure, and Heart Rate: Systematic Review. *JMIR mHealth and uHealth*, *8*(9), e18694. https://doi.org/10.2196/18694

Gillis, L., Tomkinson, G., Olds, T., Moreira, C., Christie, C., Nigg, C., Cerin, E., Van Sluijs, E., Stratton, G., Janssen, I., Dorovolomo, J., Reilly, J. J., Mota, J., Zayed, K., Kawalski, K., Andersen, L. B., Carrizosa, M., Tremblay, M., Chia, M., Hamlin, M., … Van Mechelen, W. (2013). Research priorities for child and adolescent physical activity and sedentary behaviours: an international perspective using a twin-panel Delphi procedure. *The international journal of behavioral nutrition and physical activity, 10,* 112. https://doi.org/10.1186/1479-5868-10-112

Gorzelitz, J., Farber, C., Gangnon, R., & Cadmus-Bertram, L. (2020). Accuracy of Wearable Trackers for Measuring Moderate- to Vigorous-Intensity Physical Activity: A Systematic Review and Meta-Analysis, *Journal for the Measurement of Physical Behaviour, 3*(4), 346-357. https://doi.org/10.1123/jmpb.2019-0072

Guthold, R., Stevens, G. A., Riley, L. M., & Bull, F. C. (2020). Global trends in insufficient physical activity among adolescents: a pooled analysis of 298 population-based surveys with 1.6 million participants. *The Lancet Child & Adolescent Health, 4*(1), 23-35. https://doi.org/10.1016/S2352-4642(19)30323-2

Hao, Y., Ma, X. K., Zhu, Z., & Cao, Z. B. (2021). Validity of Wrist-Wearable Activity Devices for Estimating Physical Activity in Adolescents: Comparative Study. *JMIR mHealth and uHealth, 9*(1), e18320. https://doi.org/10.2196/18320

Hartung, V., Sarshar, M., Karle, V., Shammas, L., Rashid, A., Roullier, P., Eilers, C., Mäurer, M., Flachenecker, P., Pfeifer, K., & Tallner, A. (2020). Validity of Consumer Activity Monitors and an Algorithm Using Smartphone Data for Measuring Steps during Different Activity Types. *International journal of environmental research and public health, 17*(24), 9314. https://doi.org/10.3390/ijerph17249314

Henriksen, A., Haugen Mikalsen, M., Woldaregay, A. Z., Muzny, M., Hartvigsen, G., Hopstock, L. A., & Grimsgaard, S. (2018). Using Fitness Trackers and Smartwatches to Measure Physical Activity in Research: Analysis of Consumer Wrist-Worn Wearables. *Journal of medical Internet research, 20*(3), e110. https://doi.org/10.2196/jmir.9157

Hernaez, R. (2015). Reliability and agreement studies: A guide for clinical investigators. *Gut, 64*(7), 1018–1027. https://doi.org/10.1136/gutjnl-2014-308619

Howie, E. K., & Straker, L. M. (2016). Rates of attrition, non-compliance and missingness in randomized controlled trials of child physical activity interventions using accelerometers: A brief methodological review. *Journal of* *science and medicine in sport*, *19*(10), 830–836. https://doi.org/10.1016/j.jsams.2015.12.520

International Data Corporation. (2020). *Shipments of Wearable Devices Leap to 125 Million Units, Up 35.1% in the Third Quarter, According to IDC*. https://www.idc.com/getdoc.jsp?containerId=prUS47067820.

Jia, Y., Wang, W., Wen, D., Liang, L., Gao, L., & Lei, J. (2018). Perceived user preferences and usability evaluation of mainstream wearable devices for health monitoring. *PeerJ*, *6*, e5350. https://doi.org/10.7717/peerj.5350

Johnston, W., Judice, P. B., Molina García, P., Mühlen, J. M., Lykke Skovgaard, E., Stang, J., Schumann, M., Cheng, S., Bloch, W., Brønd, J. C., Ekelund, U., Grøntved, A., Caulfield, B., Ortega, F. B., & Sardinha, L. B. (2021). Recommendations for determining the validity of consumer wearable and smartphone step count: expert statement and checklist of the INTERLIVE network. *British Journal of Sports Medicine*, *55*(14), 780 LP – 793. https://doi.org/10.1136/bjsports-2020-103147

Koo, T., & Li, M. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Kottner, J., Audigé, L., Brorson, S., Donner, A., Gajewski, B. J., Hróbjartsson, A., Roberts, C., Shoukri, M., & Streiner, D. L. (2011). Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *Journal of clinical epidemiology*, *64*(1), 96–106. https://doi.org/10.1016/j.jclinepi.2010.03.002

Lee, J. A., Williams, S. M., Brown, D. D., & Laurson, K. R. (2015). Concurrent validation of the Actigraph gt3x+, Polar Active accelerometer, Omron HJ-720 and Yamax Digiwalker SW-701 pedometer step counts in lab-based and free-living settings. *Journal of sports sciences*, *33*(10), 991–1000. https://doi.org/10.1080/02640414.2014.981848

Maher, C., Ryan, J., Ambrosi, C., & Edney, S. (2017). Users' experiences of wearable activity trackers: a cross-sectional study. *BMC public health*, *17*(1), 880. https://doi.org/10.1186/s12889-017-4888-1

Martínez-Gómez, D., Martínez-De-Haro, V., Del-Campo, J., Zapatera, B., Welk, G. J., Villagra, A., Marcos, A., & Veiga, O. L. (2009). Validez de cuatro cuestionarios para valorar la actividad física en adolescentes españoles [Validity of four questionnaires to assess physical activity in Spanish adolescents]. *Gaceta sanitaria*, *23*(6), 512–517. https://doi.org/10.1016/j.gaceta.2009.02.013

Mayorga-Vega, D., Casado-Robles, C., López-Fernández, I., & Viciana, J. (2021). A comparison of the utility of different step-indices to translate the physical activity recommendation in adolescents. *Journal of Sports Sciences*, *39*(4), 469–479. https://doi.org/10.1080/02640414.2020.1826667

Mercer, K., Giangregorio, L., Schneider, E., Chilana, P., Li, M., & Grindrod, K. (2016). Acceptance of Commercially Available Wearable Activity Trackers Among Adults Aged Over 50 and With Chronic Illness: A Mixed-Methods Evaluation. *JMIR mHealth and uHealth*, *4*(1), e7. https://doi.org/10.2196/mhealth.4225

Metcalf, B., Henley, W., & Wilkin, T. (2012). Effectiveness of intervention on physical activity of children: systematic review and meta-analysis of controlled trials with objectively measured outcomes (EarlyBird 54). *BMJ*, *345*, e5888. https://doi.org/10.1136/bmj.e5888

Michie, S., Abraham, C., Whittington, C., McAteer, J., & Gupta, S. (2009). Effective techniques in healthy eating and physical activity interventions: a meta-regression. *Health psychology : official journal of the Division of Health Psychology, American Psychological Association*, *28*(6), 690–701. https://doi.org/10.1037/a0016136

Michie, S., Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., Eccles, M. P., Cane, J., & Wood, C. E. (2013). The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Annals of behavioral medicine, 46*(1), 81–95. https://doi.org/10.1007/s12160-013-9486-6

Migueles, J. H., Cadenas-Sanchez, C., Ekelund, U., Delisle Nyström, C., Mora-Gonzalez, J., Löf, M., Labayen, I., Ruiz, J. R., & Ortega, F. B. (2017).

Accelerometer Data Collection and Processing Criteria to Assess Physical Activity and Other Outcomes: A Systematic Review and Practical Considerations. *Sports medicine*, *47*(9), 1821–1845. https://doi.org/10.1007/s40279-017-0716-0

Nunnally, J. (1978). *Psychometric theory*. McGraw-Hill.

Oliver, M., Badland, H. M., Schofield, G. M., & Shepherd, J. (2011). Identification of accelerometer nonwear time and sedentary behavior. *Research quarterly for exercise and sport*, *82*(4), 779–783. https://doi.org/10.1080/02701367.2011.10599814

Parra Saldías, M., Mayorga-Vega, D., López-Fernández, I., & Viciana, J. (2018). How many daily steps are really enough for adolescents? A cross-validation study. *Retos*, *33*, 241–246. https://doi.org/10.47197/retos.v0i33.55504

Patel, M. S., Asch, D. A., & Volpp, K. G. (2015). Wearable devices as facilitators, not drivers, of health behavior change. *JAMA*, *313*(5), 459–460. https://doi.org/10.1001/jama.2014.14781

Pino-Ortega, J., Gómez-Carmona, C. D., & Rico-González, M. (2021). Accuracy of Xiaomi Mi Band 2.0, 3.0 and 4.0 to measure step count and distance for physical activity and healthcare in adults over 65 years. *Gait & posture*, *87*, 6–10. https://doi.org/10.1016/j.gaitpost.2021.04.015

Poitras, V. J., Gray, C. E., Borghese, M. M., Carson, V., Chaput, J. P., Janssen, I., Katzmarzyk, P. T., Pate, R. R., Connor Gorber, S., Kho, M. E., Sampson, M., & Tremblay, M. S. (2016). Systematic review of the relationships between objectively measured physical activity and health indicators in school-aged children and youth. *Applied physiology, nutrition, and metabolism*, *41*(6), S197–S239. https://doi.org/10.1139/apnm-2015-0663

Rich, E., & Miah, A. (2016). Mobile, wearable and ingestible health technologies: towards a critical research agenda. Health sociology review, 26(1), 84-97. https://doi.org/10.1080/14461242.2016.1211486

Romanzini, M., Petroski, E. L., Ohara, D., Dourado, A. C., & Reichert, F. F. (2014). Calibration of ActiGraph GT3X, Actical and RT3 accelerometers in adolescents. *European journal of sport science, 14*(1), 91–99. https://doi.org/10.1080/17461391.2012.732614

Ruiz, D. C., & Goransson, A. (2015). *Professional Android Wearables*. John Wiley & Sons.

Shephard, R., & Tudor-Locke, C. (Eds.). (2016). *The objective monitoring of physical activity: Contributions of accelerometry to epidemiology, exercise science and rehabilitation*. Springer.

Šimůnek, A., Dygrýn, J., Jakubec, L., Neuls, F., Frömel, K., & Welk, G. J. (2019). Validity of Garmin Vívofit 1 and Garmin Vívofit 3 for School-Based Physical Activity Monitoring. *Pediatric exercise science, 31*(1), 130–136. https://doi.org/10.1123/pes.2018-0019

Stamm, A., & Hartanto, R. (2018). Feature extraction from mems accelerometer and motion tracking measurements in comparison with smart bands during running. *Proceedings, 2*(6), 197-202. https://doi.org/10.3390/proceedings2060197

Stewart, A., Marfell-Jones, M., Olds, T., & De Ridder, J. (2011). *International standards for anthropometric assessment*. International Society for the Advancement of Kinanthropometry.

Strath, S. J., & Rowley, T. W. (2018). Wearables for Promoting Physical Activity. *Clinical chemistry, 64*(1), 53–63. https://doi.org/10.1373/clinchem.2017.272369

Tam, K. M., & Cheung, S. Y. (2018). Validation of Electronic Activity Monitor Devices During Treadmill Walking. *Telemedicine journal and e-health : the official journal of the American Telemedicine Association, 24*(10), 782–789. https://doi.org/10.1089/tmj.2017.0263

Topalidis, P., Florea, C., Eigl, E. S., Kurapov, A., Leon, C., & Schabus, M. (2021). Evaluation of a Low-Cost Commercial Actigraph and Its Potential Use in Detecting Cultural Variations in Physical Activity and Sleep. *Sensors, 21*(11), 3774. https://doi.org/10.3390/s21113774

Trost, S. G., Loprinzi, P. D., Moore, R., & Pfeiffer, K. A. (2011). Comparison of accelerometer cut points for predicting activity intensity in youth. *Medicine and* *science in sports and exercise*, *43*(7), 1360–1368. https://doi.org/10.1249/MSS.0b013e318206476e

Tudor-Locke, C., Barreira, T. V., & Schuna, J. M., Jr (2015). Comparison of step outputs for waist and wrist accelerometer attachment sites. *Medicine and science in sports and exercise*, *47*(4), 839–842. https://doi.org/10.1249/MSS.0000000000000476

Tudor-Locke, C., Craig, C. L., Beets, M. W., Belton, S., Cardon, G. M., Duncan, S., Hatano, Y., Lubans, D. R., Olds, T. S., Raustorp, A., Rowe, D. A., Spence, J. C., Tanaka, S., & Blair, S. N. (2011). How many steps/day are enough? for children and adolescents. *The international journal of behavioral nutrition and physical activity*, *8*, 78. https://doi.org/10.1186/1479-5868-8-78

Van Hecke, L., Loyen, A., Verloigne, M., van der Ploeg, H. P., Lakerveld, J., Brug, J., De Bourdeaudhuij, I., Ekelund, U., Donnelly, A., Hendriksen, I., Deforche, B., & DEDIPAC consortium (2016). Variation in population levels of physical activity in European children and adolescents according to cross-European studies: a systematic literature review within DEDIPAC. *The international journal of behavioral nutrition and physical activity*, *13*, 70. https://doi.org/10.1186/s12966-016-0396-4

Voss, C., Gardner, R. F., Dean, P. H., & Harris, K. C. (2017). Validity of Commercial Activity Trackers in Children With Congenital Heart Disease. *The Canadian journal of cardiology*, *33*(6), 799–805.

Willmott, C., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, *30*(1), 79–82. https://doi.org/10.3354/cr030079

World Health Organization. (2018). *Physical Activity Factsheets for the 28 European Union Member States of the Who European Region*. World Health Organization

World Health Organization. (2020). *WHO guidelines on physical activity and sedentary behaviour*. World Health Organization.

Yang, X., Jago, R., Zhang, Q., Wang, Y. Y., Zhang, J., & Zhao, W. H. (2019). Validity and Reliability of Wristband Activity Monitor in Free-living Children Aged 10-17 Years. *Biomedical and environmental sciences*, *32*(11), 812–822.

Paper XI

ARE ACTIVITY WRISTBANDS AND MOBILE APPS VALID FOR ASSESSING PHYSICAL ACTIVITY IN ADOLESCENTS?

Viciana, J., Casado-Robles, C., Guijarro-Romero, S., & Mayorga-Vega, D.

Draft

Paper XI

ABSTRACT

Objective: To examine the validity of three activity wristbands (i.e., Samsung Watch Active 2, Apple Watch Series 5, and Xiaomi Mi Band 5) and six mobile apps (Pedometer and Pacer for Android and iPhone mobiles, and Google Fit for Android, and Apple Health for iPhone mobiles) for estimating students' steps and physical activity (PA) under free-living conditions.

Methods: A sample of 56 (48.2% females) and 51 students (49.0% females) participated in Study 1 and 2, respectively, all of them aged between 12-18 years old. Study 1: Students performed a 200-meter course in four different conditions while wearing three wristbands and two mobile phones. Step counting through a video record was used as the reference standard of steps. Study 2: Students wore the three wristbands during the waking time of one day. w-GT3X-BT accelerometers were used as the reference standard of daily steps and PA.

Results: Study 1: All the wearables presented excellent validity for assessing steps under structured free-living conditions (i.e., MAPE < 5%), although their validity were between poor to excellent based on ICC values (ICC, 95% IC = 0.76-1.00, 0.56-1.00). Study 2: Xiaomi Mi Band 5 and Samsung Watch Active 2 wristbands presented acceptable (MAPE = 9.4-11.4%) to excellent (ICC, 95% CI = 0.95, 0.91-0.97) validity for assessing steps under unstructured free-living conditions. However, Apple Watch Series 5 presented questionable to excellent validity (MAPE = 18.0%; ICC, 95% CI = 0.89, 0.69-0.95). Regarding MVPA and total PA, the three wristbands presented poor validity, although particularly for MVPA, questionable to excellent validity for ICC values assessed by the Apple Watch Series 5 was obtained (ICC, 95% IC = 0.86). Moreover, wristbands have excellent validity for correctly classifying adolescents as meeting or not the recommended 10,000 steps/day (P = 0.94–0.96), and only Apple Watch Series 5 for the recommended 60 minutes of MVPA/day (P = 0.96).

Conclusions: Wearables checked in this study have shown adequate validity results to measure steps in structured and unstructured free-living conditions. However, in unstructured free-living conditions and for assessing MVPA, only Apple Watch Series 5 reported valid results, and none of them was valid for measuring total PA. Therefore, depending on the user's/researcher's aim, one wristband could be more adequate, mainly because of its valid measurements and cost.

KEYWORDS

Consumer-wearable activity tracker. Video record. Criterion validity. Accelerometer. Convergent validity. Steps. Young people.

Introduction

It is well known that due to the high prevalence of physical inactivity in youth, and particularly in adolescents (Word Health Organization, WHO, 2020; Guthold et al., 2020) and their associated consequences in adulthood (i.e., obesity/adiposity, cardiometabolic biomarkers/metabolic syndrome, low levels of physical fitness, and weak bone health; Poitras et al., 2016), a current public health priority is to improve the ratio of adolescents that meet the daily physical activity (PA) recommendations of 60 minutes of moderate-to-vigorous PA (MVPA) in aerobic activities, such as brisk walking or running (WHO, 2020) or 10,000 steps per day (Mayorga-Vega et al., 2019; Tudor-Locke et al., 2011).

Additionally, interpreting the intensity level of daily activities and its relationship with MVPA is not an easy ability for adolescents (Crossley et al., 2019). In order to help the population in this task of regulating the daily PA and its quantification, the consumer-wearable activity trackers have raised strongly since they provide a display for self-monitoring of the users' PA (Strath & Rowley, 2018). Particularly, due to several advantages, two types of wearable activity trackers can be stand out as the most popular: (a) PA wristbands (including smartwatches) have shown to be one of the most valued and used type of wearable because of their characteristics, such as an attractive display, real-time feedback, low weight and price (mainly wristbands), sedentary break audible warnings, or goals alerts (Fairclough et al., 2016); and (b) mobile apps are also widely spread among adolescents and let them registering many PA parameters. Therefore, activity wristbands and mobile apps could be useful, accessible, and feasible devices to objectively assess and promote adolescents' daily PA (Da Silva et al., 2015). Moreover, these two wearable activity trackers (i.e., wristbands and mobile apps) are widely extended all over the world, and the number of users is continuous and exponentially increasing every year, reaching half a billion in 2020 with greater expectations in the future (Vailshery, 2021).

In regard to the wearable measurement, consumers (i.e., researchers and populations of different ages) need to know to what extent is accurate and valid the information regarding the PA they received from a particular device. For instance, educational researchers and Physical Education teachers in the school setting are interested in applying wearables in Physical Education-based PA promotion programs, due to the effects on the students' daily PA is a priority in the educational system of most of the countries all over the world (OECD, 2019), and consequently, valid wearables could help them to achieve this educational standard with reduced costs. However, researchers need to know if they can use activity wearables only as a strategy to increase habitual PA (independent variable) or also as a dependent variable (devices with a high quality of measurement). Therefore, criterion-related validity and convergent validity are perhaps two of the more important concepts related to the measurement of a particular wearable wristband in order to use them confidently. Criterion-related validity is referred to the accuracy of a measurement related to a 'golden standard' (i.e., a particular measurement of a construct without or almost free of error) (Bossuyt et al., 2015), while convergent validity is referred to the extent to which two or more measures capture the same phenomenon/construct.

In regard to PA intensity, the doubly labeled water assessment of energy expenditure is the golden standard (Westerterp, 2017), but oxygen uptake measurement (in a laboratory or with a portable indirect calorimetry system) substitutes this expensive method, and it is considered as 'reference standard' for assessing MVPA. Nevertheless, because of viability motives (e.g., small, light, comfortable, and reliable and valid; Migueles et al., 2017), accelerometers have been used as the most appropriate alternative in free-living conditions, relegating the use of the calorimetry system to laboratory/controlled activities (Westerterp, 2013). In fact, ActiGraph accelerometers have been the most widely used method in population and calibration studies with adolescents (Romanzini et al., 2014), as well as have been used as reference standard in most of the previous existing studies (e.g., Cosoli et al., 2020; Mayorga-Vega et al., 2019; Migueles et al., 2017; or Voss et al., 2017). In regard to the number of steps per day in free-living conditions, the accelerometer is also considered as the reference standard in most previous studies (Westerterp, 2013, 2017). Concerning structured or controlled settings, the video-taped and manual counter of the number of steps in a particular circuit or activity has been considered as the reference standard (El-Amrawy & Nounou, 2015; Stamm & Hartanto, 2018). Since it was determined that accelerometer is a reference standard measurement for PA intensity, the convergence of measurements provided by other activity wearables (e.g., wristbands and mobile apps) with it will provide new possibilities and devices to use in those contexts beyond the laboratory, such as Physical Education setting for instance, where validity and feasibility must coexist which each other (O'Neill et al., 2017).

Different cross-sectional studies with adults (e.g., Redenius et al., 2019; Toth et al., 2018) and synthesis studies (Brickwood et al., 2019; Gal et al., 2018) have been published till now with the aim of showing the validity of activity trackers and mobile apps. In regard to adolescents, the number of publications drops considerably, although there are still some publications regarding their validity (Cosoli et al., 2020; Gorzelitz, et al., 2020). Overall results of these above-mentioned publications have shown a great variety of protocols for validation, devices, and methodological characteristics, which provide irregular results barely comparable to each other. Nevertheless, depending on the context to be applied and the required accuracy of the PA parameters, different devices (i.e., wristbands and mobile apps) could be recommended. Moreover, particular measurements, such as energy expenditure or PA levels, are more complex constructs, and their measurements present significant errors when assessed by activity trackers and mobile apps. For instance, Adamakis (2020) obtained large errors in all monitors and apps used (SenseWear Armband Pro 3, Garmin Forerunner 310XT, Garmin Vivoft, Medisana Vifit, and smartphones running the Runkeeper, Runtastic, Sports Tracker, Pedometer, Accupedo, Pedometer, and Pedometer 2.0 apps), underestimating the PA energy expenditure from 13.16% (in the case of walking and the Runtastic application) to 37.46% (in the case of running and Vifit). On the contrary, the number of steps is a simpler measurement that can be assessed by mobile apps and activity trackers with a moderate correlation with the reference standard, being valid measurements (with different accuracy depending on the period of time measured and the activity tracker used). Although these outputs measured by activity wearables tend to have a lower validity for assessing intensity-related PA than steps among adults (Cosoli et al., 2020), among adolescents, the evidence is currently insufficient (Voss et al., 2017). For instance, Mayorga-Vega et al. (2019) already examined by accelerometry the cadence-based steps translation of the MVPA recommendation in adolescents, but they did not use wristbands or mobile apps in order to measure this variable (i.e., steps). Moreover, due to all the components of the devices (i.e., wristbands and smartphones) for the estimate of PA levels and steps are in continuous evolution by companies of activity wearables, to analyze new models of activity trackers and mobile apps is recommendable and desirable. Finally, Apple devices validity studies have not found, or were scarce and centered in other parameters such as heart rate or energy expenditure, or measuring steps in laboratory setting, as some previous synthesis studies have concluded (Fuller et al., 2020; Gorzelitz et al., 2020; LaMunion et al., 2020). Only Fokkema et al. (2017) studied the validity of the Apple Watch Sport in healthy adults and in structured free-living conditions, and they found that it was accurate in measurement steps, mainly when the speed walking was higher than slow. However, they did not study an adolescent's sample, an unstructured free-living condition, the measurement of PA intensities, nor the adequacy of its measurement to PA daily recommendations.

Consequently, the overall purpose of the present study was to examine the validity of the mobile apps Pedometer, Pacer, Google Fit for Androids mobiles (Samsung); Pedometer, Pacer, and Apple Health for Apple mobiles (iPhone); and the wristbands Samsung Watch Active 2, Apple Watch Series 5, and Xiaomi Mi Band 5 for estimating PA in adolescents during free-living conditions. Specifically, the purposes of this study were twofold: *Study 1:* to compare the criterion-related validity of the steps assessed by the nine wearables under structured free-living conditions. *Study 2:* to compare the convergent validity of the daily steps, total PA, and MVPA scores assessed by the three wristbands under unstructured free-living conditions.

Method

Participants

The present study is reported according to the GRRAS guidelines (Kottner et al., 2011). The protocol of the present study conforms to the Declaration of Helsinki statements (64th WMA, Brazil, October 2013) and it was approved by the Ethical Committee for Human Studies at the University of Granada. Then, the principals of two high schools chosen by convenience were contacted for the study 1 and 2. They were informed about the project, and permission to conduct the study was requested. After the approval of the school was obtained, all the adolescents and their legal guardians were fully informed about the features of the project. Adolescents' signed written informed assents and their legal guardians' signed written informed consents were obtained before taking part in the study. School A was used for study 1 (i.e., validity under structured free-living conditions) and school B for study 2 (i.e., validity under unstructured free-living conditions).

The present validity study followed a cross-sectional design. All the adolescents (i.e., 70 and 75 for the school A and B, respectively) enrolled in the selected schools (i.e., 12–16 years old) were invited to participate in the present study. The high schools were located in an urban area of the city center of Granada. For each study, the following inclusion criteria were considered: (a) being enrolled in the selected high schools (i.e., in the 7^{th} to 12^{h} grade); (b) being free of any health disorder that would make them unable to engage in PA normally; (c) presenting the corresponding signed written informed assents of the adolescents, and (d) presenting the corresponding signed written informed consents of their legal guardians. The following exclusion criteria were considered: (a) not having completed and valid data from the three activity wristbands (i.e., Samsung Galaxy Watch Active 2, Apple Watch Series 5, or Xiaomi Mi Band 5) or from the six mobile apps (study 1) or from the three activity wristbands (study 2), and (b) not having completed and valid data from the accelerometer (study 2)/the video-based step counts (study 1).

A priori sample size calculation was estimated with the Arifin´s web-based sample size calculator (Arifin, 2018). Based on steps values, parameters were set as follows: Study 1 and 2: ICC, $\rho_0 = 0.70$ (Nunnally, 1978); $\rho_1 = 0.86$ (Voss et al., 2017), $\alpha = 0.05$, $1 - \beta = 0.80$, $k = 2$, dropout = 10% (Rowlands et al., 2018). Study 2: Kappa, $k_0 = 0.40$ (Cicchetti, 2001); $k_1 = 0.85$ (Mayorga-Vega et al., 2021), $p = 0.25$ (Guthold et al., 2020), $\alpha = 0.05$, $1 - \beta = 0.80$, $k = 2$, dropout = 10% (Rowlands et al., 2018). A final sample size of at least 49 adolescents (minimum initial sample size equal to 55) was used in the final analysis. In addition to exceed the minimum required sample size, for each study sampling obtaining a sample balanced by grade and gender was aimed.

Measures

Demographic characteristics

Adolescents' gender (males/females), age (in years), and non-dominant hand (left/right) information was self-reported.

Anthropometric

Adolescents' body mass (kg) and height (cm) were first measured following the International Standards for Anthropometric Assessment (Stewart et al., 2011). Adolescents' body mass and height were measured in shorts, T-shirts, and barefoot. For the body mass measure, the adolescents stood in the centre of the scale (Seca, Ltd., Hamburg, Germany; accuracy = 0.1 kg) without support and with the weight distributed evenly on both feet. For the body height assessment, adolescents stood with the feet together with the heels, buttocks and upper part of the back touching the stadiometer (Holtain Ltd., Crymmych, Pembs, United Kingdom; accuracy = 0.1 cm), and with the head placed in the Frankfort plane. Each measurement was performed twice and the mean was recorded (Stewart et al., 2011). Then, the body mass index was calculated as body mass divided by body height squared (kg/m^2). Finally, teenagers' body weight status was categorized by gender- and age-adjusted body mass index thresholds as overweight/obesity or non-overweight/obesity (Cole et al., 2000). Body mass index and body weight status scores have shown high evidence supporting validity among adolescents (Cole et al., 2000).

Activity wristbands and mobile apps

Adolescents' PA levels and steps were estimated by the three activity wristbands [i.e., Xiaomi Mi Band 5 (Xiaomi, Pekin, China); Samsung Galaxy Watch Active 2 (Samsung, Seul, South Korea); and Apple Watch Series 5 (Apple Park, California, USA)]; and from six mobile apps [i.e., Pedometer (ITO Technologies) and Pedometer Pacer Health for both Android and iOS; and Google Fit app for Android (Samsung Galaxy S20+), and Apple Health app for iOS (iPhone 11 Pro Max)]. Regarding the number of activity wristbands, it was considered that three devices were a feasible number of them that did not interfere with the adolescents' natural arms swing during walking, brisk walking, or running (study 1), or daily physical activities (study 2). Mobile phones (for study 1) were allocated in two belt bags that change sides in each participant, which do not interfere either the adolescents' movements during the test. The total mass of the three wristbands was not high (83 grams). According to the user manual of each device brand, the activity wristbands were fit snugly on the top of adolescents' wrist, close and above the wrist bone. As regards the particular chosen activity wristbands, the criteria were to study the most worldwide used display-based wristbands, choosing the more advanced model (in that moment) of activity wristband considering a low cost of price (i.e., Xiaomi Mi Band 5 ≈ 25€) and smartwatches from the two main brands [Samsung belonging to Android (Samsung Galaxy Watch Active 2 that uses the Android Wear System) and Apple belonging to iOS (Apple Watch Series 5 that uses the WatchOS Apple System)] (based on the International Data Corporation's Worldwide Quarterly Wearable Device Tracker reports from 2021, and Henriksen et al. 2018). As regards the mobile apps [i.e., Pedometer (ITO Technologies) and Pedometer Pacer Health for Android and iOS; and Google Fit app for Android, and Apple Health app for iOS] were selected after revising the literature and examining the criteria followed previously [e.g., the following criteria both for Android and iOS: a) the most popular and used ones (due to the number of downloads and their users' ratings); b) free download apps; and c) the own apps of the corresponding mobile phones (i.e., Samsung Google Fit for Android and Apple Health for iOS].

According to the user manual of each device brand, activity wristbands were adjusted, in a random order, on the adolescent's wrist of the non-dominant hand. The three chosen devices are characterized to be small and lightweight wristbands (Xiaomi Mi Band 5: 22 g, 1.5 x 1.6 x 4.0 cm; Samsung Galaxy Watch Active 2 2: 30 g, 1.09 x 4.4 x 4.4 cm; Apple Watch Series 5: 30.8 g, 1.07 x 4.0 x 4.4 cm), based in tri-axial built-in accelerometers. Each wristband and mobile phone has its own algorithmic equation to estimate the daily steps taken and the minutes engaged in each specific intensity-related PA. The activity wristbands data and mobile phones apps were recorded immediately from the screen (for study 1), and were also synchronized via Bluetooth for the three wristbands to their specific applications in order to download and store data (for study 2) (Xiaomi Mi Band 5: Mi Fit version 5.3.2 for Android; Samsung Galaxy Watch Active 2: Samsung Health version 6.19 for Android; Apple Watch Series 5: Apple Health version 7.6.2 for iOS).

Regarding the data scoring, steps (number) were registered as directly reported in the devices. For the intensity-related PA (minutes), scores were calculated as follows. For the wristband Xiaomi Mi Band 5, the variables "slow walking" and "brisk walking" were calculated by summing the total time spent on all the bouts of "slow walking" and "brisk walking," respectively. The variable MVPA was calculated by summing the total time spent on all the bouts of "moderate activity" and "vigorous activity". Besides the default variable "total activity", the variable "walking" was also calculated by summing the total time spent on all the bouts of "slow walking" and "brisk walking". For the smartwatch Samsung Galaxy Watch Active 2, the variable total PA was calculated considering the default option "total active minutes". And for the smartwatch Apple Watch Series 5, the default option "exercise minutes" was considered as the variable MVPA.

Video-based steps count

Adolescents' reference standard of steps under structured free-living conditions was determined by step counting through a slow-motion video record. Adolescents were asked to perform a 200-meter course in four different conditions. Marking with cones and lines in the surface, the 200-meter course was performed in the school a non-slippery sport court with an oval shape with no strong turns. A digital video camera (Go Pro Hero 7, California, USA) with a tripod was situated in the middle of the sports court to easily record the participants´ lower limbs during the entire course from the sagittal plane turning around the camera. When the adolescent was in the starting line, the steps count from the activity wristbands were recorded. Then, they were requested not to move until they started walking/running. They also were asked to always start the course with the contralateral leg where the activity wristbands were wearing. For calculating the speed and steps cadence of each condition, when the adolescent started walking/running the manual stop watched, using the chronometer app of the Samsung Galaxy Watch Active 2, was activated and then it was stopped immediately after he/she crossed the finish line. Adolescents were requested to stop immediately after the finish line, so a cone was situated five meters before to remind them. Then, the steps count from the activity wristbands were recorded again. Adolescents performed the following four conditions: 1) Slow pace walking; 2) Normal pace walking (also known as self-pace walking); 3) Brisk pace walking; and 4) Jogging (slow running). Finally, the reference standard steps count for each adolescent in each condition were performed by two researchers independently through the slow-motion video record projected on a 15" screen. When disagreement occurred (8.3%), a third researcher evaluated it.

Accelerometer

Adolescents' reference standard of daily steps and intensity-related PA under unstructured free-living conditions were determined by w-GT3X-BT accelerometers (ActiGraph, LLC, Pensacola, FL, USA). The ActiGraph model w-GT3X-BT is a small (4.6 x 3.3 x 1.5 cm), lightweight (19 g), tri-axial accelerometer. Accelerometers were adjusted on the adolescents' right hips. Initializing, downloading, wear time validation and scoring were performed using the ActiLife Lifestyle Monitoring System Software version 6.13.3 (ActiGraph, LLC, Pensacola, FL, USA), and were initialized with a sample ratio of 30 Hz (Evenson et al., 2008; Trost et al., 2011). Data download was carried out with 15-second *epochs* (Evenson et al., 2008). Valid wear time was set as equal to or higher than 600 minutes per day (Migueles et al., 2017), with non-wear periods set as 60 minutes or more of consecutive zero-count *epochs* with up to two minutes spike tolerance (Oliver et al., 2011).

Regarding the data scoring, steps (number) were assessed by within-instrument processing of the number of cycles in the accelerometer signal or *cycle counts*. The time (minutes) engaged in light PA (101-2295 *counts*/min), MVPA (≥ 2,296 *counts*/ min), and total PA (≥ 101 *counts*/min) was calculated by the Evenson's thresholds (Evenson et al., 2008). According to the cross-validation study performed by Trost et al. (Trost et al., 2011), these cut-off points have demonstrated the best evidence supporting score validity for assessing intensity-related PA with one-second *epochs* among adolescents. Finally, adolescents' steps and MVPA were dichotomized as achieving or not achieving the daily recommendation of at least 10,000 steps and 60 minutes of MVPA, respectively (Mayorga-Vega et al., 2021; World Health Organization, 2020). ActiGraph accelerometer scores have shown high evidence supporting validity for assessing steps and intensity-related PA among adolescents (Hickey et al., 2016; Lee et al., 2015; Romanzini et al., 2014).

Procedure

Study 1 and 2: Data collection was carried out by the same researchers, instruments, and protocols. Firstly, adolescents' demographic characteristics, anthropometric, and self-reported habitual PA levels were registered. The three activity wristbands were adjusted so they could not move (i.e., avoiding over-tightening and clearance), in a random order, on the adolescents' wrist of the non-dominant hand. All devices were blocked in order to prevent participants from manipulating all functions and options they had, and to avoid influence in their habitual behavior.

Study 1: Evaluations were carried out during extracurricular activity time from Monday to Friday, and then data were downloaded and batteries charged during the morning. Due to the limitations of instruments, facilities, and time, an average of two or three adolescents per hour were evaluated one by one during each evaluation session. Apart from the three wristbands, the two mobile phones were inside two belt bags allocated on each adolescent's hip, changing sides for each participant. Adolescents were instructed to walk/run the 200-meter course in the four conditions specified above, at a continuous speed, and with a natural arm swing. Before starting, a demonstration in order to guide each participant was performed. During the five-minute rest between conditions, steps count data from the three wristbands and the mobile phones apps were recorded.

Study 2: The three activity wristbands and the accelerometer were adjusted from one day to the following one in each participant (i.e., 24h), and then data were downloaded. Due to the limitations of material resources, waves of three adolescents per day were carried out. For each wave, adolescents were met at 8:00 a.m. in the room allocated in the school gym, so they could go then to start their school day at the regular time (i.e., 8:30 a.m.). Additionally to the activity wristbands, an accelerometer was adjusted on the adolescents' right hip using an elastic waistband. Adolescents were instructed to wear the activity trackers (i.e., activity wristbands and accelerometer) for the whole day until bedtime maintaining their habitual PA levels, and they were asked to take them off only when they took a bath/shower, or to leave them in a plastic box inside their schoolbags just before going to bed. In case one adolescent reported that he/she was going to engage in aquatic activities that day, he/she was requested to come the next day (and another adolescent was requested instead). In the morning of the following day, the activity trackers were collected and adjusted to the next three participants following the same protocol.

Statistical analysis

Descriptive statistics for all the variables of the included participants were calculated. Firstly, all the statistical tests assumptions were check and met (e.g., histograms and Q-Q plots for normality). Furthermore, univariate (i.e., z ± 3.0) and multivariate outliers (i.e., Mahalanobis distance) were removed. Afterward, the agreement between the PA scores (i.e., continuous variables) measured by the wearables and the video (study 1)/accelerometers (study 2) were calculated as follows: a) Limits of Agreement (LOA) with its confident intervals (95% CI) (Bland & Altman, 1986); b) Mean Absolute Error (MAE) (Willmott & Matsuura, 2005); c)

Mean Absolute Percentage Error (MAPE) (Johnston et al., 2021); and d) Intraclass Correlation Coefficient (ICC), and its 95% CI, by a two-way random-effects model with absolute agreement and single measurement [also known as ICC(2,1)] (Koo & Li, 2016). Additionally, LOA plots, which are the individual participant differences between the two scores plotted against the respective individual means, were performed (Bland & Altman, 1986). Heteroscedasticity was also examined objectively by calculating the Pearson´s correlation coefficient (r) between the absolute differences and the individual means (Atkinson & Nevill, 1998). Based on Cohen's (1992) benchmarks, a correlation coefficient > 0.50 was considered as indicative of heteroscedasticity. Finally, the agreement between the PA scores dichotomized as achieving or not achieving the daily recommendations of 10,000 steps (Mayorga-Vega et al., 2021) and 60 min of MVPA (i.e., categorical variables) (World Health Organization, 2020) measured by the wearable bands and the accelerometers (study 2) were calculated as the proportion of agreement [P = number of agreements/ (number of agreements + disagreements)] and kappa coefficient (k) (Hernaez, 2015). Agreement values were interpreted as follows: MAPE, > 20.0% poor; 15.1-20.0% questionable; 10.1-15.0% acceptable; 5.1-10.0% good; and 0.0-5.0% excellent (Bai et al., 2021; Johnston et al., 2021). ICC, 0.00-0.49 unacceptable; 0.50-0.59 poor; 0.60-0.69 questionable; 0.70-0.79 acceptable; 0.80-0.89 good; and 0.90-1.00 excellent (Nunnally, 1978); k, 0.00-0.39 poor; 0.40-0.59 acceptable; 0.60-0.74 good; and 0.75-1.00 excellent (Cicchetti, 2001). Based on statistical inference, each ICC value was interpreted according to its 95% IC, that means, there was 95% chance that the true ICC value landed on any point between the 95% IC range (Koo & Li, 2016). Regarding kappa values, < 0.00 poor; 0.00-0.20 slight; 0.21-0.40 fair; 0.41-0.60 moderate; 0.61-0.80 substantial; 0.81-1.00 almost perfect (Landis & Koch, 1977). All statistical analyses were performed using the SPSS version 25.0 for Windows (IBM® SPSS® Statistics). The statistical significance level was set at $p < 0.05$.

Results

General characteristics

Figure 1 shows the flow diagram of the participants through the studies. An initial sample of 62 and 61 adolescents met the inclusion criteria and agreed to participate in the studies 1 and 2, respectively. Since some adolescents met at least one exclusion criterion, the final sample consisted of 56 and 51 participants for the studies 1 and 2, respectively (i.e., non-compliance rate of 9.7% and 16.4%, respectively). Table 1 shows the general characteristics of the included participants.

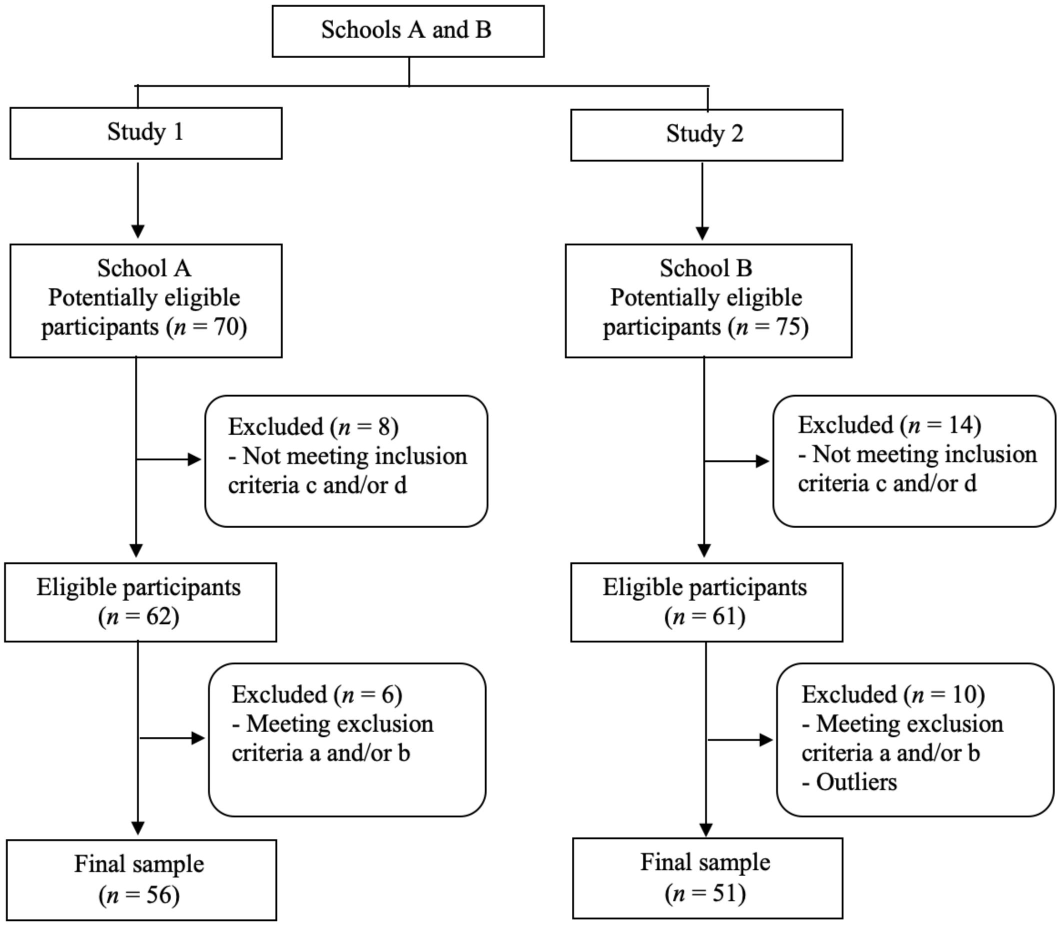


Figure 1. Flow chart of participants included in the present study

Table 1. General characteristics of the analyzed participants in each study

	Study 1 (*n* = 56)	Study 2 (*n* = 51)
Age (years)	14.7 (1.7)	14.0 (1.5)
Grade (1st/2nd/3rd/4th/5th/6th)	21.4/17.9/12.5/19.6/14.3/14.3	19.6/21.6/21.6/19.6/17.6/0.0
Gender (males/females)	51.8/48.2	51.0/49.0
Body mass (kg)	58.1 (12.9)	58.6 (13.2)
Body height (cm)	165.0 (11.3)	161.8 (8.3)
Body mass index (kg/m^2)	21.2 (3.3)	22.4 (4.8)
Overweight/obesity (no/yes)	85.7/14.3	31.4/68.6
Non-dominant hand (left/right)	87.5/12.5	90.2/9.8
Self-reported habitual PA (days/week)	3.7 (1.3)	2.4 (1.6)

Note. Data are reported as mean (standard deviation) or percentage. Study 1: Validity in structured free-living conditions; Study 2: Validity in unstructured free-living conditions.

Study 1: Validity of the activity wristbands during structured free-living conditions.

The average speed (SD) in each condition was as follows: Slow pace walking = 1.1 (0.2) m/s [3.9 (0.5) km/h]; normal pace walking = 1.5 (0.1) m/s [5.2 (0.5) km/h]; brisk pace walking = 1.9 (0.2) m/s [6.8 (0.6) km/h]; and running = 3.1 (0.4) m/s [11.0 (1.3) km/h]. The average steps cadence (SD) in each condition was as follows: Slow pace walking = 99.7 (7.3) steps/min; normal pace walking = 113.3 (5.0) steps/min; brisk pace walking = 128.7 (7.0) steps/min; and running = 167.2 (19.1) steps/min.

Table 2 shows the validity of the activity wristbands for estimating steps during structured free-living conditions. The agreement between the steps assessed by the activity wristbands and video-based count tended to be higher for slow pace walking, followed by running, normal pace walking, and brisk pace walking. Based on the MAPE values, the validity results of the steps assessed by all five devices (three wristbands and the two mobile phones with their respective apps; due to the three apps of the Apple iPhone mobile registered exactly the same parameters without distinction between them, only "Apple iPhone" has been reported in this Results section, obviating distinguishing the three apps) in the four conditions were excellent (i.e., < 5%). Regarding the validity results of the steps based on the values of ICC 95% IC, similarly, the Samsung apps Pacer and Google Fit in all conditions ranged between good to excellent. Furthermore, the validity results of the steps assessed by the Apple iPhone and Apple Watch Series 5 were excellent, and by the Samsung Pedometer and Samsung Watch Active 2 were good-excellent in the three conditions. However, their validity was worse in one condition: Apple iPhone in slow pace walking condition was acceptable-excellent; Apple Watch Series 5 in brisk pace walking was poor-excellent; Samsung Pedometer in normal pace walking condition was questionable-good; and Samsung Watch Active 2 in brisk pace walking was questionable-excellent. Similarly, while the validity results of the steps assessed by the Xiaomi Mi Band 5 were excellent in the slow pace walking and running conditions, they were acceptable-excellent and poor-good in the normal and brisk pace walking conditions, respectively. Supplementary Figure 1 shows the LOA plots. Pearson´s correlation coefficients showed that with all the instruments there was no heteroscedasticity on any walking/running condition (Supplementary Table 1).

Table 2. Validity of the wearable bands for estimating steps during structured free-living conditions (*n* = 56)

Instrument	Mean (SD)	LOA (95% CI)	MAE	MAPE	ICC (95% CI)
Slow pace walking					
Video-based count	306.6 (30.9)	-	-	-	-
Samsung Pedometer	307.6 (32.4)	-0.9 (-18.1, 16.3)	6.2	2.0	0.96 (0.94, 0.98)
Samsung Pacer	308.0 (29.8)	-1.4 (-20.2, 17.4)	5.8	1.9	0.95 (0.92, 0.97)
Samsung Google Fit	307.1 (30.2)	-0.5 (-17.9, 16.9)	5.6	1.8	0.96 (0.93, 0.98)
Apple iPhone	307.1 (28.6)	-0.5 (-32.4, 31.4)	8.2	2.7	0.85 (0.76, 0.91)
Xiaomi Mi Band 5	303.7 (31.1)	2.9 (-7.9, 13.7)	4.1	1.3	0.98 (0.95, 0.99)
Samsung Watch Active 2	305.7 (31.0)	0.9 (-8.5, 10.3)	3.9	1.3	0.99 (0.98, 0.99)
Apple Watch Series 5	304.7 (30.9)	1.9 (-4.8, 8.6)	3.1	1.0	0.99 (0.98, 1.00)
Normal pace walking					
Video-based count	261.4 (22.6)	-	-	-	-
Samsung Pedometer	259.9 (23.9)	1.5 (-28.5, 31.5)	5.6	2.1	0.79 (0.66, 0.87)
Samsung Pacer	262.0 (23.8)	-0.6 (-15.7, 14.5)	4.2	1.6	0.95 (0.91, 0.97)
Samsung Google Fit	262.8 (22.8)	-1.4 (-12.2, 9.4)	3.5	1.4	0.97 (0.95, 0.98)
Apple iPhone	262.4 (23.9)	-1.0 (-15.9, 13.9)	4.5	1.7	0.95 (0.91, 0.97)
Xiaomi Mi Band 5	256.3 (25.2)	5.1 (-10.6, 20.8)	6.0	2.4	0.92 (0.79, 0.97)
Samsung Watch Active 2	258.1 (22.3)	3.3 (-6.1, 12.7)	4.0	1.5	0.97 (0.89, 0.99)
Apple Watch Series 5	258.3 (22.7)	3.1 (-6.9, 13.1)	3.9	1.5	0.97 (0.91, 0.99)
Brisk pace walking					
Video-based count	227.6 (18.2)	-	-	-	-
Samsung Pedometer	230.5 (20.5)	-2.9 (-20.7, 14.9)	5.8	2.6	0.88 (0.80, 0.93)
Samsung Pacer	229.3 (18.8)	-1.7 (-15.2, 11.8)	3.7	1.7	0.93 (0.88, 0.96)
Samsung Google Fit	229.3 (18.9)	-1.8 (-15.5, 11.9)	3.9	1.7	0.93 (0.88, 0.96)
Apple iPhone	226.4 (18.1)	1.2 (-5.5, 7.9)	2.5	1.1	0.98 (0.96, 0.99)
Xiaomi Mi Band 5	221.6 (23.2)	6.0 (-20.7, 32.7)	7.3	3.2	0.76 (0.58, 0.86)
Samsung Watch Active 2	222.1 (18.1)	5.5 (-9.4, 20.4)	6.8	3.0	0.87 (0.62, 0.95)
Apple Watch Series 5	220.9 (21.1)	6.7 (-14.5, 27.9)	8.1	3.6	0.81 (0.56, 0.90)
Running					
Video-based count	185.0 (24.0)	-	-	-	-
Samsung Pedometer	187.7 (24.4)	-2.7 (-21.5, 16.1)	5.7	3.2	0.92 (0.86, 0.95)
Samsung Pacer	188.3 (25.6)	-3.3 (-18.0, 11.4)	4.7	2.5	0.95 (0.90, 0.97)
Samsung Google Fit	188.4 (25.7)	-3.4 (-18.1, 11.3)	4.9	2.7	0.95 (0.89, 0.97)
Apple iPhone	185.6 (23.5)	-0.6 (-12.4, 11.2)	3.3	1.8	0.97 (0.95, 0.98)
Xiaomi Mi Band 5	183.6 (24.6)	1.4 (-8.6, 11.4)	3.6	2.0	0.98 (0.96, 0.99)
Samsung Watch Active 2	183.6 (23.5)	1.5 (-5.8, 8.8)	2.8	1.5	0.99 (0.97, 0.99)
Apple Watch Series 5	184.8 (23.8)	0.2 (-4.1, 4.5)	1.5	0.8	1.00 (0.99, 1.00)

Note. SD = Standard deviation; LOA = Limits of agreement; 95% CI = 95% confident interval; MAE = Mean absolute error; MAPE = Mean absolute percentage error; ICC = Intraclass correlation coefficient.

Study 2: Validity of the activity wristbands during daily unstructured free-living conditions

Table 3 shows the validity of the activity wristbands for estimating daily PA in adolescents during unstructured free-living conditions. Based on the values of both the MAPE and 95% IC of the ICC, the validity results of the steps assessed by the Xiaomi Mi Band 5 were good and excellent, respectively. Similarly, the validity results of the steps assessed by the Samsung Watch Active 2 were acceptable and excellent based on the MAPE and 95% IC of the ICC, respectively. However, as regards the Apple Watch Series 5, the validity results showed that it was poor for the values of the MAPE and questionable-excellent for those with the 95% IC of the ICC. Regarding the validity results for the MVPA and total PA assessment, the validity results for the three activity wristbands were poor for the values of both the MAPE and 95% IC of the ICC (exceptionally was questionable-excellent for the values of 95% IC of the ICC assessed by the Apple Watch Series 5). Supplementary Figure 2 shows the LOA plots. Pearson´s correlation coefficients did not show heteroscedasticity, except with the MVPA with the Xiaomi Mi Band 5 (Supplementary Table 1).

Table 4 shows the validity of the activity wristbands for estimating the daily PA recommendations in adolescents during unstructured free-living conditions. A total of 13.7% and 17.6% of adolescents met the accelerometer-measured step- and MVPA-based recommendations, respectively. The validity results of the daily step-based recommendations assessed by the three activity wristbands were excellent. Regarding the daily MVPA-based recommendation, while the validity results with the Apple Watch Series 5 were excellent, for the Xiaomi Mi Band 5 were poor.

Table 3. Validity of the wearable bands for estimating daily physical activity during unstructured free-living conditions (*n* = 51)

Instrument	Mean (SD)	LOA (95% CI)	MAE	MAPE	ICC (95% CI)
Steps (n)					
ActiGraph wGT3X-BT	6562.1 (2662.6)	-	-	-	-
Samsung Watch Active 2	6290.2 (3002.1)	271.9 (-1349.2, 1893.0)	679.3	11.4	0.95 (0.92, 0.97)
Apple Watch Series 5	7316.0 (2886.6)	-753.9 (-2912.8, 1405.0)	1001.1	18.0	0.89 (0.69, 0.95)
Xiaomi Mi Band 5	6688.0 (3155.9)	-125.9 (-1976.7, 1724.9)	591.7	9.4	0.95 (0.91, 0.97)
Moderate-to-vigorous physical activity (min)					
ActiGraph w-GT3X-BT	41.8 (17.8)	-	-	-	-
Apple Watch Series 5	36.3 (20.8)	5.4 (-12.4, 23.2)	8.2	22.6	0.86 (0.67, 0.93)
Xiaomi Mi Band 5	65.8 (41.8)	-24.0 (-108.1, 60.1)	37.8	120.6	0.09 (0.00, 0.31)
Xiaomi Mi Band 5[a]	21.3 (15.3)	20.5 (-1.3, 42.3)	20.7	53.5	0.44 (0.00, 0.77)
Total physical activity (min)					
ActiGraph wGT3X-BT	216.1 (61.5)	-	-	-	-
Samsung Watch Active 2	70.3 (36.6)	145.8 (44.1, 247.5)	145.8	67.6	0.09 (0.00, 0.32)
Xiaomi Mi Band 5	83.0 (34.4)	133.0 (37.7, 228.3)	133.0	61.3	0.12 (0.00, 0.38)
Xiaomi Mi Band 5[b]	64.6 (38.4)	151.5 (52.3, 250.7)	151.5	70.6	0.10 (0.00, 0.34)

Note. SD = Standard deviation; LOA = Limits of agreement; 95% CI = 95% confident interval; MAE = Mean absolute error; MAPE = Mean absolute percentage error; ICC = Intraclass correlation coefficient; [a] Brisk walking time (min); [b] Slow-brisk walking time (min).

Table 4. Validity of the wearable bands for estimating the daily physical activity recommendations during unstructured free-living conditions (*n* = 51)

		ActiGraph wGT3X-BT						
		10,000 steps				60 min of MVPA		
Instrument		%TP	*P*	*k*		%TP	*P*	*k*
Samsung Watch Active 2	10,000 steps	13.7	0.96	0.83†	60 min of MVPA	-	-	-
Apple Watch Series 5		15.7	0.94	0.77†		17.6	0.96	0.87†
Xiaomi Mi Band 5		17.6	0.96	0.85†		60.8	0.41	-0.03
Xiaomi Mi Band 5[a]		-	-	-		2.0	0.84	0.17*

Note. MVPA = Moderate-to-vigorous physical activity; %TP= Percentage of total positive cases according to the recommendation; *P* = Proportion of agreement; *k* = Kappa coefficient. [a] Brisk walking time (min).
* $p < 0.05$; ‡ < 0.01; and † $p < 0.001$

Discussion

Regarding the general objective of the study referred to the validity of the nine wearables (i.e., activity wristbands and mobile apps) for estimating PA in adolescents during free-living conditions, it is necessary to differentiate the measurement of steps and other PA parameters in structured or unstructured free-living conditions. In regard to the measurement of steps, all wearables were valid under the two conditions (i.e., structured and unstructured free-living), but to measure PA parameters in unstructured free-living conditions only one option could be considered as valid (i.e., Apple Watch Series 5 was valued as questionable-excellent for estimating MVPA using the criterion of the values of 95% IC of the ICC). Previous research coincided with the present study in regard to steps counting, showing high validity outcomes in adults and youths, both with other wristbands companies such as Jawbone or Fitbit (Evenson et al., 2015; Kang et al., 2019) and with the ones used in this study, although with different models of activity wristbands (i.e., Apple, Xiaomi, and Samsung) (e.g., Fuller et al., 2020; Hao et al., 2021). However, similar to this study, other exercise parameters such as PA intensities or energy expenditure have shown inadequate validity results (e.g., DeGroote et al., 2020; Evenson et al., 2015; Feehan et al., 2018; Fuller et al., 2020; Voss et al., 2017).

Specifically, the aim of the study 1 was to compare the criterion-related validity of the steps assessed by the nine wearables [i.e., mobile apps Pedometer, Pacer, Google Fit for Androids mobiles (Samsung); Pedometer, Pacer, and Apple Health for Apple mobiles (iPhone); and the wristbands Samsung Watch Active 2, Apple Watch Series 5, and Xiaomi Mi Band 5 under] in structured free-living conditions. All cases were excellent for the four conditions (i.e., slow pace walking, normal pace walking, brisk pace walking, and running), considering the criterion of MAPE values, and good to excellent taking into account the 95% IC of the ICC criterion. Probably this is the only condition and setting where there is a higher consensus in previous literature regarding that activity wristbands are valid for measuring steps accurately (i.e., in structured free-living conditions) (e.g., Fuller et al., 2020; Fokkema et al. 2017; or Hao et al., 2021). There were no clear differences between data obtained among devices. Being the activity wristbands more comfortable than mobile phones to wear, and particularly being the wristband Xiaomi Mi Band 5 the cheapest one (around 25€), it seems logical to conclude that this is the more recommended option for measuring steps under structured free-living conditions. For instance, in the school setting where mobile phones are not allowed, together with the fact that smartwatches are more expensive, the Xiaomi Mi Band 5 could provide Physical Education teachers with the opportunity to measure the adolescents' steps in Physical Education-based health promotion programs accurately.

In regard to the specific aim of the study 2, to compare the convergent validity of the daily steps, total PA, and MVPA scores assessed by the three wristbands under unstructured free-living conditions, different results were obtained, and different conclusions could be deduced. First, for the measurement of daily steps, the Xiaomi Mi Band 5 obtained the best validity result (i.e., good and excellent for the MAPE values and 95% IC of the ICC criteria, respectively); followed by Samsung Watch Active 2 (acceptable for MAPE values and excellent for 95% IC of the ICC); and finally, the Apple Watch Series 5 (poor and questionable-excellent for MAPE and for 95% IC of the ICC criteria, respectively). Previous research also obtained good validity results of daily steps in unstructured free-living conditions with adolescents (Schneider & Chau, 2016; Šimůnek et al., 2019; Yang et al., 2019), although the first one studied the Fitbit activity wristband detecting an overestimated registration in comparison to the ActiGraph accelerometer; the second one studied the Garmin 1 and 3 activity wristband models, obtaining good levels of accuracy in comparison to the Yamax pedometer as the reference standard; and the third one studied the more used wristband in China (i.e., a previous version of Xiaomi than the one used in the present study) in comparison to the ActiGraph accelerometer. Secondly, for the MVPA measurement, the best validity result was obtained by the Apple Watch Series 5 with questionable-excellent values using the 95% IC of the ICC criterion. The Xiaomi Mi Band 5 presented unsatisfactory results, and the Samsung Watch active 2 did not provide this parameter. Previous literature with adults, such as the revision performed by Feehan et al. (2018) concluded that the more the intensity of PA increases, the more is the overestimation of the Fitbit wristband in comparison to the ActiGraph accelerometer. In regard to adolescent samples, previous literature is scarce. On the one hand, recent and previous review studies agree with the fact that some specific models of wristbands have obtained adequate validity results for MVPA (Gorzelitz et al., 2020), although they only analyzed one study with adolescents (i.e., Schneider & Chau, 2016), which examined the validity of the Fitbit activity wristband. On the other hand, another study with adolescents carried out by Yang et al. (2019) obtained inadequate validity results for MVPA (i.e., MAPE = 61.2%). Therefore, further studies are needed in order to achieve a more consensus in this variable, which probably with new models of wristbands that incorporate new algorithms may provide more accurate measurement of MVPA in the future. Third, for total PA, none of the activity wristbands analyzed provided adequate validity results, and the Apple Watch Series 5 did not provide this parameter.

Finally, for the PA daily recommendations, and taking into consideration the criterion of 10,000 steps, all devices obtained excellent values of validity. In relation to the 60' of MVPA criterion, only the Apple Watch Series 5 obtained excellent results, while the Xiaomi Mi Band 5 with poor results. The Samsung Watch Active 2 did not provide this parameter. Therefore, valuable outputs were obtained for future intervention programs because the three wristbands were valid for accurately classifying adolescents according to whether or not they met the recommended 10,000 steps per day, and also the Apple Watch Series 5 for accurately classifying them according to minutes in MVPA, which provides insight into the effectiveness of the program for adolescents to achieve the international recommendations.

In relation to the mobile apps specifically, it is important to denote that the Apple iPhone 11 used the same algorithm for assessing PA parameters in the three evaluated apps (i.e., Pedometer, Pacer, and Apple Health), providing the same results for all of them. Consequently, the possibilities offered by the Apple Company related to PA parameters assessed by mobile apps are reduced in comparison with Android mobile phones.

To sum up, according to this study, activity wearables are valid to report adolescents' steps, daily MVPA, and PA recommendations (i.e., 10,000 steps or 60' of MVPA per day). Consequently, these wearables are also useful for reporting valid feedbacks for being used in adolescents' progressive challenges to achieve daily recommendations in health-promotion interventions (included in the Physical Education setting, where Physical Education teachers do have not enough feasible resources to apply in out-of-school time for measuring PA with secondary students). Wearables characteristics of lightweight, easy to use and understand, and economically affordable (Parra Saldías et al., 2018), together with the results obtained in the present study, makes these devices feasible and applicable to consumers (researchers or not), and a great option for PA promotion as much as a strategy to promote PA behavior as accurate measurement of PA (e.g., using them as a control instrument of the daily PA recommendations).

A strength of this study was, to our knowledge, the first study to examine the validity of more advanced and used models of wearables (at the moment of the data collection, International Data Corporation's Worldwide Quarterly Wearable Device Tracker Report, 2021) in structured and unstructured free-living conditions and with adolescents, analyzing both PA parameters (i.e., steps, MVPA, and total PA) and consensus daily PA recommendations criteria (i.e., 10,000 steps and 60' MVPA). Furthermore, since the main goal of wearables is to assess adolescents' daily PA levels or to use them as a motivating tool during daily life, the evaluation in free-living conditions as in the present study are closer to reality, and therefore they are meaningful and useful (Duncan et al., 2018). Taking into account that the majority of the previous studies have been performed in adults or special samples such as people with different health diseases (e.g., Feehan et al., 2018 or Fokkema et al., 2017), this study covers an important gap in the scientific knowledge, which is also applicable to practical interventions in order to improve health promotion and behavior change programs in regard to PA, both in Physical Education and in other health or sport contexts (e.g., Duncan et al., 2012; Strath & Rowley, 2018).

Regarding the limitations, this study used a non-probability and relatively small sample, which limits the generalizability of the obtained results to this particular context (i.e., adolescents with similar characteristics and PA behavior). Secondly, Actigraph accelerometers have been shown as the most common and valid method for objectively assessing adolescents' PA levels during free-living conditions (e.g., Shephard & Tudor-Locke, 2016; Van Hecke et al., 2016; Hickey et al., 2016; Trost et al., 2011; Migueles et al., 2017). However, some methodological issues have not achieved an evidence-based consensus, despite the best current evidence-based decisions were adopted in the present study (Migueles et al., 2017). Therefore, it may contribute to the variability of wearables convergent validity obtained results. Finally, the third limitation is related to the models of wearable evaluated and its associated consequences. Two repercussions could be mentioned: (a) at this moment, more advanced models of some wearables are in the market (i.e., the wristband

Xiaomi Mi Band 6, the mobile iPhone 12, the Apple Watch Series 6, or the Samsung Watch Active 4, for instance); and (b) some Companies did not provide their algorithms to classify PA at different intensities (e.g., Xiaomi or Apple companies). Because of all of these mentioned limitations, further studies should be performed in the future in order to improve the knowledge regarding the accuracy and valid registration of PA parameters by wearables, both for consumers' and researchers' decisions of their personal use. Moreover, as commented by other previous studies (e.g., Gorzelitz et al., 2020), a standardization of validation protocols (i.e., validation methods, standard references, as well as reporting outcomes) are needed to address in future studies in order to allow for comparability across studies and to report strong evidence regarding the different activity wearable measurements.

Conclusions

Wearables checked in this study have shown adequate validity results in order to measure steps in structured and unstructured free-living conditions. However, in unstructured free-living conditions and for assessing MVPA, only Apple Watch Series 5 reported valid results, and none of them was valid for measuring total PA. Therefore, depending on the user's/researcher's aim and context, one or another wristband could be more adequate, mainly because of their valid measurements and their costs.

Acknowledgements

This work is funded by FEDER/*Junta de Andalucía-Consejería de Economía y Conocimiento* (Reference: B-SEJ-029-UGR18). Carolina Casado-Robles is also supported by a research grant from the Spanish Ministry of Universities [grant number: FPU16/03314].

Supplementary Files

Scan or click here to see Supplementary Files [Escanea el código o pincha aquí para ver los Archivos Suplementarios]

Supplementary Files

References

Adamakis, M. (2020). Criterion validity of wearable monitors and smarphone applications to measure physical activity energy expenditure in adolescents. *Sport Sciences for Health, 16*, 755–763. https://doi.org/10.1007/s11332-020-00654-2

Arifin, W. (2018). A web-based sample size calculator for reliability studies. *Education in Medicine Journal, 10*(3), 67–76. https://doi.org/10.21315/eimj2018.10.3.8

Atkinson, G., & Nevill, A. M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports medicine, 26*(4), 217–238. https://doi.org/10.2165/00007256-199826040-00002

Bai, Y., Tompkins, C., Gell, N., Dione, D., Zhang, T., & Byun, W. (2021). Comprehensive comparison of Apple Watch and Fitbit monitors in a free-living setting. *PloS one, 16*(5), e0251975. https://doi.org/10.1371/journal.pone.0251975

Bland, J., & Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet, 1*(8476), 307–310.

Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., Gatsonis, C. A., Glasziou, P. P., Irwig, L., Lijmer, J. G., Moher, D., Rennie, D., de Vet, H. C., Kressel, H. Y., Rifai, N., Golub, R. M., Altman, D. G., Hooft, L., Korevaar, D. A., Cohen, J. F., & STARD Group (2015). STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ, 351*, h5527. https://doi.org/10.1136/bmj.h5527

Brickwood, K. J., Watson, G., O'Brien, J., & Williams, A. D. (2019). Consumer-Based Wearable Activity Trackers Increase Physical Activity Participation: Systematic Review and Meta-Analysis. *JMIR mHealth and uHealth*, *7*(4), e11819. https://doi.org/10.2196/11819

Cicchetti, D. (2001). The precision of reliability and validity estimates re-visited: Distinguishing between clinical and statistical significance of sample size requirements. *Journal of Clinical and Experimental Neuropsychology, 23*(5), 695–700.

Cohen, J. (1992). A power primer. *Psychological bulletin, 112*(1), 155–159. https://doi.org/10.1037//0033-2909.112.1.155

Cole, T. J., Bellizzi, M. C., Flegal, K. M., & Dietz, W. H. (2000). Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ, 320*, 1240-1243. https://doi.org/10.1136/bmj.320.7244.1240

Cosoli, G., Spinsante, S., & Scalise, L. (2020). Wrist-worn and chest-strap wearable devices: Systematic review on accuracy and metrological characteristics. *Measurement, 159,* 107789. https://doi.org/10.1016/j.measurement.2020.107789

Crossley, S., McNarry, M. A., Rosenberg, M., Knowles, Z. R., Eslambolchilar, P., & Mackintosh, K. A. (2019). Understanding Youths' Ability to Interpret 3D-Printed Physical Activity Data and Identify Associated Intensity Levels: Mixed-Methods Study. *Journal of medical Internet research, 21*(2), e11253. https://doi.org/10.2196/11253

Da Silva, M. P., Fontana, F. E., Callahan, E., Mazzardo, O., & De Campos, W. (2015). Step-Count Guidelines for Children and Adolescents: A Systematic Review. *Journal of physical activity & health, 12*(8), 1184–1191. https://doi.org/10.1123/jpah.2014-0202

Degroote, L., Hamerlinck, G., Poels, K., Maher, C., Crombez, G., De Bourdeaudhuij, I., Vandendriessche, A., Curtis, R. G., & DeSmet, A. (2020). Low-Cost Consumer-Based Trackers to Measure Physical Activity and Sleep Duration Among Adults in Free-Living Conditions: Validation Study. *JMIR mHealth and uHealth, 8*(5), e16674. https://doi.org/10.2196/16674

Duncan, M. J., Wunderlich, K., Zhao, Y., & Faulkner, G. (2018). Walk this way: validity evidence of iphone health application step count in laboratory and free-living conditions. *Journal of sports sciences, 36*(15), 1695–1704. https://doi.org/10.1080/02640414.2017.1409855

Duncan, M., Birch, S., & Woodfield, L. (2012). Efficacy of an integrated school curriculum pedometer intervention to enhance physical activity and to reduce weight status in children. *European Physical Education Review, 18*(3), 396–407. https://doi.org/10.1177/1356336X12450799

El-Amrawy, F., & Nounou, M. I. (2015). Are Currently Available Wearable Devices for Activity Tracking and Heart Rate Monitoring Accurate, Precise, and Medically Beneficial?. *Healthcare informatics research*, *21*(4), 315–320. https://doi.org/10.4258/hir.2015.21.4.315

Evenson, K. R., Catellier, D. J., Gill, K., Ondrak, K. S., & McMurray, R. G. (2008). Calibration of two objective measures of physical activity for children. *Journal of sports sciences*, *26*(14), 1557–1565. https://doi.org/10.1080/02640410802334196

Evenson, K. R., Goto, M. M., & Furberg, R. D. (2015). Systematic review of the validity and reliability of consumer-wearable activity trackers. *The international journal of behavioral nutrition and physical activity*, *12*, 159. https://doi.org/10.1186/s12966-015-0314-1

Fairclough, S. J., Noonan, R., Rowlands, A. V., Van Hees, V., Knowles, Z., & Boddy, L. M. (2016). Wear Compliance and Activity in Children Wearing Wrist- and Hip-Mounted Accelerometers. *Medicine and science in sports and exercise*, *48*(2), 245–253. https://doi.org/10.1249/MSS.0000000000000771

Feehan, L. M., Geldman, J., Sayre, E. C., Park, C., Ezzat, A. M., Yoo, J. Y., Hamilton, C. B., & Li, L. C. (2018). Accuracy of Fitbit Devices: Systematic Review and Narrative Syntheses of Quantitative Data. *JMIR mHealth and uHealth*, *6*(8), e10527. https://doi.org/10.2196/10527

Fokkema, T., Kooiman, T. J., Krijnen, W. P., VAN DER Schans, C. P., & DE Groot, M. (2017). Reliability and Validity of Ten Consumer Activity Trackers Depend on Walking Speed. *Medicine and science in sports and exercise*, *49*(4), 793–800. https://doi.org/10.1249/MSS.0000000000001146

Fuller, D., Colwell, E., Low, J., Orychock, K., Tobin, M. A., Simango, B., Buote, R., Van Heerden, D., Luan, H., Cullen, K., Slade, L., & Taylor, N. (2020). Reliability and Validity of Commercially Available Wearable Devices for Measuring Steps, Energy Expenditure, and Heart Rate: Systematic Review. *JMIR mHealth and uHealth*, *8*(9), e18694. https://doi.org/10.2196/18694

Gal, R., May, A. M., Van Overmeeren, E. J., Simons, M., & Monninkhof, E. M. (2018). The Effect of Physical Activity Interventions Comprising Wearables and Smartphone Applications on Physical Activity: a Systematic Review and Meta-analysis. *Sports Medicine - Open, 4*(1). https://doi.org/10.1186/s40798-018-0157-9

Gorzelitz, J., Farber, C., Gangnon, R., & Cadmus-Bertram, L. (2020). Accuracy of Wearable Trackers for Measuring Moderate- to Vigorous-Intensity Physical Activity: A Systematic Review and Meta-Analysis, *Journal for the Measurement of Physical Behaviour, 3*(4), 346-357. https://doi.org/10.1123/jmpb.2019-0072

Guthold, R., Stevens, G. A., Riley, L. M., & Bull, F. C. (2020). Global trends in insufficient physical activity among adolescents: a pooled analysis of 298 population-based surveys with 1.6 million participants. *The Lancet Child & Adolescent Health, 4*(1), 23-35. https://doi.org/10.1016/S2352-4642(19)30323-2

Hao, Y., Ma, X. K., Zhu, Z., & Cao, Z. B. (2021). Validity of Wrist-Wearable Activity Devices for Estimating Physical Activity in Adolescents: Comparative Study. *JMIR mHealth and uHealth, 9*(1), e18320. https://doi.org/10.2196/18320

Henriksen, A., Haugen Mikalsen, M., Woldaregay, A. Z., Muzny, M., Hartvigsen, G., Hopstock, L. A., & Grimsgaard, S. (2018). Using Fitness Trackers and Smartwatches to Measure Physical Activity in Research: Analysis of Consumer Wrist-Worn Wearables. *Journal of medical Internet research, 20*(3), e110. https://doi.org/10.2196/jmir.9157

Hernaez, R. (2015). Reliability and agreement studies: A guide for clinical investigators. *Gut, 64*(7), 1018–1027. https://doi.org/10.1136/gutjnl-2014-308619

Hickey, A., John, D., Sasaki, J. E., Mavilia, M., & Freedson, P. (2016). Validity of Activity Monitor Step Detection Is Related to Movement Patterns. *Journal of physical activity & health, 13*(2), 145–153. https://doi.org/10.1123/jpah.2015-0203

International Data Corporation. (2021). *European Wearable Market Showed Positive Growth in 1Q21, Says IDC.* https://www.idc.com/getdoc.jsp?containerId=prEUR147995121

Johnston, W., Judice, P. B., Molina García, P., Mühlen, J. M., Lykke Skovgaard, E., Stang, J., Schumann, M., Cheng, S., Bloch, W., Brønd, J. C., Ekelund, U., Grøntved, A., Caulfield, B., Ortega, F. B., & Sardinha, L. B. (2021). Recommendations for determining the validity of consumer wearable and smartphone step count: expert statement and checklist of the INTERLIVE network. *British Journal of Sports Medicine*, *55*(14), 780 – 793. https://doi.org/10.1136/bjsports-2020-103147

Kang, S., Kim, Y., Byun, W., Suk, J., & Lee, J. M. (2019). Comparison of a Wearable Tracker with Actigraph for Classifying Physical Activity Intensity and Heart Rate in Children. *International journal of environmental research and public health*, *16*(15), 2663. https://doi.org/10.3390/ijerph16152663

Koo, T., & Li, M. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Kottner, J., Audigé, L., Brorson, S., Donner, A., Gajewski, B. J., Hróbjartsson, A., Roberts, C., Shoukri, M., & Streiner, D. L. (2011). Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *Journal of clinical epidemiology*, *64*(1), 96–106. https://doi.org/10.1016/j.jclinepi.2010.03.002

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174.

Lee, J. A., Williams, S. M., Brown, D. D., & Laurson, K. R. (2015). Concurrent validation of the Actigraph gt3x+, Polar Active accelerometer, Omron HJ-720 and Yamax Digiwalker SW-701 pedometer step counts in lab-based and free-living settings. *Journal of sports sciences*, *33*(10), 991–1000. https://doi.org/10.1080/02640414.2014.981848

Mayorga-Vega, D., Casado-Robles, C., López-Fernández, I., & Viciana, J. (2021). A comparison of the utility of different step-indices to translate the physical activity recommendation in adolescents. *Journal of Sports Sciences*, *39*(4), 469–479. https://doi.org/10.1080/02640414.2020.1826667

Mayorga-Vega, D., Casado-Robles, C., Viciana, J., & López-Fernández, I. (2019). Daily Step-Based Recommendations Related to Moderate-to-Vigorous Physical Activity and Sedentary Behavior in Adolescents. *Journal of Sports Science & Medicine, 18*(4), 586–595.

Migueles, J. H., Cadenas-Sanchez, C., Ekelund, U., Delisle Nyström, C., Mora-Gonzalez, J., Löf, M., Labayen, I., Ruiz, J. R., & Ortega, F. B. (2017). Accelerometer Data Collection and Processing Criteria to Assess Physical Activity and Other Outcomes: A Systematic Review and Practical Considerations. *Sports medicine*, *47*(9), 1821–1845. https://doi.org/10.1007/s40279-017-0716-0

Nunnally, J. (1978). *Psychometric theory*. McGraw-Hill.

O'Neill, B., McDonough, S. M., Wilson, J. J., Bradbury, I., Hayes, K., Kirk, A., Kent, L., Cosgrove, D., Bradley, J. M., & Tully, M. A. (2017). Comparing accelerometer, pedometer and a questionnaire for measuring physical activity in bronchiectasis: a validity and feasibility study?. *Respiratory research*, *18*(1), 16. https://doi.org/10.1186/s12931-016-0497-2

OECD. (2019). *OECD Future of Education 2030. Making Physical Education dynamic and inclusive for 2030. International curriculum analysis*. OECD Publishing.

Oliver, M., Badland, H. M., Schofield, G. M., & Shepherd, J. (2011). Identification of accelerometer nonwear time and sedentary behavior. *Research quarterly for exercise and sport*, *82*(4), 779–783. https://doi.org/10.1080/02701367.2011.10599814

Parra Saldías, M., Mayorga-Vega, D., López-Fernández, I., & Viciana, J. (2018). How many daily steps are really enough for adolescents? A cross-validation study. *Retos, 33*, 241–246. https://doi.org/10.47197/retos.v0i33.55504

Poitras, V. J., Gray, C. E., Borghese, M. M., Carson, V., Chaput, J. P., Janssen, I., Katzmarzyk, P. T., Pate, R. R., Connor Gorber, S., Kho, M. E., Sampson, M., & Tremblay, M. S. (2016). Systematic review of the relationships between objectively measured physical activity and health indicators in school-aged children and youth. *Applied physiology, nutrition, and metabolism*, *41*(6), S197–S239. https://doi.org/10.1139/apnm-2015-0663

Redenius, N., Kim, Y., & Byun, W. (2019). Concurrent validity of the Fitbit for assessing sedentary behavior and moderate-to-vigorous physical activity. *BMC medical research methodology*, *19*(1), 29. https://doi.org/10.1186/s12874-019-0668-1

Romanzini, M., Petroski, E. L., Ohara, D., Dourado, A. C., & Reichert, F. F. (2014). Calibration of ActiGraph GT3X, Actical and RT3 accelerometers in adolescents. *European journal of sport science*, *14*(1), 91–99. https://doi.org/10.1080/17461391.2012.732614

Rowlands, A. V., Harrington, D. M., Bodicoat, D. H., Davies, M. J., Sherar, L. B., Gorely, T., Khunti, K., & Edwardson, C. L. (2018). Compliance of Adolescent Girls to Repeated Deployments of Wrist-Worn Accelerometers. *Medicine and science in sports and exercise*, *50*(7), 1508–1517. https://doi.org/10.1249/MSS.0000000000001588

Schneider, M., & Chau, L. (2016). Validation of the Fitbit Zip for monitoring physical activity among free-living adolescents. *BMC research notes*, *9*(1), 448. https://doi.org/10.1186/s13104-016-2253-6

Shephard, R., & Tudor-Locke, C. (Eds.). (2016). *The objective monitoring of physical activity: Contributions of accelerometry to epidemiology, exercise science and rehabilitation*. Springer.

Šimůnek, A., Dygrýn, J., Jakubec, L., Neuls, F., Frömel, K., & Welk, G. J. (2019). Validity of Garmin Vívofit 1 and Garmin Vívofit 3 for School-Based Physical Activity Monitoring. *Pediatric exercise science*, *31*(1), 130–136. https://doi.org/10.1123/pes.2018-0019

Stamm, A., & Hartanto, R. (2018). Feature extraction from mems accelerometer and motion tracking measurements in comparison with smart bands during running. *Proceedings*, *2*(6), 197-202. https://doi.org/10.3390/proceedings2060197

Stewart, A., Marfell-Jones, M., Olds, T., & De Ridder, J. (2011). *International standards for anthropometric assessment*. International Society for the Advancement of Kinanthropometry.

Strath, S. J., & Rowley, T. W. (2018). Wearables for Promoting Physical Activity. *Clinical chemistry*, *64*(1), 53–63. https://doi.org/10.1373/clinchem.2017.272369

Toth, L. P., Park, S., Springer, C. M., Feyerabend, M. D., Steeves, J. A., & Bassett, D. R. (2018). Video-Recorded Validation of Wearable Step Counters under Free-living Conditions. *Medicine and science in sports and exercise*, *50*(6), 1315–1322. https://doi.org/10.1249/MSS.0000000000001569

Trost, S. G., Loprinzi, P. D., Moore, R., & Pfeiffer, K. A. (2011). Comparison of accelerometer cut points for predicting activity intensity in youth. *Medicine and science in sports and exercise*, *43*(7), 1360–1368. https://doi.org/10.1249/MSS.0b013e318206476e

Tudor-Locke, C., Craig, C. L., Beets, M. W., Belton, S., Cardon, G. M., Duncan, S., Hatano, Y., Lubans, D. R., Olds, T. S., Raustorp, A., Rowe, D. A., Spence, J. C., Tanaka, S., & Blair, S. N. (2011). How many steps/day are enough? for children and adolescents. *The international journal of behavioral nutrition and physical activity*, *8*, 78. https://doi.org/10.1186/1479-5868-8-78

Vailshery, L. S. (2021). *Wearables Statistics & facts*. Technology & Telecommunications. https://www.statista.com/topics/1556/wearable-technology

Van Hecke, L., Loyen, A., Verloigne, M., van der Ploeg, H. P., Lakerveld, J., Brug, J., De Bourdeaudhuij, I., Ekelund, U., Donnelly, A., Hendriksen, I., Deforche, B., & DEDIPAC consortium (2016). Variation in population levels of physical activity in European children and adolescents according to cross-European studies: a systematic literature review within DEDIPAC. *The international journal of behavioral nutrition and physical activity*, *13*, 70. https://doi.org/10.1186/s12966-016-0396-4

Voss, C., Gardner, R. F., Dean, P. H., & Harris, K. C. (2017). Validity of Commercial Activity Trackers in Children With Congenital Heart Disease. *The Canadian journal of cardiology*, *33*(6), 799–805. https://doi.org/10.1016/j.cjca.2016.11.024

Westerterp, K. R. (2013). Physical activity and physical activity induced energy expenditure in humans: measurement, determinants, and effects. *Frontiers in physiology*, *4*, 90. https://doi.org/10.3389/fphys.2013.00090

Westerterp, K. R. (2017). Doubly labelled water assessment of energy expenditure: principle, practice, and promise. *European journal of applied physiology*, *117*(7), 1277–1285. https://doi.org/10.1007/s00421-017-3641-x

Willmott, C., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, *30*(1), 79–82. https://doi.org/10.3354/cr030079

World Health Organization. (2020). *WHO guidelines on physical activity and sedentary behaviour*. World Health Organization.

Yang, X., Jago, R., Zhang, Q., Wang, Y. Y., Zhang, J., & Zhao, W. H. (2019). Validity and Reliability of the Wristband Activity Monitor in Free-living Children Aged 10-17 Years. *Biomedical and environmental sciences*, *32*(11), 812–822. https://doi.org/10.3967/bes2019.103

LIMITATIONS AND FUTURE DIRECTIONS

[LIMITACIONES Y ESTUDIOS FUTUROS]

LIMITATIONS AND FUTURE DIRECTIONS [LIMITACIONES Y ESTUDIOS FUTUROS]

The main limitations and future research studies of the present Doctoral Thesis are summarized below. For further information of any study, please check the corresponding section of each paper.

Papers I, III-V, and *VIII-XI*: Firstly, all the intervention and validity studies were conducted using a convenience sample, limiting the obtained outcomes' generalizability to the particular studied setting. However, due to human, time, and material resource restrictions, a probabilistic and larger sample could not be examined. Therefore, future research studies should include a larger sample involving a wide range of students' ages and characteristics, which provide a higher generalization of the obtained outcomes in the present papers.

Papers I, and *III-V*: Regarding intervention studies, the length of the program may have been a limitation to achieve more significant effects on the objectively measured PA and SB levels. However, given the large volume of objectives that have to be developed throughout the school year with a very limited time for the PE subject, the intervention studies were adjusted to the mean length of PE teaching units (Hardman et al., 2014). Secondly, regarding the contents, the innovative programs were developed with physical fitness and traditional and alternative games and sports contents, which dominate the students' PA experiences globally in PE (Hardman et al., 2014) and could be the most applicable to students' free-time (individually or in small groups with friends), but these effects should also be studied with other PE contents. Therefore, further studies are suggested for verifying the findings obtained in *Papers I* and *III-V* in several different contexts (e.g., other PE contents, older and earlier ages of participants, or more extended interventions). Specifically, according to the recent systematic review published by Neil-Sztramko et al. (2021), longer school-based interventions may be needed (around 12 weeks and up) in order to achieve real changes in habitual PA and SB levels.

Moreover, new applications of the innovative teaching units, such as the irregular or alternated teaching units tested in the present Doctoral Thesis, are necessary to provide further empirical evidence of their effectiveness in promoting students' healthy PA and SB levels. For example, future irregular teaching units, such as that carried out in *Paper III*, should include explicit strategies of PA promotion (e.g., advice about PA benefits and problems related to SB, goal setting strategies, or the resolution of barriers to PA practice) which make students feel that they are making an informed decision about their health and, therefore, they may be more likely to change their behavior in order to pursue long-term improvements in PA levels and not only during the intervention (Jago et al., 2015; Neil-Sztramko et al., 2021). For example, these orientations could be included throughout the season phase, sharing the time spent in fair-play analysis during PE lessons with this counseling time.

Besides that, in order to achieve significant long-term benefits for increasing students' habitual objectively measured PA levels and reduce their SB, it may be needed further studies complementing these innovative teaching units in *Papers III and IV-V* with a structured PA plan for their leisure time (Neil-Sztramko et al., 2021). For example, complementing the irregular teaching unit carried out in *Paper III* with the maintenance of voluntary structured programs at school recess during the whole academic year. It could allow checking whether the increase in students' PA levels happens only in the first weeks or if it remains for a desirably prolonged period to achieve sustainable change in students' habitual PA levels. Moreover, complementing the alternated program carried out in *Paper V* with a proposal of an extracurricular PA plan for leisure time (e.g., school recess, weekends or after-school time) could also help to achieve significant benefits in objectively measured students' habitual PA and SB levels. In this sense, an interesting future study would be to carry out an experimental and controlled study about the intermittent teaching unit proposed in *Paper VI*, which includes an adaptive plan of individualized daily PA goals for students (e.g., from 5,000 steps per day in week 1 to achieving 10,000 steps per day in week 10), as well as their daily PA monitoring with the help of activity trackers, to test their effectiveness for promoting students' healthy PA and SB levels.

Finally, another limitation is the non-compliance rate in objectively measured PA variables, widely acknowledged as one of the most important methodological limitations with accelerometry (Shephard & Tudor-Locke, 2016). However, as analyses showed no differences in students' general characteristics in included and excluded participants, the non-compliance rate may not bias the findings. Furthermore, the non-compliance rates values found are typical for this kind of studies (Howie & Straker, 2016).

Paper VII: Firstly, although randomized controlled trials have higher methodological quality, the present systematic review included several study designs. Nevertheless, sensitivity analyses were also performed comparing randomized controlled trials and non-randomized trials, showing no differences in students' daily MVPA, total PA, and SB levels. Secondly, although the inclusion of a wide range of intervention types, populations, sample size, and study designs had some advantages regarding the generalizability of conclusions, it supposes a high level of heterogeneity. Therefore, it makes the independent contribution of any intervention features and, thus, the establishment of solid intervention conclusions challenging to establish. However, subgroup analyses and meta-regression of the a priori hypothesized moderators were also performed in addition to the overall effect size. Therefore, not only general effect results are provided, but also for each specific group based on the characteristics of the programs and students. Furthermore, the present systematic review investigated effectiveness right at the end of the program (i.e., short-term), but future studies should investigate long-term effectiveness to assess actual behavioral changes some months after the program, and not only short-term effects. However, due to the very limited evidence, it was not performed in the present systematic review. Moreover, coding some study outcomes was problematic due to authors not reporting them. Although authors were contacted, many of them did not reply, and the particular study outcome had to be omitted. However, this is a common problem in most systematic reviews, and a great effort was made in contacting authors, recalculating data, or estimating values from figures. Finally, in some cases, consumer-based activity trackers were both used as a motivational instrument during the intervention and to objectively measure PA, which could affect results by increasing their actual PA levels in the control group or during baseline assessments.

Therefore, as a result of the findings obtained in *Paper VII*, further research is needed to determine the effectiveness of consumer-wearable activity tracker-based programs among school-aged children using robust designs with low risk of bias, and which compare the effect of different intervention characteristics in the same study. For example, future studies could include different experimental groups that compared some intervention characteristics (e.g., one experimental group including counseling and another without counseling; or one experimental group including goals and another without goals) to establish causal-effect relationships between intervention characteristics and the effect of the programs.

Paper VIII: Due to the nature of the study, the CEPAF test was validated in a specific setting, which may limit its generalizability and applicability to other contexts. Therefore, future research studies could design and validate other knowledge test of the environment for practicing PA but using general environments and spaces. That is, focusing the questions on objects or environments commonly present in every place or city (e.g., hills, ramps, stairs, banks, or avenues, without specifying their location). This would allow higher applicability of the built test so PE teachers from different places can use it.

Paper IX: The measurements were transversal applying the adapted SSI-PE only during one specific PE class (i.e., SSI-PE-lesson), and possible changes between lessons were not considered in this study. Therefore, further applications of this questionnaire applying the construct validity analysis (e.g., between two different PE contents/activities) in longitudinal studies should be considered. Moreover, feasible future research from *Paper IX* could be to adapt the SSI-PE-lesson to a non-paper based format using new technologies (e.g., Additio or Socrative gamification apps) in order to obtain immediate feedback from the students on each part of the lesson (i.e., warm-up, main part, and cool-down) using their smartphones or educative tablets to answer the questionnaire.

Furthermore, the scale could also be adapted to be more attractive for younger students (e.g., using thumbs up or down, or happy or sad faces), improving their understanding and motivation. Besides that, the SSI-PE and SSI-PE-lesson questionnaires only assess whether the students are having fun or not. However, future studies could design and validate other questionnaires that answer the question "Why is the student having fun?" in order to continue studying the conceptualization of the term enjoyment. This questionnaire would be focused on identifying which elements of the PE lesson make the student enjoy or get bored during the lesson so that the teacher could use their answers to improve students' motivation towards PE and, ultimately, toward the PA practice.

Papers X-XI: Firstly, although Actigraph accelerometers have been highlighted as the most common and valid method for objectively assessing high-school students' PA levels during free-living conditions (e.g., Colley et al., 2019; or Romanzini et al., 2014), today there is no solid evidence-based consensus about many methodological issues (e.g., device placement, sample rate, epoch length, non-wear period definition, minimum wear time per day or cut-off points; Migueles et al., 2017). Therefore, it may contribute to the variability of wristbands validity results, although the best current evidence-based decisions were adopted (Migueles et al., 2017). Lastly, another limitation is related to the fact that brands (e.g., Xiaomi, Apple, or Samsung) did not provide transparent information about their algorithms to classify PA at different intensities (i.e., cut-off points), which may bias the present results.

Therefore, further research investigating consumer-wearable activity trackers' validity and reliability for measuring steps and intensity-related PA levels among high-school students is needed, especially if brands provided transparent information about their algorithms to classify PA at different intensities, to obtain more robust and reliable results. Moreover, these studies have investigated consumer-wearable activity trackers' validity only under unstructured free-living settings for measuring students' daily steps and intensity-related PA levels (*Papers X-XI*), and under structured free-living settings for measuring students' daily steps (*Paper XI*). Therefore, further studies should assess their validity for measuring students' intensity-related PA levels under a structured free-living environment and measuring students' daily steps and intensity-related PA levels under laboratory conditions.

CONCLUSIONS

[CONCLUSIONES]

CONCLUSIONS

The overall results of the present Doctoral Thesis show the effectiveness of innovative teaching units for improving some factors influencing the acquisition of students' healthy PA and SB habits, overcoming some limitations related to planning in PE. However, improvements in objectively measured PA and SB levels are not achieved after the application of innovative short-term PE-based programs. Therefore, other strategies such as proposed longer programs, complementing them with other extracurricular PA plans for leisure time, structured programs for recess, or including consumer-wearable activity trackers for providing students real-time feedback seem necessary to improve students' habitual objective PA and SB levels. This knowledge could help PE teachers to design effective and feasible intervention programs to increase students' daily PA levels and reduce SB from the PE setting.

Specific conclusions

The specific conclusions of the present Doctoral Thesis were the following:

I. A short-term Sport Education-based program carried out during 12 PE lessons only improves students' desire and willingness to participate in sports competitions, but it is not enough to improve students' self-reported nor objective habitual PA and SB levels. Therefore, in addition to the Sport Education Model, the PE teachers should include other specific strategies focused on transference students' PA practice from the PE classroom to the out-of-school context. (*Paper I*)

II. The proposal of innovative teaching units (intermittent, alternated, reinforced, and irregular) could help PE teachers to design effective and viable programs to increase habitual PA and decrease habitual SB levels among students, overcoming some limitations related to planning in PE. However, future experimental and controlled studies are needed to test their effectiveness among high-school students. (*Paper II and VI*)

III. An irregular teaching unit based on the Sport Education Model during PE lessons and using school recess to develop the formal competition phase effectively increases objectively measured high-school students' PA levels and reduces SB levels during this extracurricular period. Moreover, students were interested in continue participating in sport-based activities during school recess, showing that these programs are feasible and in accord with students' interests, although their PA and SB levels during school recess reverted back to the baseline values after the intervention. Therefore, it could help PE teachers to design effective and feasible programs that increase high-school students' PA levels and reduce SB levels during school recess and contribute to the daily PA and SB recommendation and their consequent health benefits on high-school students. However, it is desirable to convert these short-term programs into continuous voluntary programs, structured and guided by qualified PE teachers at school recess to avoid losing the gain obtained during the intervention in students' PA and SB levels. (*Paper III*)

IV. Alternated inside-outside school teaching units are effective for improving students' knowledge about the environment for practicing PA, perceived autonomy, enjoyment in PE, autonomous motivation towards PA, the intention to be physically active, and self-reported PA, although they did not improve objectively measured PA nor SB levels. Therefore, the application of alternated teaching units could effectively improve factors influencing students' PA behavior (e.g., making them competent for practicing PA masterfully on their own during leisure time, or improving their intention to be physically active). However, complementing the intervention with a proposal of an extracurricular structured PA plan for after-school time or weekends may be necessary to achieve real changes in students' objectively measured PA and SB levels. (*Papers IV and V*)

V. Consumer-wearable activity tracker-based programs seem to effectively promote students' daily total steps and MVPA levels, especially for females and those that are physically inactive before the intervention program. These programs should include specific goal-setting, educational counseling, and wrist-worn trackers as especially effective strategies. However, due to the certainty of evidence being from "Low" to "Moderate", future well-designed primary research studies about the topic are needed. (*Paper VII*)

VI. The designed CEPAF test is a valid and reliable measuring instrument to gather information about the knowledge of the immediate environment for the students' practice of PA. The present study represents a significant contribution to check whether PE subject is providing students with the necessary tools to be active autonomously outside the school, as well as to check the effectiveness of intervention programs to promote PA practice. (*Paper VIII*)

VII. The adapted versions of the SSI-PE referred to both the PE subject and only one PE class are valid and reliable scales to be applied among high-school students. Furthermore, having valid tools to assess students' enjoyment during PE lessons allows comparing between teaching methodologies or activities. This provides an opportunity to know which ones are more interesting for students and, therefore, PE teachers could adapt their teaching process to improve students' autonomous motivation and ultimately promote their PA practice. (*Paper IX*)

VIII. The mobile apps Pedometer, Pacer, Google Fit and Apple Health, and the wristbands Samsung Watch Active 2, Apple Watch Series 5, and Xiaomi Mi Band 2, 3, 4, and 5 seem valid activity trackers for monitoring high-school students' steps under free-living conditions. Moreover, although caution is warranted when using these wristbands for assessing students' total PA and MVPA levels, the three wristbands accurately classified adolescents as meeting or not meeting the recommendation of 10,000 steps per day, and the Xiaomi Mi Band and Apple Watch Series 5 wristbands also for the recommendation of 60 minutes of MVPA per day. Therefore, wristbands are a feasible alternative to the research-grade accelerometers, including for being used by teachers in the PE setting. Furthermore, they could be used during healthy habits promotion programs to provide accurate feedback to students (especially for steps output), as well as to set specific goals based on 10,000 daily steps and/or 60 minutes of MVPA per day to ensure their accomplishment with the international recommendations. (*Papers X and XI*)

CONCLUSIONES

Los resultados principales de la presente Tesis Doctoral muestran la efectividad de las unidades didácticas innovadoras para mejorar algunos factores que influyen en la adquisición de hábitos saludables de AF y CS de los escolares, superando diversas limitaciones relacionadas con la planificación de la asignatura de EF. Sin embargo, no se han conseguido mejoras en los niveles de AF y CS de los escolares medidos de forma objetiva tras la aplicación de estos programas innovadores de corta duración en la asignatura de EF. Por lo tanto, para mejorar los niveles objetivos de AF habitual y CS de los estudiantes, parece necesario aplicar otras estrategias adicionales, tales como proponer programas más largos, complementarlos con otros planes de AF extracurriculares para realizar en su tiempo libre o incluir los monitores portátiles de fitness como herramienta motivacional para proporcionar a los estudiantes retroalimentación en tiempo real sobre sus niveles de AF y CS. Este conocimiento podría ayudar a los profesores de EF a diseñar programas de intervención viables y eficaces para aumentar los niveles de AF diaria y reducir la CS de los escolares desde la asignatura de EF.

Conclusiones específicas

Las conclusiones específicas de la presente Tesis Doctoral fueron las siguientes:

I. Un programa de corta duración llevado a cabo durante 12 sesiones de EF siguiendo el Modelo de Educación Deportiva sólo mejora el deseo y la voluntad del alumnado para participar en competiciones deportivas, pero no es suficiente para mejorar los niveles auto-reportados ni objetivos de AF habitual y de CS de los escolares. Por lo tanto, además del Modelo de Educación Deportiva, los profesores de EF deberían incluir otras estrategias específicas centradas en conseguir la transferencia de la práctica de AF del alumnado desde el aula de EF a su tiempo libre. (*Artículo I*)

II. La propuesta de unidades didácticas innovadoras (intermitentes, alternadas, irregulares y reforzadas) podría ayudar a los profesores de EF a diseñar programas efectivos y viables para incrementar los niveles de AF habitual y disminuir la CS de los escolares, superando algunas limitaciones relacionadas con la planificación en EF. Sin embargo, se necesitan futuros estudios experimentales y controlados para comprobar su eficacia en estudiantes de Educación Secundaria. (*Artículo II y VI*)

III. Una unidad didáctica irregular siguiendo el Modelo de Educación Deportiva durante las clases de EF, y utilizando el recreo escolar para desarrollar la fase de competición formal del modelo, es eficaz para incrementar los niveles objetivos de AF y reducir la CS del alumnado de Educación Secundaria durante este periodo extraescolar. Además, los escolares se mostraban interesados en continuar participando en actividades deportivas durante el recreo escolar, lo que demuestra que estos programas son viables y se ajustan a los intereses de los estudiantes, aunque sus niveles de AF y CS durante el recreo escolar volvieron a los valores basales después de la intervención. Por lo tanto, estos resultados podrían ayudar a los profesores de EF a diseñar programas eficaces y viables que permitan incrementar los niveles de AF de los estudiantes de Educación Secundaria durante el recreo escolar y contribuir así a alcanzar las recomendaciones diarias de AF y CS y obtener sus consiguientes beneficios en la salud de los estudiantes. Sin embargo, sería deseable convertir estos programas de corta duración en programas voluntarios continuos, estructurados y guiados por los profesores de EF, durante los recreos escolares para evitar que se pierda la ganancia obtenida durante la intervención en los niveles de AF y CS del alumnado. (*Artículo III*)

IV. Las unidades didácticas alternadas dentro-fuera del centro escolar son efectivas para mejorar el conocimiento de los estudiantes de Educación Secundaria sobre su entorno para la práctica de AF, la autonomía percibida, el disfrute durante la EF, la motivación autónoma hacia la AF, la intención de ser físicamente activo y los niveles de AF auto-reportados, aunque no mejoraron los niveles de AF ni de CS medidos objetivamente. Por lo tanto, la aplicación de unidades didácticas alternadas podría ser una forma efectiva de mejorar algunos factores que influyen en el comportamiento de AF de los estudiantes (por ejemplo, hacerlos competentes para practicar AF de forma autónoma durante su tiempo libre, o mejorar la intención de ser físicamente activos). Sin embargo, parece necesario complementar esta intervención con una propuesta de plan de AF estructurado para el tiempo extraescolar o los fines de semana para lograr cambios reales en los niveles objetivos de AF y CS de los estudiantes. (*Artículos IV y V*)

V. Los programas de intervención que incluyen monitores portátiles de fitness parecen ser eficaces para incrementar los pasos diarios y los niveles de AFMV de los escolares, especialmente para las mujeres y los que son físicamente inactivos antes de la intervención. Estos programas deben incluir el establecimiento de objetivos específicos, sesiones de asesoramiento educativo y la utilización de monitores de muñeca como estrategias especialmente eficaces. Sin embargo, debido a que la certeza de la evidencia es de "Baja" a "Moderada", se necesitan futuros estudios de investigación primarios bien diseñados sobre el tema. (*Artículo VII*)

VI. La prueba CEPAF es un instrumento de medida válido y fiable para recoger información sobre el conocimiento que poseen los escolares de Educación Secundaria de su entorno próximo para la práctica de AF. El presente estudio representa una importante contribución para comprobar si la asignatura de EF está proporcionando a los alumnos las herramientas necesarias para ser activos fuera de la escuela autónomamente, así como para comprobar la eficacia de los programas de intervención centrados en la promoción de la práctica de AF. (*Artículo VIII*)

VII. Las versiones adaptadas del *Sport Satisfaction Instrument*, referidas tanto a la asignatura de EF en general como a una única clase de EF, son escalas válidas y fiables para ser aplicadas con estudiantes de Educación Secundaria. Disponer de herramientas válidas para evaluar el disfrute de los alumnos durante las clases de EF permite comparar entre metodologías de enseñanza o actividades. Esto permite conocer cuáles son más interesantes para los alumnos y, por tanto, que el profesor de EF pueda adaptar su proceso de enseñanza para mejorar la motivación autónoma del alumnado y, en última instancia, promover la práctica de AF. (*Artículo IX*)

VIII. Las aplicaciones móviles Pedometer, Pacer, Google Fit y Apple Health, y las pulseras de fitness Samsung Watch Active 2, Apple Watch Series 5 y Xiaomi Mi Band 2, 3, 4 y 5 parecen ser dispositivos válidos para monitorizar los pasos realizados por los estudiantes de Educación Secundaria en condiciones de vida libre. Además, aunque se debe tener cuidado al utilizar estas pulseras de fitness para evaluar los niveles de AF total y AFMV de los estudiantes, las tres pulseras de fitness clasifican correctamente a los adolescentes como que cumplen o no la recomendación de 10.000 pasos al día, así como, las pulseras Xiaomi Mi Band y Apple Watch Series 5 también para la recomendación de 60 minutos de AFMV al día. Por lo tanto, estas pulseras de fitness son una alternativa viable a los acelerómetros de investigación, incluso para ser utilizados por los profesores en la asignatura de EF. Además, podrían utilizarse durante los programas de promoción de hábitos saludables para proporcionar información precisa a los estudiantes (especialmente en cuanto al recuento de pasos), así como para establecer objetivos específicos basados en 10.000 pasos diarios y/o 60 minutos de AFMV al día para garantizar el cumplimiento de las recomendaciones internacionales de AF. (*Artículos X y XI*)

REFERENCES

[REFERENCIAS BIBLIOGRÁFICAS]

REFERENCES [REFERENCIAS BIBLIOGRÁFICAS]

Abós, Á., Sevil, J., Julián, J. A., Abarca-Sos, A., & García-González, L. (2017). Improving students' predisposition towards physical education by optimizing their motivational processes in an acrosport unit. *European Physical Education Review*, *23*(4), 444-460. https://doi.org/10.1177/1356336X16654390

ActiGraph Corporation. (2021). *Activity monitors comparisons*. http://actigraphcorp.com

Adamo, K. B., Prince, S. A., Tricco, A. C., Connor-Gorber, S., & Tremblay, M. (2009). A comparison of indirect versus direct measures for assessing physical activity in the pediatric population: a systematic review. *International journal of pediatric obesity*, *4*(1), 2–27. https://doi.org/10.1080/17477160802315010

Ainsworth, B., Cahalin, L., Buman, M., & Ross, R. (2015). The current state of physical activity assessment tools. *Progress in cardiovascular diseases*, *57*(4), 387–395. https://doi.org/10.1016/j.pcad.2014.10.005

Alley, S., Schoeppe, S., Guertler, D., Jennings, C., Duncan, M. J., & Vandelanotte, C. (2016). Interest and preferences for using advanced physical activity tracking devices: results of a national cross-sectional survey. *BMJ open*, *6*(7), e011243. https://doi.org/10.1136/bmjopen-2016-011243

Altenburg, T. M., Kist-van Holthe, J., & Chinapaw, M. J. (2016). Effectiveness of intervention strategies exclusively targeting reductions in children's sedentary time: a systematic review of the literature. *The international journal of behavioral nutrition and physical activity*, *13*, 65. https://doi.org/10.1186/s12966-016-0387-5

Association for Physical Education. (2020). *Health Position Paper*. Association for Physical Education.

Bai, Y., Allums-Featherston, K., Saint-Maurice, P. F., Welk, G. J., & Candelaria, N. (2018). Evaluation of Youth Enjoyment Toward Physical Activity and Sedentary Behavior. *Pediatric exercise science*, *30*(2), 273–280. https://doi.org/10.1123/pes.2017-0101

Bandura, A. (2004). Health promotion by social cognitive means. *Health education & behavior, 31*(2), 143-164.

Barkoukis, V., Hagger, M. S., Lambropoulos, G., & Tsorbatzoudis, H. (2010). Extending the trans-contextual model in physical education and leisure-time contexts: examining the role of basic psychological need satisfaction. *The*

British journal of educational psychology, 80(4), 647–670. https://doi.org/10.1348/000709910X487023

Bessa, C., Hastie, P., Araújo, R., & Mesquita, I. (2019). What do we know about the development of personal and social skills within the sport education model: A systematic review. *Journal of Sports Science & Medicine, 18*(4), 812-829.

Borde, R., Smith, J. J., Sutherland, R., Nathan, N., & Lubans, D. R. (2017). Methodological considerations and impact of school-based interventions on objectively measured physical activity in adolescents: a systematic review and meta-analysis. *Obesity reviews : an official journal of the International Association for the Study of Obesity, 18*(4), 476–490. https://doi.org/10.1111/obr.12517

Brickwood, K. J., Watson, G., O'Brien, J., & Williams, A. D. (2019). Consumer-Based Wearable Activity Trackers Increase Physical Activity Participation: Systematic Review and Meta-Analysis. *JMIR mHealth and uHealth, 7*(4), e11819. https://doi.org/10.2196/11819

Bronikowski, M., Bronikowska, M., & Glapa, A. (2016). Do They Need Goals or Support? A Report from a Goal-Setting Intervention Using Physical Activity Monitors in Youth. *International journal of environmental research and public health, 13*(9), 914. https://doi.org/10.3390/ijerph13090914

Brooke, H. L., Corder, K., Atkin, A. J., & van Sluijs, E. M. (2014). A systematic literature review with meta-analyses of within- and between-day differences in objectively measured physical activity in school-aged children. *Sports medicine, 44*(10), 1427–1438. https://doi.org/10.1007/s40279-014-0215-5

Burgueño, R., Cueto-Martín, B., Morales-Ortiz, E., & Medina-Casaubón, J. (2020). Influence of sport education on high school students' motivational response: A gender perspective [Influencia de la educación deportiva sobre la respuesta motivacional del alumnado de bachillerato: Una perspectiva de género]. *Retos, 37*, 546-555. https://doi.org/10.47197/retos.v37i37.70880

Carson, R. L., Castelli, D. M., Beighle, A., & Erwin, H. (2014). School-based physical activity promotion: a conceptual framework for research and practice. *Childhood obesity, 10*(2), 100–106. https://doi.org/10.1089/chi.2013.0134

Carson, V., Hunter, S., Kuzik, N., Gray, C. E., Poitras, V. J., Chaput, J. P., Saunders, T. J., Katzmarzyk, P. T., Okely, A. D., Connor Gorber, S., Kho, M. E., Sampson, M., Lee, H., & Tremblay, M. S. (2016). Systematic review of sedentary behaviour and health indicators in school-aged children and youth: an update. *Applied physiology, nutrition, and metabolism, 41*(6), S240–S265. https://doi.org/10.1139/apnm-2015-0630

Caspersen, C. J., Powell, K. E., & Christenson, G. M. (1985). Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public health reports, 100*(2), 126-131.

Centers for Disease Control and Prevention & Society of Health and Physical Educators. (2017). S*trategies for recess in schools*. Centers for Disease Control and Prevention.

Chaput, J. P., Willumsen, J., Bull, F., Chou, R., Ekelund, U., Firth, J., Jago, R., Ortega, F. B., & Katzmarzyk, P. (2020). 2020 WHO guidelines on physical activity and sedentary behaviour for children and adolescents aged 5–17 years: summary of the evidence. *International Journal of Behavioral Nutrition and Physical Activity 17,* 141. https://doi.org/10.1186/s12966-020-01037-z

Chatzisarantis, N. L., Hagger, M. S., & Brickell, T. (2008). Using the construct of perceived autonomy support to understand social influence within the theory of planned behavior. *Psychology of Sport and Exercise, 9*(1), 27-44. https://doi.org/10.1016/j.psychsport.2006.12.003

Cheon, S. H., Reeve, J., Lee, Y., & Lee, J. W. (2018). Why autonomy-supportive interventions work: Explaining the professional development of teachers' motivating style. *Teaching and Teacher Education, 69,* 43-51. https://doi.org/10.1016/j.tate.2017.09.022

Chicote-López, J., Abarca-Sos, A., Gallardo, L. O., & García-González, L. (2018). Social antecedents in physical activity: Tracking the self-determination theory sequence in adolescents. *Journal of Community Psychology, 46*(3), 356-373. https://doi.org/10.1002/jcop.21945

Chu, T. L., & Zhang, T. (2018). Motivational processes in Sport Education programs among high school students: A systematic review. *European Physical Education Review, 24,* 372-394. https://doi.org/10.1177/1356336X17751231

Colley, R. C., Butler, G., Garriguet, D., Prince, S. A., & Roberts, K. C. (2019). Comparison of self-reported and accelerometer-measured physical activity among Canadian youth. *Health reports, 30*(7), 3–12. https://doi.org/10.25318/82-003-x201900700001-eng

Coolkens, R., Ward, P., Seghers, J., & Iserbyt, P. (2018). Effects of Generalization of Engagement in Parkour from Physical Education to Recess on Physical Activity. *Research quarterly for exercise and sport, 89*(4), 429–439. https://doi.org/10.1080/02701367.2018.1521912

Cooper, H., Hedges, L. V., Valentine, J.C. (2019). *The Handbook of Research Synthesis and Meta-Analysis. 3rd ed*. Russell Sage Foundation.

Cosoli, G., Spinsante, S., & Scalise, L. (2020). Wrist-worn and chest-strap wearable devices: Systematic review on accuracy and metrological characteristics. *Measurement, 159,* 107789. https://doi.org/10.1016/j.measurement.2020.107789

Cox, A., Smith, A., & Williams, L. (2008). Change in physical Education motivation and physical activity behavior during middle school. *The Journal of Adolescent Health, 43*(5), 506-513. https://doi.org/10.1016/j.jadohealth.2008.04.020

Cuevas, R., García-López, L. M., & Serra-Olivares, J. (2016). Sport education model and self-determination theory: An intervention in secondary school children. *Kinesiology, 48*(1), 30–38. https://doi.org/10.26582/k.48.1.15

Da Silva, M. P., Fontana, F. E., Callahan, E., Mazzardo, O., & De Campos, W. (2015). Step-Count Guidelines for Children and Adolescents: A Systematic Review. *Journal of physical activity & health, 12*(8), 1184–1191. https://doi.org/10.1123/jpah.2014-0202

Deci, E. L., & Ryan, R. M. (2008). Facilitating optimal motivation and psychological wellbeing across life's domains. *Canadian Psychology, 49,* 14–23. https://doi.org/10.1037/0708-5591.49.1.14

Department of Education and Sports of the Andalusian Regional Government. (2021). Order of January 15, 2021, which develops the curriculum corresponding to the Compulsory Secondary Education stage in the Autonomous Community of Andalusia [Orden de 15 de enero de 2021, por la que se desarrolla el currículo correspondiente a la etapa de Educación Secundaria Obligatoria en la Comunidad Autónoma de Andalucía]. *Official Gazette of the Andalusian Regional Government, 7,* 656-793.

Dhurandhar, N. V., Schoeller, D., Brown, A. W., Heymsfield, S. B., Thomas, D., Sørensen, T. I., Speakman, J. R., Jeansonne, M., Allison, D. B., & Energy Balance Measurement Working Group (2015). Energy balance measurement:

when something is not better than nothing. *International journal of obesity, 39*(7), 1109–1113. https://doi.org/10.1038/ijo.2014.199

Duncan, M., Birch, S., & Woodfield, L. (2012). Efficacy of an integrated school curriculum pedometer intervention to enhance physical activity and to reduce weight status in children. *European Physical Education Review, 18*(3), 396–407. https://doi.org/10.1177/1356336X12450799

Dyrstad, S. M., Hansen, B. H., Holme, I. M., & Anderssen, S. A. (2014). Comparison of self-reported versus accelerometer-measured physical activity. *Medicine and science in sports and exercise*, *46*(1), 99–106. https://doi.org/10.1249/MSS.0b013e3182a0595f

Ekelund, U., Tomkinson, G., & Armstrong, N. (2011). What proportion of youth are physically active? Measurement issues, levels and recent time trends. *British journal of sports medicine*, *45*(11), 859–865. https://doi.org/10.1136/bjsports-2011-090190

Ennis, C. D. (2015). Knowledge, transfer, and innovation in physical literacy curricula. *Journal of Sport and Health Science, 4*(2), 119–124. https://doi.org/10.1016/j.jshs.2015.03.001

European Commission/EACEA/Eurydice. (2013). *Physical Education and sport at school in Europe Eurydice Report*. Publications Office of the European Union.

Evangelio, C., Sierra Díaz, J., González Víllora, S., & Fernández-Río, F. J. (2018). The sport education model in elementary and secondary education: A systematic review. *Movimento, 24*(3), 931-946. https://doi.org/10.22456/1982-8918.81689

Evans, E. W., Abrantes, A. M., Chen, E., & Jelalian, E. (2017). Using Novel Technology within a School-Based Setting to Increase Physical Activity: A Pilot Study in School-Age Children from a Low-Income, Urban Community. *BioMed research international*, *2017,* 4271483. https://doi.org/10.1155/2017/4271483

Evenson, K. R., Goto, M. M., & Furberg, R. D. (2015). Systematic review of the validity and reliability of consumer-wearable activity trackers. *The international journal of behavioral nutrition and physical activity*, *12,* 159. https://doi.org/10.1186/s12966-015-0314-1

Eyre, E. L., Cox, V. M., Birch, S. L., & Duncan, M. J. (2016). An integrated curriculum approach to increasing habitual physical activity in deprived

South Asian children. *European journal of sport science, 16*(3), 381–390. https://doi.org/10.1080/17461391.2015.1062565

Franssen, W., Franssen, G., Spaas, J., Solmi, F., & Eijnde, B. O. (2020). Can consumer wearable activity tracker-based interventions improve physical activity and cardiometabolic health in patients with chronic diseases? A systematic review and meta-analysis of randomised controlled trials. *The international journal of behavioral nutrition and physical activity, 17*(1), 57. https://doi.org/10.1186/s12966-020-00955-2

Friedrich, R. R., Polet, J. P., Schuch, I., & Wagner, M. B. (2014). Effect of intervention programs in schools to reduce screen time: a meta-analysis. *Jornal de pediatria, 90*(3), 232–241. https://doi.org/10.1016/j.jped.2014.01.003

Fuller, D., Colwell, E., Low, J., Orychock, K., Tobin, M. A., Simango, B., Buote, R., Van Heerden, D., Luan, H., Cullen, K., Slade, L., & Taylor, N. (2020). Reliability and Validity of Commercially Available Wearable Devices for Measuring Steps, Energy Expenditure, and Heart Rate: Systematic Review. *JMIR mHealth and uHealth, 8*(9), e18694. https://doi.org/10.2196/18694

Garcia, J. M., Sirard, J. R., Larsen, R., Bruening, M., Wall, M., & Neumark-Sztainer, D. (2016). Social and psychological factors associated with adolescent physical activity. *Journal of physical activity and health, 13*(9), 957-963. https://doi.org/10.1123/jpah.2015-0224

Garriguet, D., & Colley, R. C. (2014). A comparison of self-reported leisure-time physical activity and measured moderate-to-vigorous physical activity in adolescents and adults. *Health reports, 25*(7), 3–11.

Gillis, L., Tomkinson, G., Olds, T., Moreira, C., Christie, C., Nigg, C., Cerin, E., Van Sluijs, E., Stratton, G., Janssen, I., Dorovolomo, J., Reilly, J. J., Mota, J., Zayed, K., Kawalski, K., Andersen, L. B., Carrizosa, M., Tremblay, M., Chia, M., Hamlin, M., … Van Mechelen, W. (2013). Research priorities for child and adolescent physical activity and sedentary behaviours: an international perspective using a twin-panel Delphi procedure. *The international journal of behavioral nutrition and physical activity, 10,* 112. https://doi.org/10.1186/1479-5868-10-112

González-Cutre, D., Ferriz, R., Beltrán-Carrillo, V. J., Andrés-Fabra, J. A., Montero-Carretero, C., Cervelló, E., & Moreno-Murcia, J. A. (2014).

Promotion of autonomy for participation in physical activity: a study based on the trans-contextual model of motivation. *Educational Psychology, 34*(3), 367-384. https://doi.org/10.1080/01443410.2013.817325

González-Cutre, D., Sierra, A. C., Beltrán-Carrillo, V. J., Peláez-Pérez, M., & Cervelló, E. (2018). A school-based motivational intervention to promote physical activity from a self-determination theory perspective. *Journal of Educational Research, 111*(3), 320–330. https://doi.org/10.1080/00220671.2016.1255871

Gorzelitz, J., Farber, C., Gangnon, R., & Cadmus-Bertram, L. (2020). Accuracy of Wearable Trackers for Measuring Moderate- to Vigorous-Intensity Physical Activity: A Systematic Review and Meta-Analysis, *Journal for the Measurement of Physical Behaviour, 3*(4), 346-357. https://doi.org/10.1123/jmpb.2019-0072

Grao-Cruces, A., Ruiz-López, R., Moral-García, J. E., Ruiz-Ariza, A., & Martínez-López, E. J. (2016). Effects of a steps/day programme with evaluation in physical education on body mass index in schoolchildren 11-12 years of age. *Kinesiology, 48*(1), 132–41. https://doi.org/10.26582/k.48.1.2

Guijarro-Romero, S., Casado-Robles, C., & Mayorga-Vega, D. (2019). Unidades didácticas reforzadas, alternadas, irregulares e intermitentes como herramienta para desarrollar y mantener un nivel saludable de capacidad cardiorrespiratoria a través de la Educación Física. In S. Alonso García, J. Romero Rodríguez, C. Rodríguez-Jiménez, & J. Sola Reche (Eds.), *Investigación, Innovación docente y TIC. Nuevos horizontes educativos* (pp. 232–244). Dykinson

Guijarro, E., Rocamora, I., Evangelio, C., & Víllora, S. G. (2020). Sport Education Model in Spain: a systematic review [El modelo de educación deportiva en España: Una revisión sistemática]. *Retos, 38*, 886-894. https://doi.org/10.47197/retos.v38i38.77249

Guijarro, E., Rocamora, I., González-Víllora, S., & Arias-Palencia, N.M. (2020). The role of physical Education in the achievement of international recommendations: A study based on pedagogical models. *Journal of Human Sport and Exercise, 15*(4), 849-860. https://doi.org/10.14198/jhse.2020.154.12

Guthold, R., Stevens, G. A., Riley, L. M., & Bull, F. C. (2020). Global trends in insufficient physical activity among adolescents: a pooled analysis of 298 population-based surveys with 1.6 million participants. *The Lancet Child & Adolescent Health, 4*(1), 23-35. https://doi.org/10.1016/S2352-4642(19)30323-2

Hagger, M. S., & Chatzisarantis, N. L. (2016). The Trans-Contextual Model of Autonomous Motivation in Education: Conceptual and Empirical Issues and Meta-Analysis. *Review of educational research, 86*(2), 360–407. https://doi.org/10.3102/0034654315585005

Hagger, M., Chatzisarantis, N. L., Hein, V., Soós, I., Karsai, I., Lintunen, T., & Leemans, S. (2009). Teacher, peer and parent autonomy support in physical education and leisure-time physical activity: A trans-contextual model of motivation in four nations. *Psychology & health, 24*(6), 689–711. https://doi.org/10.1080/08870440801956192

Hardman, C. A., Horne, P. J., & Fergus Lowe, C. (2011). Effects of rewards, peer-modelling and pedometer targets on children's physical activity: a school-based intervention study. *Psychology & health, 26*(1), 3–21. https://doi.org/10.1080/08870440903318119

Hardman, K., Murphy, C., Routen, A., & Tones, S. (2014). *UNESCO-NWCPEA: World-wide survey of school Physical Education.* United Nations Educational, Scientific and Cultural Organization.

Henriksen, A., Haugen Mikalsen, M., Woldaregay, A. Z., Muzny, M., Hartvigsen, G., Hopstock, L. A., & Grimsgaard, S. (2018). Using Fitness Trackers and Smartwatches to Measure Physical Activity in Research: Analysis of Consumer Wrist-Worn Wearables. *Journal of medical Internet research, 20*(3), e110. https://doi.org/10.2196/jmir.9157

Hills, A. P., Dengel, D. R., & Lubans, D. R. (2015). Supporting public health priorities: recommendations for physical education and physical activity promotion in schools. *Progress in cardiovascular diseases, 57*(4), 368–374. https://doi.org/10.1016/j.pcad.2014.09.010

Hollis, J. L., Sutherland, R., Williams, A. J., Campbell, E., Nathan, N., Wolfenden, L., Morgan, P. J., Lubans, D. R., Gillham, K., & Wiggers, J. (2017). A systematic review and meta-analysis of moderate-to-vigorous physical activity levels in secondary school physical education lessons. *The international journal of behavioral nutrition and physical activity, 14*(1), 52. https://doi.org/10.1186/s12966-017-0504-0

Howie, E. K., & Straker, L. M. (2016). Rates of attrition, non-compliance and missingness in randomized controlled trials of child physical activity interventions using accelerometers: A brief methodological review. *Journal of science and medicine in sport*, *19*(10), 830–836. https://doi.org/10.1016/j.jsams.2015.12.520

Institute of Medicine. (2013). *Educating the student body: Taking physical activity and physical education to school*. National Academies Press.

International Data Corporation. (2020). *Shipments of Wearable Devices Leap to 125 Million Units, Up 35.1% in the Third Quarter, According to IDC*. https://www.idc.com/getdoc.jsp?containerId=prUS47067820.

International Sport and Culture Association. (2015). *The economic cost of Physical inactivity in Europe. An ISCA/ Cebr report.* ISCA/ Cebr Office.

Jago, R., Baranowski, T., Baranowski, J. C., Thompson, D., Cullen, K. W., Watson, K., & Liu, Y. (2006). Fit for Life Boy Scout badge: outcome evaluation of a troop and Internet intervention. *Preventive medicine*, *42*(3), 181–187. https://doi.org/10.1016/j.ypmed.2005.12.010

Jago, R., Rawlins, E., Kipping, R. R., Wells, S., Chittleborough, C., Peters, T. J., Mytton, J., Lawlor, D. A., & Campbell, R. (2015). Lessons learned from the AFLY5 RCT process evaluation: implications for the design of physical activity and nutrition interventions in schools. *BMC public health*, *15*, 946. https://doi.org/10.1186/s12889-015-2293-1

Kriemler, S., Meyer, U., Martin, E., van Sluijs, E. M., Andersen, L. B., & Martin, B. W. (2011). Effect of school-based interventions on physical activity and fitness in children and adolescents: a review of reviews and systematic update. *British journal of sports medicine*, *45*(11), 923–930. https://doi.org/10.1136/bjsports-2011-090186

Lonsdale, C., Rosenkranz, R. R., Peralta, L. R., Bennie, A., Fahey, P., & Lubans, D. R. (2013). A systematic review and meta-analysis of interventions designed to increase moderate-to-vigorous physical activity in school physical education lessons. *Preventive medicine*, *56*(2), 152–161. https://doi.org/10.1016/j.ypmed.2012.12.004

Lubans, D. R., Morgan, P. J., Aguiar, E. J., & Callister, R. (2011). Randomized controlled trial of the Physical Activity Leaders (PALs) program for adolescent boys from disadvantaged secondary schools. *Preventive medicine*, *52*(3-4), 239–246. https://doi.org/10.1016/j.ypmed.2011.01.009

Lubans, D., Richards, J., Hillman, C., Faulkner, G., Beauchamp, M., Nilsson, M., Kelly, P., Smith, J., Raine, L., & Biddle, S. (2016). Physical Activity for Cognitive and Mental Health in Youth: A Systematic Review of Mechanisms. *Pediatrics*, *138*(3), e20161642. https://doi.org/10.1542/peds.2016-1642

Martínez de Ojeda, D., Méndez-Giménez, A., & Valverde, J. J. (2016). Sport Education Model effects in the social climate classroom, perceived competence and intent to be physically active: an extended study in Primary Education [Efectos del modelo Educación Deportiva en el clima social del aula, la competencia percibida y la intención de ser físicamente activo: un estudio prolongado en Primaria]. *SPORT TK-Revista EuroAmericana de Ciencias del Deporte, 5*(2), 153-166.

Mavropoulou, A., Barkoukis, V., Douka, S., Alexandris, K., & Hatzimanouil, D. (2018). The role of autonomy supportive activities on students' motivation and beliefs toward out-of-school activities. *The Journal of Educational Research, 112*(2), 223-233. https://doi.org/10.1080/00220671.2018.1503580

Mayorga-Vega, D., Martínez-Baena, A., & Viciana, J. (2018). Does school physical education really contribute to accelerometer-measured daily physical activity and non sedentary behaviour in high school students? *Journal of sports sciences*, *36*(17), 1913–1922. https://doi.org/10.1080/02640414.2018.1425967

McDavid, L., Cox, A. E., & Amorose, A. J. (2012). The relative roles of physical education teachers and parents in adolescents' leisure-time physical activity motivation and behavior. *Psychology of Sport and Exercise*, *13*(2), 99-107. https://doi.org/10.1016/j.psychsport.2011.10.003

McKenzie, T. L., & Lounsbery, M. A. (2014). The pill not taken: revisiting Physical Education Teacher Effectiveness in a Public Health Context. *Research quarterly for exercise and sport*, *85*(3), 287–292. https://doi.org/10.1080/02701367.2014.931203

Mercer, K., Giangregorio, L., Schneider, E., Chilana, P., Li, M., & Grindrod, K. (2016). Acceptance of Commercially Available Wearable Activity Trackers Among Adults Aged Over 50 and With Chronic Illness: A Mixed-Methods

Evaluation. *JMIR mHealth and uHealth, 4*(1), e7. https://doi.org/10.2196/mhealth.4225

Metcalf, B., Henley, W., & Wilkin, T. (2012). Effectiveness of intervention on physical activity of children: systematic review and meta-analysis of controlled trials with objectively measured outcomes (EarlyBird 54). *BMJ, 345,* e5888. https://doi.org/10.1136/bmj.e5888

Metzler, M. (2017). *Instructional models for Physical Education (3 rd ed.).* Routledge.

Michie, S., Abraham, C., Whittington, C., McAteer, J., & Gupta, S. (2009). Effective techniques in healthy eating and physical activity interventions: a meta-regression. *Health psychology : official journal of the Division of Health Psychology, American Psychological Association, 28*(6), 690–701. https://doi.org/10.1037/a0016136

Michie, S., Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., Eccles, M. P., Cane, J., & Wood, C. E. (2013). The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Annals of behavioral medicine, 46*(1), 81–95. https://doi.org/10.1007/s12160-013-9486-6

Migueles, J. H., Cadenas-Sanchez, C., Ekelund, U., Delisle Nyström, C., Mora-Gonzalez, J., Löf, M., Labayen, I., Ruiz, J. R., & Ortega, F. B. (2017). Accelerometer Data Collection and Processing Criteria to Assess Physical Activity and Other Outcomes: A Systematic Review and Practical Considerations. *Sports medicine, 47*(9), 1821–1845. https://doi.org/10.1007/s40279-017-0716-0

Murad, M. H., Asi, N., Alsawas, M., & Alahdab, F. (2016). New evidence pyramid. *Evidence-based medicine, 21*(4), 125–127. https://doi.org/10.1136/ebmed-2016-110401

Murillo-Pardo, B., García-Bengoechea, E., Generelo-Lanaspa, E., Bush, P. L., Zaragoza-Casterad, J., Julián-Clemente, J. A., & García-González, L. (2013). Promising school-based strategies and intervention guidelines to increase physical activity of adolescents. *Health education research, 28*(3), 523–538. https://doi.org/10.1093/her/cyt040

Neil-Sztramko, S. E., Caldwell, H., & Dobbins, M. (2021). School-based physical activity programs for promoting physical activity and fitness in children and adolescents aged 6 to 18. *The Cochrane database of systematic reviews*, *9*(9), CD007651. https://doi.org/10.1002/14651858.CD007651.pub3

Parrish, A. M., Okely, A. D., Stanley, R. M., & Ridgers, N. D. (2013). The effect of school recess interventions on physical activity : a systematic review. *Sports medicine*, *43*(4), 287–299. https://doi.org/10.1007/s40279-013-0024-2

Patel, M. S., Asch, D. A., & Volpp, K. G. (2015). Wearable devices as facilitators, not drivers, of health behavior change. *JAMA*, *313*(5), 459–460. https://doi.org/10.1001/jama.2014.14781

Perlman, D. (2012). The influence of the Sport Education Model on amotivated students' in-class physical activity. *European Physical Education Review, 18*, 335-345. https://doi.org/10.1177/1356336X12450795

Poitras, V. J., Gray, C. E., Borghese, M. M., Carson, V., Chaput, J. P., Janssen, I., Katzmarzyk, P. T., Pate, R. R., Connor Gorber, S., Kho, M. E., Sampson, M., & Tremblay, M. S. (2016). Systematic review of the relationships between objectively measured physical activity and health indicators in school-aged children and youth. *Applied physiology, nutrition, and metabolism*, *41*(6), S197–S239. https://doi.org/10.1139/apnm-2015-0663

Prince, S. A., Cardilli, L., Reed, J. L., Saunders, T. J., Kite, C., Douillette, K., Fournier, K., & Buckley, J. P. (2020). A comparison of self-reported and device measured sedentary behaviour in adults: a systematic review and meta-analysis. *The international journal of behavioral nutrition and physical activity*, *17*(1), 31. https://doi.org/10.1186/s12966-020-00938-3

Puente-Maxera, F., Méndez-Giménez, A., & Martínez de Ojeda, D. (2020). Physical activity levels during a Sport Education season of games from around the world. *European Physical Education Review, 27*(2), 240-257. https://doi.org/10.1177/1356336X20939591

Pulido, J. J., Sánchez-Oliva, D., Amado, D., González-Ponce, I., & Sánchez-Miguel, P. A. (2014). Influence of motivational processes on enjoyment, boredom and intention to persist in young sportspersons. *South African journal for research in sport, physical education and recreation, 36*(3), 135-149.

Regional Ministry of Education and Youth of the Government of Madrid. (2020). Decree 59/2020, of July 29, of the Governing Council, amending Decree 48/2015, of May 14, of the Governing Council, which establishes for the

Community of Madrid the curriculum for Compulsory Secondary Education. *Official Gazette of the Madrid Regional Government, 184, 11-19.*

Reilly, J. J., Penpraze, V., Hislop, J., Davies, G., Grant, S., & Paton, J. Y. (2008). Objective measurement of physical activity and sedentary behaviour: review with new data. *Archives of disease in childhood, 93*(7), 614–619. https://doi.org/10.1136/adc.2007.133272

Rhodes, R. E., Janssen, I., Bredin, S., Warburton, D., & Bauman, A. (2017). Physical activity: Health impact, prevalence, correlates and interventions. *Psychology & health, 32*(8), 942–975. https://doi.org/10.1080/08870446.2017.1325486

Rich, E., & Miah, A. (2016). Mobile, wearable and ingestible health technologies: towards a critical research agenda. Health sociology review, 26(1), 84-97. https://doi.org/10.1080/14461242.2016.1211486

Rocamora, I., González-Víllora, S., Fernández-Río, J., & Arias-Palencia, N.M. (2019). Physical activity levels, game performance and friendship goals using two different pedagogical models: Sport Education and Direct Instruction. *Physical Education and Sport Pedagogy, 24*(1), 87-102. https://doi.org/10.1080/17408989.2018.1561839

Romanzini, M., Petroski, E. L., Ohara, D., Dourado, A. C., & Reichert, F. F. (2014). Calibration of ActiGraph GT3X, Actical and RT3 accelerometers in adolescents. *European journal of sport science, 14*(1), 91–99. https://doi.org/10.1080/17461391.2012.732614

Ruiz, D. C., & Goransson, A. (2015). *Professional Android Wearables*. John Wiley & Sons.

Ruiz, J. R., Ortega, F. B., Martínez-Gómez, D., Labayen, I., Moreno, L. A., De Bourdeaudhuij, I., Manios, Y., Gonzalez-Gross, M., Mauro, B., Molnar, D., Widhalm, K., Marcos, A., Beghin, L., Castillo, M. J., Sjöström, M., & HELENA Study Group (2011). Objectively measured physical activity and sedentary time in European adolescents: the HELENA study. *American journal of epidemiology, 174*(2), 173–184. https://doi.org/10.1093/aje/kwr068

Russ, L. B., Webster, C. A., Beets, M. W., & Phillips, D. S. (2015). Systematic Review and Meta-Analysis of Multi-Component Interventions Through Schools to Increase Physical Activity. *Journal of physical activity & health, 12*(10), 1436–1446. https://doi.org/10.1123/jpah.2014-0244

Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology, 61,* 101860. https://doi.org/10.1016/j.cedpsych.2020.101860

Sánchez-Oliva, D., Sánchez-Miguel, P. A., Leo, F. M., Kinnafick, F. E., & García-Calvo, T. (2014). Physical education lessons and physical activity intentions within Spanish secondary schools: A self-determination perspective. *Journal of Teaching in Physical Education*, *33*(2), 232-249. https://doi.org/10.1123/jtpe.2013-0043

Santos-Lozano, A., Santín-Medeiros, F., Cardon, G., Torres-Luque, G., Bailón, R., Bergmeir, C., Ruiz, J. R., Lucia, A., & Garatachea, N. (2013). Actigraph GT3X: validation and determination of physical activity intensity cut points. *International journal of sports medicine*, *34*(11), 975–982. https://doi.org/10.1055/s-0033-1337945

Scheuer, C., Antala, B., & Holzweg, M. (2014). *Physical Education: Quality in Management and Teaching*. Logos Verlag Berlin

Schmidt, F.L., & Hunter, J.E. (2021). *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings. 3rd Edition*. SAGE Publications.

Sevil, J., Abós, Á., Aibar, A., Julián, J. A., & García-González, L. (2016). Gender and corporal expression activity in physical education: Effect of an intervention on students' motivational processes. *European Physical Education Review*, *22*(3), 372-389. https://doi.org/10.1177/1356336X15613463

Sevil, J., Aibar, A., Abós, Á., Generelo, E., & García-González, L. (2020). Improving motivation for physical activity and physical education through a schoolbased intervention. *The Journal of Experimental Education*, 1–21. https://doi.org/10.1080/00220973.2020.1764466

Sheeran, P., Klein, W. M., & Rothman, A. J. (2017). Health Behavior Change: Moving from Observation to Intervention. *Annual review of psychology*, *68*, 573–600. https://doi.org/10.1146/annurev-psych-010416-044007

Shephard R. J. (2003). Limits to the measurement of habitual physical activity by questionnaires. *British journal of sports medicine*, *37*(3), 197–206. https://doi.org/10.1136/bjsm.37.3.197

Shephard, R., & Tudor-Locke, C. (Eds.). (2016). *The objective monitoring of physical activity: Contributions of accelerometry to epidemiology, exercise science and rehabilitation*. Springer.

Siedentop, D. (1994). *Sport education: Quality physical education through positive sport experiences*. Human Kinetics.

Siedentop, D., Hastie, P., & Van der Mars, H. (2019). *Complete guide to sport education: Edition 3*. Human Kinetics.

Šimůnek, A., Dygrýn, J., Jakubec, L., Neuls, F., Frömel, K., & Welk, G. J. (2019). Validity of Garmin Vívofit 1 and Garmin Vívofit 3 for School-Based Physical Activity Monitoring. *Pediatric exercise science, 31*(1), 130–136. https://doi.org/10.1123/pes.2018-0019

Sirard, J. R., & Pate, R. R. (2001). Physical activity assessment in children and adolescents. *Sports medicine, 31*(6), 439–454. https://doi.org/10.2165/00007256-200131060-00004

Skender, S., Ose, J., Chang-Claude, J., Paskow, M., Brühmann, B., Siegel, E. M., Steindorf, K., & Ulrich, C. M. (2016). Accelerometry and physical activity questionnaires - a systematic review. *BMC public health, 16*, 515. https://doi.org/10.1186/s12889-016-3172-0

Slootmaker, S. M., Schuit, A. J., Chinapaw, M. J., Seidell, J. C., & van Mechelen, W. (2009). Disagreement in physical activity assessed by accelerometer and self-report in subgroups of age, gender, education and weight status. *The international journal of behavioral nutrition and physical activity, 6*, 17. https://doi.org/10.1186/1479-5868-6-17

Society of Health and Physical Educators. (2014). *National standards & grade-level outcomes for K-12 physical education.* Human Kinetics.

Spanish Ministry of Education and Vocational Training. (2020). Organic Law 3/2020, of December 29, amending Organic Law 2/2006, of May 3, 2006, on Education. [Ley Orgánica 3/2020, de 29 de diciembre, por la que se modifica la Ley Orgánica 2/2006, de 3 de mayo, de Educación]. *Goverment Gazette, 340*, 122868-122953.

Spanish Ministry of Education, Culture and Sport. (2015). Royal Decree 1105/2014, laying down the curriculum for Secondary Education and Baccalaureate [Real Decreto 1105/2014 de 26 de diciembre, por el que se establece el currículo básico de la Educación Secundaria Obligatoria y del Bachillerato].

Goverment Gazette, *3*, 169–546.

Sriram, K., Mulder, H. S., Frank, H. R., Santanam, T. S., Skinner, A. C., Perrin, E. M., Armstrong, S. C., Peterson, E. D., Pencina, M. J., & Wong, C. A. (2021). The Dose-Response Relationship Between Physical Activity and Cardiometabolic Health in Adolescents. *American journal of preventive medicine*, *60*(1), 95–103. https://doi.org/10.1016/j.amepre.2020.06.027

Strath, S. J., & Rowley, T. W. (2018). Wearables for Promoting Physical Activity. *Clinical chemistry*, *64*(1), 53–63. https://doi.org/10.1373/clinchem.2017.272369

Teixeira, P. J., Carraça, E. V., Markland, D., Silva, M. N., & Ryan, R. M. (2012). Exercise, physical activity, and self-determination theory: a systematic review. *The international journal of behavioral nutrition and physical activity*, *9*, 78. https://doi.org/10.1186/1479-5868-9-78

Tremblay, M. S., Aubert, S., Barnes, J. D., Saunders, T. J., Carson, V., Latimer-Cheung, A. E., Chastin, S., Altenburg, T. M., Chinapaw, M., & SBRN Terminology Consensus Project Participants. (2017). Sedentary Behavior Research Network (SBRN) - Terminology Consensus Project process and outcome. *The international journal of behavioral nutrition and physical activity*, *14*(1), 75. https://doi.org/10.1186/s12966-017-0525-8

Tremblay, M. S., Carson, V., Chaput, J., Connor Gorber, S., Dinh, T., Duggan, M., Faulkner, G., Gray, C. E., Gruber, R., Janson, K., Janssen, I., Katzmarzyk, P. T., Kho, M. E., Latimer-Cheung, A. E., LeBlanc, C., Okely, A. D., Olds, T., Pate, R. R., Phillips, A., Poitras, V. J., Rodenburg, S., Sampson, M., Saunders, T. J., Stone, J. A., Stratton, G., Weiss, S. K. & Zehr, L. (2016). Canadian 24-hour movement guidelines for children and youth: An integration of physical activity, sedentary behaviour, and sleep. *Applied Physiology, Nutrition, and Metabolism* *41*(6), S311-327. https://doi.org/10.1139/apnm-2016- 0151

Troiano, R. P., McClain, J. J., Brychta, R. J., & Chen, K. Y. (2014). Evolution of accelerometer methods for physical activity research. *British journal of sports medicine*, *48*(13), 1019–1023. https://doi.org/10.1136/bjsports-2014-093546

Trost, S. G., Loprinzi, P. D., Moore, R., & Pfeiffer, K. A. (2011). Comparison of accelerometer cut points for predicting activity intensity in youth. *Medicine and* *science in sports and exercise*, *43*(7), 1360–1368. https://doi.org/10.1249/MSS.0b013e318206476e

Tudor-Locke, C., Craig, C. L., Beets, M. W., Belton, S., Cardon, G. M., Duncan, S., Hatano, Y., Lubans, D. R., Olds, T. S., Raustorp, A., Rowe, D. A., Spence, J. C., Tanaka, S., & Blair, S. N. (2011). How many steps/day are enough? for children and adolescents. *The international journal of behavioral nutrition and physical activity*, *8*, 78. https://doi.org/10.1186/1479-5868-8-78

Uddin, R., Salmon, J., Islam, S., & Khan, A. (2020). Physical education class participation is associated with physical activity among adolescents in 65 countries. *Scientific reports*, *10*(1), 22128. https://doi.org/10.1038/s41598-020-79100-9

United Nations Educational, Scientific and Cultural Organization (UNESCO). (2015). *Quality Physical Education (QPE): Guidelines for policy makers*. UNESCO.

United Nations Educational, Scientific and Cultural Organization (UNESCO). (1962). *New methods and techniques in education. Report of a meeting of experts*. https://unesdoc.unesco.org/ark:/48223/pf0000126329

Van Hecke, L., Loyen, A., Verloigne, M., van der Ploeg, H. P., Lakerveld, J., Brug, J., De Bourdeaudhuij, I., Ekelund, U., Donnelly, A., Hendriksen, I., Deforche, B., & DEDIPAC consortium (2016). Variation in population levels of physical activity in European children and adolescents according to cross-European studies: a systematic literature review within DEDIPAC. *The international journal of behavioral nutrition and physical activity*, *13*, 70. https://doi.org/10.1186/s12966-016-0396-4

Viciana, J. (2002). *Planificar en Educación Física*. INDE.

Viciana, J., & Mayorga-Vega, D. (2016). Innovative teaching units applied to Physical Education – changing the curriculum management for authentic outcomes. *Kinesiology*, *48*(1), 142–152. https://doi.org/doi:10.26582/k.48.1.1.

Viciana, J., & Mayorga-Vega, D. (2018). The three-axes model of planning in physical education. *Retos*, *33*, 313–319. https://doi.org/10.47197/retos.v0i33.54533

Viciana, J., Martínez-Baena, A., & Mayorga-Vega, D. (2015). Contribution of physical education to daily recommendations of physical activity in adolescents according to gender; a study with accelerometry [Contribución de la educación física a las recomendaciones diarias de actividad física en adolescentes según el género; un estudio con acelerometría]. *Nutricion Hospitalaria*, *32*(3), 1246–1251. https://doi.org/10.3305/nh.2015.32.3.9363.

Viciana, J., Mayorga-Vega, D., & Merino-Marban, R. (2014). Physical education-based planning for developing and maintaining students' health-related physical fitness levels. In R. Todaro (Ed.), *Handbook of physical education research: Role of school programs, children's attitudes and health implications* (pp. 237–252). Nova Science Publishers.

Wallhead, T. L., Garn, A. C., & Vidoni, C. (2014). Effect of a sport education program on motivation for physical education and leisure-time physical activity. *Research quarterly for exercise and sport*, *85*(4), 478–487. https://doi.org/10.1080/02701367.2014.961051

Wallhead, T. L., Hagger, M., & Smith, D. T. (2010). Sport education and extracurricular sport participation: an examination using the trans-contextual model of motivation. *Research quarterly for exercise and sport*, *81*(4), 442–455. https://doi.org/10.1080/02701367.2010.10599705

Wallhead, T., Gran, A. C., & Vidoni, C. (2013). Sport Education and social goals in physical Education: relationships with enjoyment, relatedness, and leisure-time physical activity. *Physical Education and Sport Pedagogy, 18*(4), 427-441. https://doi.org/10.1080/17408989.2012.690377

Wang, L. (2017). Using the self-determination theory to understand Chinese adolescent leisure-time physical activity. *European journal of sport science*, *17*(4), 453–461. https://doi.org/10.1080/17461391.2016.1276968

Wang, Y., & Chen, A. (2019). Two Pathways Underlying the Effects of Physical Education on Out-of-School Physical Activity. *Research quarterly for exercise and sport*, *91*(2), 197–208. https://doi.org/10.1080/02701367.2019.1656325

Welk, G. J. (1999). The youth physical activity promotion model: A conceptual bridge between theory and practice. *Quest, 51*(1), 5-23. https://doi.org/10.1080/00336297.1999.10484297

World Health Organization. (2008). *WHO global strategy on diet, physical activity and health: a framework to monitor and evaluate implementation*. World Health Organization.

World Health Organization. (2018). *Physical Activity Factsheets for the 28 European Union Member States of the Who European Region*. World Health Organization

World Health Organization. (2018b). *Promoting physical activity in the education sector*. World Health Organization

World Health Organization. (2020). *WHO guidelines on physical activity and sedentary behaviour*. World Health Organization.

Yli-Piipari, S., Layne, T., Hinson, J., & Irwin, C. (2018). Motivational pathways to leisure-time physical activity participation in urban physical education: A cluster-randomized trial. *Journal of Teaching in Physical Education*, *37*(2), 123-132. https://doi.org/10.1123/jtpe.2017-0099

ACKNOWLEDGMENTS

[AGRADECIMIENTOS]

ACKNOWLEDGMENTS [AGRADECIMIENTOS]

Después de estos cuatro años parece mentira que llegue el momento de escribir estas líneas, ¡pellizcarme que no me lo creo! Los que me conocéis sabréis que este es el apartado que más me ha costado escribir de toda la Tesis Doctoral. Sin embargo, ha sido sin duda el más gratificante. Escribiendo estas líneas me he dado cuenta de la suerte que he tenido de recorrer este camino acompañada por muchas personas, sin las cuáles estoy segura que no podría haber llegado hasta aquí. He tenido unos directores y compañeros de trabajo geniales, que siempre han estado dispuestos a ayudar. Unos amigos que hacen que me olvide de todo cuando estoy con ellos y que siempre convierten pequeños momentos en grandes historias. Una familia que se preocupa por mí, y que siempre está a mi lado incondicionalmente. Y por supuesto, un compañero de vida que siempre me ha estado apoyando y que ha sufrido el esfuerzo dedicado a esta Tesis tanto como yo. Por ello, no puedo acabar la Tesis sin agradecerles a todos ellos que hayan estado a mi lado.

En primer lugar, mis palabras de agradecimiento van para mis directores, **Jesús** y **Dani**, lo que escriba en estas líneas se queda corto para agradeceros a ambos todo el tiempo y esfuerzo que habéis puesto en ayudarme a llegar hasta aquí.

Jesús, tu me abriste la puerta al grupo hace 5 años y me iniciaste en este mundo. Recuerdo que yo estaba convencida en prepararme las oposiciones a Secundaria y... ¡me trastocaste los planes! Gracias por hacerlo, por confiar en mí, por verme capaz de conseguirlo y por recordármelo cada día de este proceso. Por convertirte en un "padre académico" para mí, por compartir todo tu conocimiento conmigo y por abrirme las puertas de tu casa y estar disponible siempre que lo he necesitado. Gracias también por seguir trabajando desinteresadamente para los que venimos detrás, y por preocuparte en conseguir proyectos que nos sigan abriendo puertas aunque te desborde el trabajo. Y por supuesto, gracias por tu guía en la docencia, por hacer sencillos mis primeros pasos en la asignatura de Fundamentos de la Educación Física y por ayudarme a ser mejor profesora.

Dani, mi director en la sombra, si Jesús ha sido el padre académico tu has sido el "hermano mayor". Gracias por tu orientación continua en mi formación, por compartir conmigo tus inalcanzables conocimientos de metodología y estadística, por tu paciencia para explicarme las cosas las veces que fuesen necesarias, y también (aunque a veces me explotase la cabeza) por tus minuciosas correcciones de los artículos. Gracias por enseñarme a ser rigurosa en todo el proceso de investigación y por contagiarme esas ganas por investigar. Pero sobre todo, gracias por preocuparte por mi a nivel personal. Porque cada llamada de teléfono (aunque la mitad de las veces no te lo coja), siempre empiezan por preguntar como estoy y terminan con un "¡Venga que tú puedes, ya queda nada!". Mil gracias por tus palabras de ánimo constantes durante esta aventura.

En definitiva, gracias a ambos por darme la oportunidad de formar parte de este equipo, por vuestra confianza, por vuestro apoyo a lo largo de estos años, y por estar siempre disponibles al teléfono fuese la hora que fuese. Sin vuestra ayuda nada de esto hubiera sido posible. Si como profesionales sois de Q1, como personas sois aún más grandes. He aprendido tantas cosas de vosotros, que de hecho creo que en cada conversación aprendo algo nuevo. Y, por supuesto, no puedo olvidarme de darles las gracias también a vuestras familias, especialmente a **Lucía**, **Marco** y **Paula** que también han sufrido que les robe a sus papis muchos días. Espero haber estado a la altura de vuestras expectativas y que hayáis disfrutado tanto como yo.

Gracias también a mi compañero de grupo y amigo, **Santi**. Por tu apoyo, por ser optimista con cada rechazo que nos llevábamos (y llevamos) en los artículos, por preparar numerosos congresos y cursos juntos, por motivarme a seguir hacia delante y por estar dispuesto a compartir todo lo que sabes. Eres un ejemplo de que trabajando duro al final llega la recompensa. Eres un compañero fantástico y espero que sigamos trabajando juntos muchos años. Gracias también a los compañeros que en algún momento han formado parte de esta pequeña familia, **Emilio** y **Alejandro**. Gracias por ayudarme en las tediosas tomas de datos con acelerómetros y pulseras y hacerlas más llevaderas. De igual modo, agradecer a los alumnos del Máster de Investigación que han colaborado en algunas tomas de datos de esta Tesis Doctoral, en especial a **Ángel** y **Alejandro**.

Un agradecimiento muy especial a todos los alumnos por su participación voluntaria y desinteresada en nuestros estudios. Gracias por vuestra paciencia rellenado tantos cuestionarios y llevando los aparatitos cada día. Agradezco también a todos los miembros de los centros escolares que nos han permitido llevar a cabo nuestros proyectos. En especial, gracias a los profesores de Educación Física **Maribel**, **Inés**, **José Antonio** y **Jorge** por vuestra colaboración y predisposición para ayudarme siempre, y por enseñarme lo más bonito de esta profesión. Gracias a todos por compartir esta experiencia conmigo, sin vosotros nada de esto sería posible.

*Thanks to all the people that helped me during my research stays abroad in Olomouc. Especially thanks to my tutors, **František Chmelík** and **Jan Dygrýn**, and also thanks to the other members of the team, **Karel Frömel**, **Lukáš**, **Aleš**, **Jana**, and **Mike**, for hosting me as a friend. Thank you for every coffee, lunches and talks we shared together. On the other hand, I can't forget to thank **Vanda**, and her beautiful children **Oskar** and **Ester**. Thank you for opening the doors of your home to us, and for all the laughs and trips you shared with us. And of course, thanks to my friends **Vicky** and **Dana**. For the great moments trying to teach us Czech, for immersing us in your culture, for every night drinking wine and singing the Macarena and Despacito with your Slovak-Spanish accents, and also for laughing at me in yoga class. All of you made my research stays unforgettable, and made me fall in love with Olomouc. I hope we can continue sharing good moments in the future, ¡Děkuji!*

También quería agradecer a la **Facultad de Ciencias del Deporte de Granada** y a todos los profesores y compañeros que forman parte de ella. Ha sido mi casa durante 9 años, en la que he crecido muchísimo desde que entré en ella, y de la que me llevo grandes personas. En especial, agradecer a mis amigas y futuras Doctoras, **Patri**, **Nuria**, **Lucía** y **Cristina**. Fuimos las locas que se metieron en el doble Máster y qué orgullo ver donde hemos llegado juntas. **Patri**, gracias por acompañarme siempre en este camino desde que entramos en la Facultad, en lo académico pero sobre todo en lo personal, y por ser siempre la llamada de socorro para desahogarnos. Espero seguir compartiendo muchos más momentos a tu lado. **Nuria**, que aunque desertaras a Cádiz te seguimos queriendo igual, gracias por poner siempre el punto de locura al grupo, por todos los cotilleos que compartimos y por conseguir que nos veamos todas cada vez que vienes a Granada. **Lucía**, gracias por tu ayuda incondicional, por contagiarme la pasión que le pones a todo lo que haces y por compartir la locura de que graduarnos en Educación Primaria a la vez que hacíamos la Tesis era buena idea. **Cristina**, a la chica tímida que tanto me ha ayudado desde que comencé la carrera, gracias por poner siempre el grado de cordura y tranquilidad a este grupo. Y tampoco puedo olvidarme de mi compañero **Abel**, que desde que lo conocí supe que iba a llegar tan lejos como se propusiera en la investigación. Gracias a todos por siempre intentar sacar un hueco entre tantas evaluaciones, reuniones, estancias y clases para compartir al menos un café y muchas risas, aunque también hayamos compartido nuestros agobios y penas a partes iguales. Creo que ninguna promoción puede presumir de tener tantos futuros Doctores/as y encima haberlo hecho todos con financiación. Es un orgullo haber compartido esta aventura con vosotros y estoy segura que vais a conseguir todo lo que os propongáis. La primera Tesis ya ha caído... ¡Las próximas son las vuestras!

No podría olvidarme de mi gran compañero, **Borja**. Desde el primer minuto que pisamos la Facultad hemos compartido muchísimos trabajos, horas de estudio, y estrés pre-exámenes juntos, pero también muchos buenos ratos de coche, charlas, fútbol y celebraciones varias. Sé que si hoy he llegado hasta aquí ha sido por tu ayuda y tus palabras de ánimo todos estos años, ya que partido a partido hemos llegado a alcanzar las metas propuestas. Muchas gracias por todo. Estoy muy orgullosa de todo lo que has conseguido, y espérame en el patio, que en unos años seremos compañeros.

Gracias también a los compañeros con los que he podido compartir la docencia en la asignatura de Fútbol durante estos años, **Javi**, **Juanma**, **Javi Sevil** y **Pablo**, y al jefe **Alfonso**. Gracias por compartir todo vuestro conocimiento conmigo y por los ratos que hemos pasado organizando clases y corrigiendo. Agradecer también el apoyo del equipo ActiveBrains, en especial a **Fran**, **Jairo** y **Cristina**, por cedernos el material que ha hecho posible llevar a cabo estas tomas de datos. Por último, gracias a todos los compañeros que han ido pasando por la **Sala de Becarios** y que han sido una excusa perfecta para apartar la mirada del ordenador y echar unas risas.

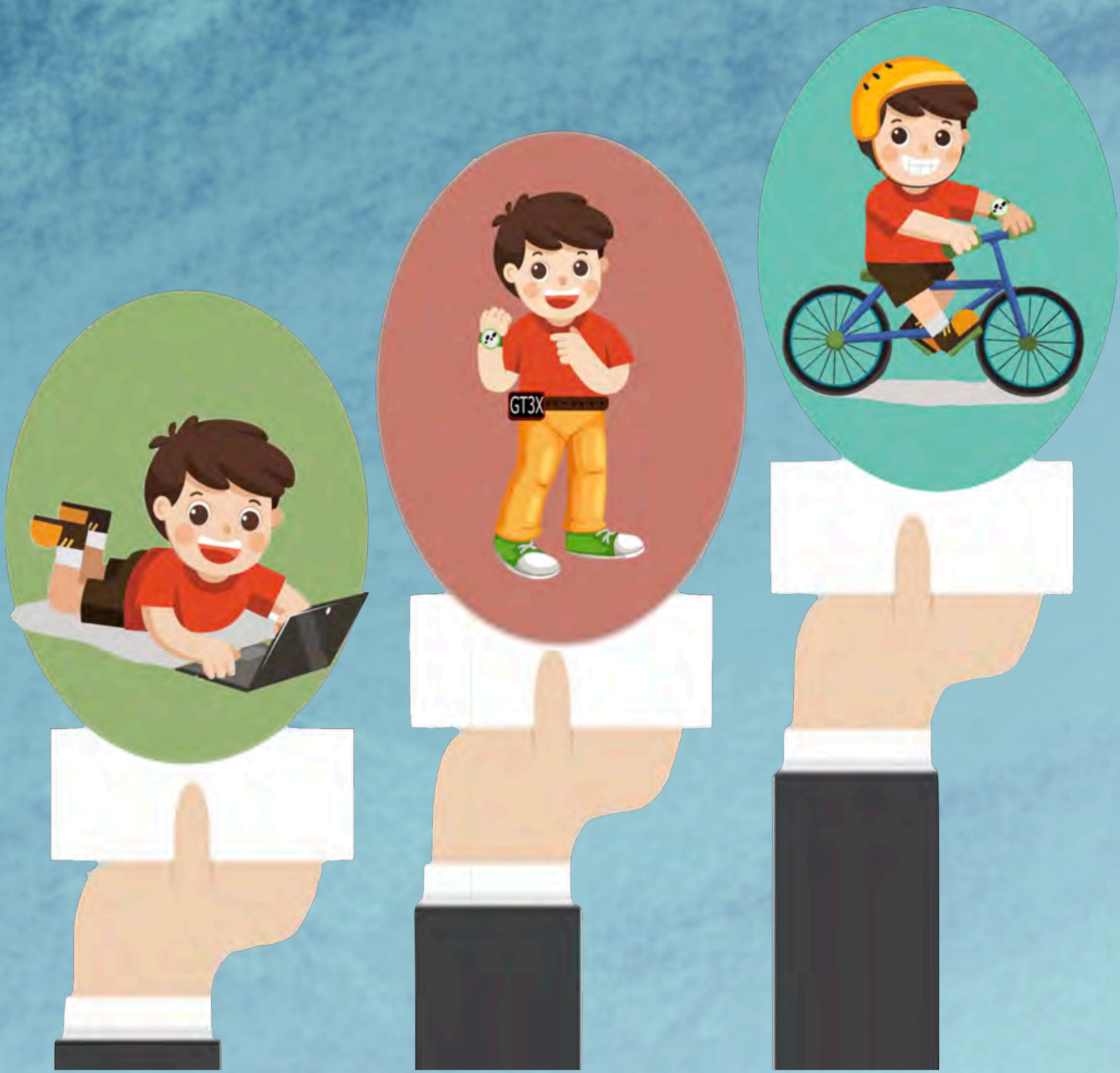
No puedo dejar de dar las gracias a la familia que se elige, a mis amigos del "poblao", por la inmensa suerte que he tenido con vosotros. A los @amiguitos **Marta**, **María**, **Patri**, **Clarilla**, **Julia**, **María Ruíz** y **Fidel**, y a mi trío de cabesas locas, **Cristina**, **Antonio** y **Jesús**. Cada uno llegasteis en un momento distinto a mi vida, pero juntos hemos compartido una gran cantidad de aventuras como una pequeña familia. Gracias por cada rato de desconexión en los que me habéis separado del ordenador estos años. Por todas las barbacoas, cortijos, tardes de juegos, días de playa, luchas entre barcos, ataques de risa sin sentido, las videollamadas confinados y un largo etcétera de momentos "*hygge*" que hemos compartido juntos. Gracias por estar siempre a mi lado, haciendo este camino más llevadero y por animarme cada día a llegar hasta la meta, aunque yo os dijera que no me daba la vida. Por todo el apoyo en los buenos y malos momentos durante todo este tiempo, espero que sigamos compartiendo muchos más y celebrando los logros de los demás del grupo como nuestros propios. Escribiendo estas líneas no puedo olvidarme de ti, **Lore**. Gracias por enseñarme el verdadero significado de luchar y ser valiente. Gracias por demostrarme que hay que llevar la sonrisa por bandera por muy dura que sea la vida. Ojalá estuvieras aquí y escucharte decir bien alto y con orgullo ese "¡Carito, que eres doctora!" que tanto te gustaba gritar. Parece que aún lo escucho. Ojalá estuvieses brindando con nosotros con un tintillo en la mano y contagiándonos tu risa hasta que se nos salten las lágrimas. Espero que hoy estés orgullosa viéndome defender la Tesis desde allí arriba. Seguiremos cumpliendo nuestros sueños y tu estarás presente siempre, desde algún sitio, sonriéndonos. Te echo mucho de menos, amiga.

A mi compañero de viaje desde hace ya 10 años, **Damián**. Gracias por cruzarte en mi camino de forma inesperada y no soltarme la mano desde entonces. Gracias por todo tu amor, cariño y apoyo desde que nos conocimos y en especial en esta última etapa. Has sido el que sin duda más veces has sufrido el tener que aplazar planes porque tenía que sacar la Tesis adelante, siento todas las veces que no salimos a comer, a hacer deporte, a la playa, al cine o a pasar el fin de semana. Mil gracias por tu paciencia infinita y como tu dirías "por aguantarme". Gracias por animarme con tus tonterías y sacarme una sonrisa en los momentos que más lo he necesitado. Por quitarme el ordenador de encima solo para darme un abrazo y que desconectase, y por obligarme a salir a pasear cuando estaba un día entero sin salir de casa. Gracias por todas las veces que aunque a regañadientes, hiciste la compra, cocinaste y limpiaste para que yo pudiese seguir trabajando. Gracias por embarcarte conmigo en la aventura de Olomouc y hacer que fuese una experiencia inolvidable. Por todos y cada uno de los viajes y aventuras varias que hemos realizado, ¡incluso por ser capaz de convencerme para tirarme de un barranco abajo! Gracias por compartir cada una de mis alegrías y, sobre todo, por hacerme sentir mejor a tu lado. Aún nos quedan muchas metas que cumplir juntos, te quiero muchísimo. Gracias también a tu familia por acogerme como una más, en especial a tu madre, y a tu hermana, por hacerme tita de este niño que nos vuelve locos. **Ignacio**, tú no lo sabes, pero has sido la excusa perfecta para cambiar muchos ratos de trabajo por jugar contigo a los "*brum brum*".

Gracias a toda mi familia. En especial a mis tíos **Franci**, **Miguel**, **Mª Carmen** y **Antonio**, a mis primos **Manuel**, **Poy**, **Dani** y **Noru,** y a los pequeñajos (aunque ya no lo son tanto) **Yanira** y **Daniel**. Gracias por todo lo que habéis hecho por mí desde bien pequeña, por vuestro cariño, por hacerme sentir especial, y por cada escapada a diferentes ciudades, a la sierra, a la montaña o al cortijo en las que siempre habéis contado conmigo. Un agradecimiento muy especial a mis abuelos, **Pepe** y **Pepa**. Gracias por vuestra paciencia, vuestro inmenso amor y cariño, y por haberme cuidado siempre desde mi infancia. Gracias por estar pendiente de mí y por vuestros ánimos estos años, aunque no entendieseis porqué tenía que estar tantas horas estudiando y trabajando si ya había acabado la carrera. Qué haría yo sin vuestros mimos, sin esas llamadas por la tarde-noche aunque fuesen para no hablar de nada, esas comidas tan ricas de la abuela o sin las batallitas del abuelo. Os quiero mucho. No puedo olvidarme de agradecer también a mis abuelos **Antonio** y **Carolina**, que aunque ya no estén aquí, seguro que me ven con orgullo desde el cielo.

Por último agradecer a los más importantes, mi hermano y mis padres. A mi hermano pequeño **Jose**, que aunque a menudo estemos peleándonos solo es por no perder la costumbre de llevarnos como el perro y el gato (o mejor dicho, como la Molly y la Babi). Desde pequeño siempre has sido un trasto, pero qué bien lo he pasado contigo. Crecer a tu lado ha sido una de las experiencias más bonitas que me ha regalado la vida. Gracias por estar siempre dispuesto a ayudar a los que te rodeamos, tienes un corazón enorme. Y gracias por todas las pequeñas cosas que haces por mí, como traerme el cargador cuando tengo el ordenador al 1% aunque me quieras matar, hacerme la comida cuando llegaba de la Facultad (aunque algo sencillito tampoco nos pongamos exquisitos), e incluso por sacar a la Babi a pasear para que yo siguiese trabajando. Vas a llegar tan lejos como te propongas enano, y yo estaré siempre a tu lado para ayudarte en lo que necesites. Gracias a mis padres, **Víctor** y **Mª Jesús**. No habría suficientes palabras para agradeceros todo lo que hacéis por mí. Gracias por estar siempre a mi lado, por apoyarme en todos los pasos que he ido dando y por ayudarme en todo lo que podéis y más. Gracias por haberme llevado desde chiquitita a donde hiciese falta para jugar a cualquier deporte que me había dado por probar, pero sobre todo al fútbol sala. Tenéis gran parte de culpa de haber potenciado mi pasión por el deporte y haber podido elegir esta carrera. Gracias por sufrir también la dedicación que me ha llevado esta Tesis, por todas las veces que os he dicho que bajaseis el volumen de la tele, que no hicieseis ruido porque tenía clase online, de las no llamadas por si estaba liada, y por todos los planes a los que he tenido que decir que no estos años. Si hoy puedo estar terminando esta Tesis es gracias a vuestro esfuerzo y la educación que me habéis dado desde pequeña. Para mí sois un modelo y ejemplo a seguir. Gracias por confiar siempre en mí. Soy afortunada de tener los padres que tengo, os quiero muchísimo.

Gracias a todos y todas los que me han acompañado en algún momento de esta aventura. Sin vosotros no habría llegado hasta la meta. Gracias por seguir ahí.



UNIDADES DIDÁCTICAS INNOVADORAS Y USO DE MONITORES PORTÁTILES DE FITNESS PARA LA PROMOCIÓN DE HÁBITOS DE ACTIVIDAD FÍSICA SALUDABLES EN EDUCACIÓN FÍSICA

Tesis Doctoral Internacional
Programa de Doctorado en Ciencias de la Educación
Departamento de Educación Física y Deportiva
Universidad de Granada
2021