

Diseño y validación de un instrumento para analizar los estereotipos corporales representados en el software y aplicaciones digitales (Apps)

Design and validation of an instrument for the analysis of images of digital applications (Apps) related to Physical Activity

José Díaz-Barahona, Teresa Valverde-Esteve, Irene Moya-Mata
Universidad de Valencia (España)

Resumen. La preocupación por el cuerpo y los estereotipos asociados a la actividad física que transmiten las aplicaciones y el software digital es un fenómeno sustancial pero poco investigado. En el presente estudio, se muestra la validación del proceso de adaptación inicial de un instrumento de medida y de los ítems de la escala, denominado REICAF (Representación de Estereotipos sobre la Imagen Corporal y la Actividad Física), que analiza los estereotipos corporales y de actividad física asociados a las imágenes que aparecen en las Apps y software de carácter deportivo. Se utilizó una metodología descriptiva-transversal siguiendo el protocolo Delphi. La elaboración y validación del instrumento se llevó a cabo gracias a la revisión, tres rondas iterativas realizadas por un panel de seis expertos y al posterior pilotaje del instrumento. La consistencia interna y la confiabilidad de la herramienta reportaron valores muy altos ($\alpha=.926$), mostrando valores de $K \geq .626$ en todos los ítems. Este estudio concluye que la herramienta puede ser utilizada por docentes e investigadores y creadores de software para adquirir un juicio crítico sobre los estereotipos mencionados.

Palabras clave: alfabetización visual, estereotipos de género, estereotipos corporales, formación docente, instrumentos de evaluación, técnica Delphi

Abstract. The concern for the body and the stereotypes associated with the physical activity transmitted by digital applications introduced by teachers to energize teaching is a substantial topic, although it has been little researched. This study shows the validation of the initial adaptation process of a measuring instrument and the items of the scale, called 'REICAF' (Stereotypes of Body Image and Physical Activity), to analyse the body and physical activity stereotypes associated with the images that appear in Physical Activity Apps and software. We used a descriptive-cross-sectional methodology through Delphi technique, with the participation of a panel of 6 experts ($n=6$). The elaboration and validation of the instrument was carried out after two rounds of the analysis by the panel of experts and the piloting of the questionnaire. The internal consistency and reliability of the tool reported very high values ($\alpha=.926$), thus showing values of $K \geq .626$ in all items. This study concludes that the tool can be used for teachers and researchers and software creators to acquire critical judgment about the mentioned stereotypes.

Keywords: evaluation instruments, questionnaire, body stereotypes, gender stereotype, values, teacher training, Delphi technique.

Fecha recepción: 06-06-22. Fecha de aceptación: 19-09-22

José Díaz Barahona
jose.diaz-barahona@uv.es

Marco teórico

La literatura evidencia que la institución escolar y su profesorado están sujetos, ahora más que nunca, a procesos, voluntarios o involuntarios, de inmersión digital asociados al uso masivo de medios y servicios digitales, especialmente móviles (Díaz-Barahona et al., 2019), donde la tecnología se percibe como una poderosa herramienta institucionalmente relevante y aceptada (Casey et al., 2016). En este contexto de sobreexposición digital, la institución educativa está integrando nuevos materiales y prácticas curriculares asociadas al uso tecnológico; aunque como advierten Traxler (2016) o Díaz-Barahona (2020), el profesorado todavía no dispone de criterios tecnopedagógicos para integrarlas con eficacia en sus prácticas de aula. A las acreditadas brechas digitales del profesorado (de género, etarias, de posesión, uso o de formación tecnológica) se podría sumar otra vulnerabilidad pedagógica consistente en integrar las TIC sin la suficiente reflexión sobre sus derivas éticas, legales y socioeducativas (Díaz-Barahona, 2020; Lupton, 2014).

Entre las numerosas tecnologías educativas existentes, el mobile learning y sus infinitas aplicaciones (app) y gadgets (tablets, smartphones o wearables de salud) parecen consolidarse en los entornos docentes (Apps Annie, 2021; Ditrendia, 2020). Su empleo como material curricular

crece exponencialmente ayudado por la popularización de las apps, la posesión generalizada de dispositivos móviles y por el potencial didáctico de incontables aplicaciones existentes en el mercado (Caeiro et al., 2020). Estos procesos de inmersión digital educativos, evidenciados en la etapa prepandémica, se han acelerado por la crisis sanitaria mundial por la COVID-19; como refleja el desarrollo sin precedentes de descargas y uso educativo de apps relacionadas con la salud y la actividad físico-deportiva (App Annie, 2021; Ditrendia, 2020), que ya son conocidas o usadas por el 18% del profesorado (Menescardi et al., 2021).

El uso de apps como material curricular en la enseñanza-aprendizaje de la educación físico-deportiva (EFD) está ampliamente documentado (Aznar et al., 2019; Díaz-Barahona, 2020; Goodyear et al., 2019); por las oportunidades que generan en la promoción y en la práctica de la EFD saludable y emancipadora (Díaz-Barahona, 2020; Kerner & Goodyear, 2017) o por los efectos sobre la motivación, la satisfacción corporal, al aumento en la práctica de actividad física en adolescentes (Bice et al., 2016; Kerner & Goodyear, 2017; Gil-Quintana et al., 2022). Sin embargo, la literatura muestra limitaciones y problemas en su integración; relacionados con la privacidad, la seguridad y la protección de datos personales, la cuantificación de los cuerpos o el automonitoreo permanente y control de los

usuarios (Crescenzi-Lanna et al., 2019; Lupton, 2014; Rich & Miah, 2016).

En relevantes trabajos como el de Bourdieu (1998), se advierte que los mensajes asociados al cuerpo masculino y femenino transmitidos por instituciones como la familia y la escuela siguen perpetuando estereotipos y situaciones de discriminación. En el ámbito de la EFD, también se evidencia que el profesorado y el alumnado viene legitimando ciertos estereotipos corporales (Spittle et al., 2012). La importancia del tema hace que, en la última década, proliferen estudios sobre la construcción crítica del cuerpo en diversos materiales curriculares: apps educativas (Díaz-Barahona, et al., 2022), materiales genéricos de aula (Martínez-Bello et al., 2021) o realizados sobre los libros de texto de Educación Física (Moya, 2020). En esta línea, Moya-Mata et al. (2017, 2018) encontraron, después de analizar la presencia de estereotipos corporales en 3836 imágenes, la prevalencia de un modelo masculino fuertemente estereotipado: de raza blanca, joven, con cuerpo ectomorfo, sin discapacidad y realizando actividad físico-deportiva. O de manera más específica, imágenes en donde se detectaban estereotipos de género relacionados con la práctica físico-deportiva: actividad vigorosa asociada al género masculino y actividad menos vigorosa y/o sedentaria, estética, asociada al género femenino (Martínez-Bello et al., 2021; Díaz-Barahona et al., 2022).

Diferentes investigaciones utilizan, para el análisis de las imágenes impresas los estudios observacionales, donde además de una estructura básica propia del método científico, se elaboran instrumentos de observación *ad hoc* (Anguera et al., 2011). Concretamente, en el área de Educación Física (EF) son varios los sistemas de categorías encontrados para el análisis de las imágenes, pero ninguno relacionado con las imágenes que aparecen en las apps deportivas. Es el caso del instrumento SAIMEF (Sistema de Análisis de Imágenes en EF), diseñado por Moya-Mata et al. (2018), para analizar las imágenes de los libros de texto de EF, relacionadas con las características corporales y la actividad física realizada. Concretamente se compone por tres dimensiones relacionadas con las características técnicas de la imagen, las características del personaje y las características de la actividad física. Otro instrumento utilizado en el análisis de la representación visual corporal es el de Martínez-Bello et al. (2021), centrado en el análisis de los estereotipos asociados al cuerpo en el aula de Educación Infantil, analizando variables como el nivel de actividad, la edad, la ropa, los colores estereotipados, la postura corporal o el tipo de material curricular encontrado en estas aulas.

Para Haegele y Zhu (2019), los mensajes subliminales de las imágenes pueden conllevar concepciones negativas de personas o grupos sociales y perpetuar los prejuicios, estereotipos y discriminaciones porque condicionan el pensamiento. Si como advierte Area (2017) asistimos a una metamorfosis de los materiales didácticos consecuencia del proceso de digitalización, es imprescindible trascender al estudio de los materiales curriculares tradiciona-

les (como los libros de texto) y poner el foco investigador en materiales curriculares digitales como las apps, sobre las que no existe control ni regulación institucional y tampoco un uso selectivo y crítico por parte del profesorado (Díaz-Barahona, 2020). Por lo tanto, asumiendo que las imágenes son sistemas de representaciones complejas (Lowe, 2007), necesitan ser objeto de enseñanza específica a través de la alfabetización o competencia visual (López-Manjón & Postigo, 2016) e interpretarlas desde una perspectiva crítica.

A pesar del uso exponencial de software digital relacionado con la actividad física y la salud, el profesorado y alumnado todavía carece de instrumentos y criterios sobre su elección y uso didáctico. Los docentes, de manera masiva pero voluntariosa, están integrando las apps como mediadores didácticos ajenos a los estereotipos que transmiten y legitiman (Díaz-Barahona et al., 2022). Por ello, en este trabajo nos planteamos como objetivo general, diseñar y validar un instrumento para estudiar los estereotipos asociados a la imagen corporal y a la actividad física presentes en las imágenes de aplicaciones (App) usadas en la EFD. Los objetivos específicos son: a) utilizar el protocolo Delphi para validar el contenido; b) realizar una selección de imágenes y realizar una prueba piloto para comprobar la validez del constructo y analizar la consistencia interna y la fiabilidad del instrumento.

Metodología

Se trata de un estudio descriptivo, transversal e instrumental (Almonacid-Fierro et al., 2018; Ato et al., 2013), a través del cual se determinó la validez de contenido y fiabilidad de un instrumento para analizar los estereotipos de género asociados a la representación del cuerpo y a la actividad física. Debido a que es una etapa capital en la construcción del cuestionario, el proceso en la consecución de cada una de las etapas se ha realizado con el máximo rigor y control experimental (Fernández et al., 2014).

Procedimiento: Diseño del cuestionario

El proceso llevado a cabo para la validación del instrumento siguió las directrices de Landeta (1999, 2006) realizando dos tipos de análisis. En primer lugar, aseguramos la validez externa mediante una profunda revisión documental y definición del concepto teórico sobre el que se basaba el instrumento, tomando como referencia cuestionarios validados en estudios similares. En segundo lugar, garantizamos la validez interna mediante el análisis de la utilidad y validez de contenido usando la técnica Delphi para establecer las categorías e indicadores del instrumento. Finalmente, realizamos una selección de apps e imágenes candidatas e hicimos una prueba piloto para comprobar la validez, la consistencia interna y la fiabilidad.

Validación del contenido: Técnica Delphi

Para proceder a la validación interna nos decantamos

por el protocolo Delphi, tradicional proceso de validación de contenido mediante juicio de expertos ampliamente usado en ciencias sociales. Con la validación de contenido, buscábamos evaluar el grado en el que los ítems del instrumento son representativos y relevantes en relación con el constructo que se desea medir (Salcines, 2016), garantizar la inclusión de todos los elementos explicativos de la temática a observar (Okoli & Pawlowski, 2004) y evaluar el grado en el que los ítems del instrumento son representativos y relevantes (Salcines, 2016).

La validación se articuló en tres rondas, realizadas a través de videoconferencia síncrona a través de Blackboard Collaborate (BBC). Optamos por el formato online síncrono por la situación sanitaria provocada por la COVID-19, con sesiones con una duración de 120 minutos máximo, realizadas consecutivamente en tres meses (marzo y junio de 2021). Para la recogida de información, se adoptó la técnica del Focus Grup (Wilkinson, 1998) integrando al grupo de expertos, dos investigadores encargados de moderar y recoger la información relevante (reflexión y determinación sobre las dimensiones, categorías y las propuestas de redacción sobre los indicadores). La información de las sesiones se grabó en BBC y se usó un registro sistemático. Posteriormente se transcribieron las grabacio-

nes y la hoja de registro sistemático para tener en cuenta las opiniones y valoraciones de los jueces en torno a los ítems. De esta forma, se verificó que cada ítem cumpliera los criterios de “adecuación”, “redacción” y “comprensión cualitativa”; es decir, que fueran pertinentes y precisos, además de recoger las aportaciones de los expertos (García-Santos & Ibáñez, 2016).

Para la selección del panel de expertas/os determinamos los criterios selección, calidad e idoneidad del panel y el número necesario, siguiendo las recomendaciones de López (2018) y Landeta (1999), que fijan como condiciones imprescindibles la experiencia, la disposición y su conocimiento con el objetivo del proceso y con el tema en estudio. También consideramos las propuestas de Escobar y Cuervo (2008) y Díaz-Barahona et al. (2019) y cumplieron con al menos 4 de los 5 criterios establecidos: a) ser doctores; b) profesores universitarios o de educación obligatoria en activo; c) con experiencia en formación pedagógica asociada a la tecnología educativa; d) experiencia mínima de 10 años en docencia; e) tener publicaciones sobre la materia. De los 9 jueces inicialmente contactados, finalmente seleccionamos 6 (cuatro hombres y dos mujeres), declinando el resto por diversos motivos (disponibilidad o desmotivación) (Tabla 1).

Tabla 1

Listado y características de jueces y juezas expertas

Experto	Sexo	Formación	Criterios cumplidos	Motivos específicos de selección
1.	Mujer	Dra. Didáctica de la Educación Física	5 de 5	Experta en investigación sobre imagen corporal e innovación en educación y TIC
2.	Mujer	Dra. Pedagogía	5 de 5	Experta en TIC, didáctica general y educación en valores e innovación, educación secundaria
3.	Hombre	Dr. Didáctica de la Educación Física	5 de 5	Experto en investigación didáctica sobre las TIC en educación superior.
4.	Hombre	Dr. Filosofía y Ciencias de la Educación	4 de 5	Experto en TIC en el ámbito educativo no universitario.
5.	Hombre	Dr. Ingeniería	4 de 5	Experto en TIC en aulas universitarias.
6.	Hombre	Dr. Didáctica de la Educación Física	4 de 5	Experto en investigación e innovación en didáctica y mobile learning

Proceso iterativo: protocolo Delphi

Como fase precedente al proceso iterativo del protocolo Delphi, los investigadores, siguiendo a Huertas-Delgado et al. (2019), realizamos una búsqueda y revisión documental de contenido de literatura especializada, desde mayo a junio de 2020, usando descriptores básicos como “estereotipos corporales”, “estereotipos de género” y “estereotipos de práctica física” entre otros, en bases de datos de ciencias sociales como (p.e. Scopus, Web of Science). También recopilamos literatura e instrumentos sobre estereotipos de género asociados a la representación corporal y a la práctica de actividad físico-deportiva con los que se diseñó un pre-cuestionario con 65 indicadores y siete dimensiones, que sirvieron como referente previo.

En la primera ronda se informó a los y las expertas de los objetivos del cuestionario y se sometió el pre-cuestionario a discusión valorando la idoneidad de los indicadores, las categorías y la coherencia entre ambas. Recogimos y anotamos en el registro sistemático, otros indicadores, descriptores y sugerencias aportadas por los expertos. Con la discusión inicial, considerando todos los puntos de vista, recogimos y organizamos dicha información haciendo la

primera versión del instrumento. En esta ronda, como en el resto, consideramos como nivel adecuado de consenso y convergencia de opiniones del panel de expertos un mínimo el 80%, tal como proponen Pozo et al. (2007).

En la segunda ronda, procedimos a la valoración individual de cada indicador. Los expertos debían determinar si eran o no coherentes, relevantes y/o claros como propone Landeta (2006). *La coherencia*: indicaba si la descripción y el indicador evaluado tenía relación lógica con la dimensión o concepto evaluado; *la relevancia*: si el indicador evaluado debía ser incluido por su importancia, adecuación o significado para representar a la dimensión evaluada; *la claridad*: si el enunciado del indicador se entendía y el vocabulario, la sintaxis y la semántica eran claras. También se sometieron a discusión y evaluación, los criterios de aplicabilidad y conveniencia de cada categoría.

Como resultado de dicho proceso, el panel de expertos acordó la modificación de las siguientes categorías e indicadores: a) En la dimensión I: «características técnicas de la imagen» se incluye la categoría «tipo de imagen» con los indicadores «fotografía» y «dibujo»; b) En la dimensión II: «características del cuerpo humano», se modifica la cate-

goría «sexo-agrupación» por «género-agrupación»; en la categoría «edad» se agrupan los indicadores «niñez/adolescencia»; en la categoría «diversidad funcional» se añade el indicador «no se distingue»; la categoría «indumentaria» es modificada por «indumentaria deportiva», especificando en la definición que se hace referencia a la vestimenta superior del cuerpo humano, y se incluyen los indicadores «camiseta de tirantes» y «camiseta de manga corta»; en la categoría «color de la indumentaria deportiva» se especifica en la categoría «azul (colores oscuros)», en la categoría «rosa (colores suaves)» y se redefine la categoría «neutro»; y en la categoría «cabello» se incluyen los indicadores «largo» y «coleta», con las respectivas definiciones; c) Por último, en la dimensión III: «características de la actividad física», se incluye en la categoría «nivel de actividad», el indicador «media» y su definición correspondiente.

La tercera ronda, siguiendo los criterios para la finalización del protocolo Delphi que propone Landeta (1999), buscamos alcanzar el consenso basado en la convergencia entre las opiniones de los participantes y la estabilidad

en sus respuestas. Para ello, se facilitó el cuestionario al panel de expertos y un moderador-investigador dirigió la discusión final aclarando las dudas y buscando la aprobación. Se decidió incluir, en todas las dimensiones y categorías, un nuevo indicador denominado “no se distingue” para identificar aquellas imágenes que generaban dudas sobre su clasificación. Finalmente alcanzamos el consenso, sin nuevas modificaciones, sobre las dimensiones, las categorías y la redacción de los indicadores del instrumento final, denominándolo: Representación de Estereotipos sobre la Imagen Corporal y la Actividad Física (REICAF).

El instrumento

REICAF se estructura en torno a tres dimensiones: 1) Características técnicas de la imagen, dividida en dos categorías y nueve indicadores; 2) Características del cuerpo humano representado, dividida en nueve categorías y 52 indicadores; y 3) Características de la actividad física representada, dividida en dos categorías y 12 indicadores de análisis. En total, REICAF se estructura en tres dimensiones, 13 categorías y 73 indicadores (Tabla 2).

Tabla 2

Instrumento: Representación de Estereotipos de Imagen Corporal y de Actividad Física (REICAF)

Dimensiones, categorías e indicadores			
Dimensiones	Categorías	Descripción	Indicadores
Características técnicas de la imagen	Tipo de imagen	Formato de la imagen (2 indicadores)	Fotografía Dibujo
	Tipo de plano	Proporción que tiene la imagen dentro del encuadre (7 indicadores)	General
			Entero
			Americano o ¾ Medio Medio corto Primer plano o detalle No se distingue
Sexo-agrupación	Condición biológica y natural que diferencia entre hombres y mujeres, y distribución (6 indicadores)	Masculino Femenino Grupo de hombres Grupo de mujeres Grupo mixto No se distingue	
Características del cuerpo humano	Edad	Tiempo vivido por un sujeto, determinado por su apariencia externa (6 indicadores)	Niñez/Adolescencia Juventud Madurez Vejez Varias No se distingue
	Raza	Grupo en que se subdivide la raza humana biológicamente según el color de la piel, los ojos y/o ancho de la nariz (5 indicadores)	Blanca Negra Otras Varias No se distingue
	Somatotipo	Forma, estructura o morfología corporal de un sujeto (5 indicadores)	Ectomorfo Endomorfo Mesomorfo Varios No se distingue
	Diversidad funcional	Funcionamiento de una persona al realizar las tareas habituales, de manera diferente a la mayoría de la población (3 indicadores)	Presencia Ausencia No se distingue
Indumentaria deportiva	Vestimenta superior de los cuerpos representados (8 indicadores)		Camiseta de tirantes Camiseta de manga corta Top Torso desnudo Desnuda Varias Otras No se distingue
			Color de la indumentaria deportiva

		Neutro Varios No se distingue
		Sin cabello Corto Media melena Largo Coleta Moño Varios Otros No se distingue
	Cabello	Formas de los cabellos/pelos de la cabeza de una persona (9 indicadores)
	Postura corporal	Posición adoptada por el cuerpo en relación con la acción o actividad física realizada (5 indicadores)
		Estereotipo masculino Estereotipo femenino Neutral Varias No se distingue
Características de la actividad física	Actividad realizada	Intencionalidad de la acción ejecutada por la persona, en relación con la mejora de las cualidades físicas (6 indicadores)
	Nivel de actividad	Intensidad Gasto o consumo energético que representa la imagen o cuerpo representado (6 indicadores)
		Fuerza Resistencia Velocidad Flexibilidad Varias Sin actividad física Sedentaria Ligera Media Intensa Varias No se distingue

Fiabilidad y validez: la prueba piloto del instrumento

Obtención de la muestra de apps e imágenes y la prueba piloto

Para garantizar la fiabilidad y validez del instrumento realizamos una prueba piloto, seleccionando apps e imágenes candidatas utilizadas en el ámbito de la EFD, hibridando los protocolos propuestos por Stoyanov et al. (2015) y Moher et al. (2009).

Las apps se seleccionaron mediante una búsqueda sistemática de aplicaciones a través del navegador *Google Chrome*, motor de búsqueda Google, en la base de datos Google Play Store, utilizando dos ordenadores de sobremesa (HPi5). Como la Google Play Store agrupa las apps en 35 categorías generales, se utilizaron cinco descriptores específicos relacionados con el objeto del estudio: «ejercicio físico», «fitness», «entrenamiento físico», «condición física» y «calentamiento» acotando la búsqueda. Los criterios de inclusión fueron: 1) seleccionar las primeras 250 apps disponibles en la web de Google Play Store en España, 2) gratuitas; 3) en idioma castellano, francés o inglés; y 4) dedicadas al desarrollo de la condición física (fuerza, resistencia, velocidad y/o la flexibilidad). Se excluyeron: 1) apps duplicadas; 2) con nombres, funciones irrelevantes y las no relacionadas con el objeto de estudio (estadísticas, marcadores, marcas deportivas, aspectos técnicos-tácticos o similares; 3) con menos de 10.000 descargas; 4) con una valoración de menos de 3,5 estrellas; y 5) no actualizadas en los 12 meses anteriores a la fecha de selección. De las 1,250 apps iniciales, correspondientes a los cinco descriptores buscados, tras aplicar los filtros de inclusión-exclusión, quedaron finalmente 161 apps reclutadas.

Para la selección de las imágenes candidatas, de cada app reclutada, se hizo una captura de pantalla de la imagen del logo y las tres primeras imágenes (fotografías e imágenes

vectoriales) que representaban cuerpos humanos realizando actividad física. Cuando en las imágenes se mostraba más de una figura humana, se seleccionaba la situada en la parte superior izquierda. Mediante observación directa se numeraron y clasificaron cada una de las imágenes. Las imágenes que presentaron dudas acerca del tipo de actividad física representada fueron eliminadas. Este exhaustivo procedimiento se hizo para cumplir las expectativas de rigor recomendadas por Stoyanov et al. (2015) para trabajos similares.

La prueba piloto

Para conocer las propiedades psicométricas del instrumento, del total de 161 apps, tomamos una muestra aleatoria del 20% (n= 33) a través del portal web de aleatorización (<https://www.randomizer.org/>), que se utilizaron con fines de entrenamiento y pilotaje.

El procedimiento de entrenamiento se realizó como proponen Stoyanov et al. (2015), mediante análisis de dos expertos investigadores: (E1) profesor universitario de EF con 15 años de experiencia en didáctica de la EF, experto en tecnología educativa y autor de numerosas publicaciones sobre mobile learning y (E2), doctora en EF profesora universitaria con 7 años de experiencia docente general y en TIC. Primero analizaron conjuntamente 30 imágenes, hasta lograr unidad de criterio y consenso (sobre conceptos, expresiones, comprensión, agrupación de ítems) como proponen Landis & Koch (1977) y después, por separado, realizaron el análisis piloto de las 33 apps y sus respectivas imágenes (n=120).

Tras ello, realizamos el pilotaje sobre una muestra representativa (20%) de las aplicaciones, realizado por los dos expertos, en dos rondas anteriores a la definitiva, mostrada en el apartado de resultados. Su finalidad era verificar la validez del instrumento y reconsiderar la elimi-

nación de aquellos ítems cuyos valores de concordancia entre expertos, evaluada a través de la Kappa de Cohen, fueran inferiores a .60 (Blackman & Koval, 2000; Landis & Kolsh, 1977), así como la posibilidad de considerar un ítem como constante cuando su valor siempre fuera $K=1.0$. Los valores de la Kappa de Cohen en las dos vueltas previas a la definitiva fueron de .810 (vuelta 1) y .962 (vuelta 2).

Análisis y tratamiento de los datos

El análisis estadístico fue llevado a cabo a través del software SPSS, versión 26.0 (IBM, Chicago, IL, USA). El análisis fue realizado en las tres rondas, mostrando así los resultados definitivos en este estudio. La consistencia interna de la escala fue analizada a través de la Alpha de Cronbach (α) (Santos, 1999), considerando la adecuación de aquellos ítems cuya $\alpha \geq .7$ (Celina & Campo-Oviedo, 2005). La fiabilidad del instrumento inter-observadores, entendida como la observación de un mismo elemento llevada a cabo por dos observadores distintos (Anguera et al., 2018), fue analizada mediante dos vías: a) a través del coeficiente de correlación intraclase (ICC) (Fleiss & Cohen, 1973), considerando el valor $< .5$: mala confiabilidad; $.5-.75$: moderada confiabilidad; $.75$ y $.9$: buena fiabilidad; $> .90$: fiabilidad excelente (Koo y Li, 2016); b) también por el índice Kappa de Cohen (McHugh, 2012). Para la interpretación de los valores Kappa, seguimos los criterios establecidos por Landis & Kolsh (1977): $< .20$ (ligera), $.21-.40$ (baja), $.41-.60$ (moderada), $.61-.80$ (buena) y $.81-1.00$ (muy buena). En todos los casos, se estableció el nivel de intervalo de confianza al 95% y nivel de significación de $p < .05$.

Resultados

Análisis de la fiabilidad

Los valores que se muestran en este apartado corresponden a la tercera y última vuelta del análisis de los resultados mostrados por los dos expertos. En esta ocasión, el valor de la consistencia de todo el instrumento fue muy elevado ($\alpha = .926$), siendo los valores más altos observados en la raza ($\alpha = .998$), la edad ($\alpha = .958$), el sexo-agrupación ($\alpha = .949$), el cabello ($\alpha = .948$), y los más bajos en la indumentaria ($\alpha = .832$) o la actividad realizada ($\alpha = .888$) (Tabla 3).

En la Tabla 4, se muestra el valor de la fiabilidad del instrumento, a través del índice de Kappa de Cohen y su nivel de significación. Como se puede observar, los valores más altos se encontraron en las características del cuerpo humano, especialmente en la raza ($K = .956$; $p \leq .001$), la edad ($K = .919$; $p \leq .001$) o la agrupación del sexo ($K = .935$; $p \leq .001$).

Tomando como referencia los valores estadísticos propuestos, el instrumento presenta un alto grado de consistencia interna y fiabilidad, siendo todos los valores altamente significativos ($p \leq .001$).

Tabla 3

Resultados de la Alpha de Cronbach, el coeficiente de correlación intraclase y el intervalo de confianza entre los elementos del instrumento REICAF

	Alpha de Cronbach	Coefficiente de correlación intraclase	Intervalo de confianza
Características técnicas de la imagen			
Tipo de plano	.929	.929	.898 – .951
Características del cuerpo humano			
Sexo-agrupación	.949	.949	.927 – .964
Edad	.958	.958	.940 – .971
Raza	.998	.998	.998 – .999
Somatotipo	.861	.861	.801 – .903
Diversidad funcional	.907	.907	.867- .935
Indumentaria	.832	.834	.761 – .884
Color de la indumentaria deportiva	.888	.888	.840 – .922
Cabello	.948	.947	.924 - .963
Postura Corporal	.892	.890	.841 - .923
Características de la actividad física			
Actividad realizada	.876	.874	.818 - .912
Nivel AF	.906	.904	.861 - .933

Tabla 4

Resultados del valor de la Kappa de Cohen y el nivel de significación entre los ítems

	Kappa de Cohen	Nivel de significación
Características técnicas de la imagen		
Tipo de plano	.646	$p \leq .001$
Características del cuerpo humano		
Sexo-agrupación	.935	$p \leq .001$
Edad	.919	$p \leq .001$
Raza	.956	$p \leq .001$
Somatotipo	.762	$p \leq .001$
Diversidad funcional	.829	$p \leq .001$
Indumentaria deportiva	.626	$p \leq .001$
Color de la indumentaria deportiva	.748	$p \leq .001$
Cabello	.895	$p \leq .001$
Postura Corporal	.725	$p \leq .001$
Características de la actividad física		
Actividad realizada	.781	$p = .001$
Nivel AF	.736	$p = .001$

Discusión

La validación de REICAF por juicio de expertos (método Delphi) ha mejorado la pertinencia de sus categorías y dimensiones, aportando coherencia y claridad a los ítems y a la estructura del instrumento. La participación complementaria de expertos tecnólogos y educadores físicos proporcionó al instrumento una dimensión específica, de la que carecen en otros cuestionarios esencialmente de carácter tecnocéntrico (Aesaert et al., 2014; Ambrós et al., 2013), al integrar el conocimiento tecnológico y el pedagógico.

El proceso de entrenamiento y pilotaje, así como el análisis estadístico de la fiabilidad de la muestra fue realizado en dos expertos investigadores, siguiendo los criterios recomendados por Neuendorf (2011) y Stoyanov et al. (2015). Los resultados derivados del análisis del Alpha de Cronbach para todo el instrumento fueron de $\alpha = .926$, lo que demuestra que el instrumento es altamente consistente. En este sentido, en el estudio de Moya-Mata et al. (2018) se halló un valor de $\alpha = .63$, considerado también como fiable y válido, lo que sugiere que la precisión de los instrumentos ha sido mejorada en los últimos años.

Asimismo, los valores de la Kappa de Cohen, también fueron altamente fiables, presentando valores en relación a las características de la imagen ($K = .646$), características

del cuerpo humano ($.626 \leq K \leq .956$), o características de la actividad física ($.736 \leq K \leq .781$). Estos resultados son similares a los estudios de Moya-Mata, et al. (2018), quienes hallaron un valor de $K = .81$ en el instrumento SAIMEF o el de Moya-Mata et al. (2016), quienes hallaron un valor de $K = .83$ al analizar “las otras imágenes” en libros de texto de la especialidad de EF. En el estudio realizado por Bernabé et al. (2021), en el que analizaron los estereotipos de género en libros de texto del área musical, los valores de fiabilidad oscilaron entre $K = .64$ y $K = .90$ para las variables analizadas, siendo asimismo el rango entre el que se encuentran los resultados de nuestro estudio.

Como cualquier investigación, el estudio tiene lógicas limitaciones, como analizar únicamente imágenes de apps de la Play Store de España, y no integrar app de otros repositorios internacionales, que en ocasiones ofrecen diferente información. Pero también puede tener aplicabilidad entre la comunidad científica y educativa. A los docentes, les puede orientar en la selección de materiales curriculares (apps) con criterios educativos y servir para el análisis de otros materiales curriculares tradicionales como libros de texto. A las y los investigadores, les propone líneas de estudio, debate y reflexión sobre la equidad, la perspectiva de género y los estereotipos subyacentes en los materiales curriculares. A los y las creadoras de apps puede ayudarles a diseñar productos que, además de la usabilidad tecnológica o la privacidad, también integren referentes educativos que ayuden eliminar los estereotipos corporales.

Conclusiones

Los resultados del proceso de validación y fiabilidad seguidos en la investigación confirman que el instrumento para el estudio de estereotipos de imagen corporal y de actividad física (REICAF), resulta válido y fiable para analizar la representación de los estereotipos de imagen corporal y de actividad física presentes en las aplicaciones móviles empleadas en la EFD ya que los ítems ofrecen altos niveles de consistencia interna, rigor y de fiabilidad interobservadores.

Como proponen Tirado-Morueta & Aguaded-Gómez (2014) o Díaz-Barahona et al. (2022), el discurso educativo posmoderno no debe fundamentarse solo en un cambio instrumental, de apropiación tecnológica o de acceso a servicios y medios digitales. También conviene analizar los estereotipos, los valores y las creencias que legitima el profesorado de educación físico-deportiva al usar materiales curriculares. Para dicho análisis, necesitamos instrumentos que orienten sobre las repercusiones educativas y éticas de los procesos de inmersión digital.

Referencias

- Almonacid-Fierro, A., Molina, S. F., & Carrizosa, M. V. (2018). Validación de un cuestionario para medir el Conocimiento Didáctico del Contenido en el profesorado de Educación Física. *Retos: nuevas tendencias en educación física, deporte y recreación*, (34), 132-137.
- Anguera, M.ª T., Blanco, A., Hernández-Mendo, A., & Losada, J. L. (2011). Diseños observacionales: ajuste y aplicación en psicología del deporte. *Cuadernos de psicología del deporte*, 11(2), 63-76.
- App Annie (Ed) (2021). State of mobile 2021. <https://bit.ly/3vAzP6I>.
- Area, M. (2017). La metamorfosis digital del material didáctico tras el paréntesis Gutenberg. *Revista Latinoamericana De Tecnología Educativa - RELATEC*, 16(2), 13-28. <https://doi.org/10.17398/1695-288X.16.2.13>
- Ato, M., López-García, JJ, & Benavente, A. (2013). Un sistema de clasificación para los diseños de investigación en psicología. *Anales de Psicología*, 29(3), 1038-1059.
- Aznar, I., Cáceres, M.P., Trujillo, J.M., & Romero, J.M. (2019). Impacto de las apps móviles en la actividad física: un metaanálisis. *RETOS: nuevas tendencias en educación física, deporte y recreación*, (36), 52-57.
- Bernabé-Villodre, M. D. M. B., Gil, D. G., & Martínez-Bello, V. (2021). Pervivencia de estereotipos de género en las imágenes de los libros de texto de música españoles. *Educação e Pesquisa*, 47, 1-25. <https://doi.org/10.1590/S1678-4634202147236340>
- Bice, M.R., Ball, J.W., & McClaran, S. (2016). Technology and physical activity motivation. *International Journal of Sport and Exercise Psychology*, 14(4), 295-304. <https://doi.org/10.1080/1612197X.2015.1025811>.
- Blackman, N. J. M., & Koval, J. J. (2000). Interval estimation for Cohen's kappa as a measure of agreement. *Statistics in medicine*, 19(5), 723-741.
- Bourdieu, P. (1998). *La domination masculine*. Éditions de Seuil. París.
- Caeiro, M., Ordóñez, F. F., Callejón, M. D., & Castro-León, E. (2020). Diseño de un instrumento de evaluación de aplicaciones digitales (Apps) que permiten desarrollar la competencia artística. *Pixel-Bit. Revista de Medios y Educación*, 58, 7-25. <https://doi.org/10.12795/pixelbit.74071>.
- Casey, A., Goodyear, V. A., & Armour, K. M. (2016). Rethinking the relationship between pedagogy, technology and learning in health and physical education. *Sport, education and society*, 22(2), 288-304. <https://doi.org/10.1080/13573322.2016.1226792>.
- Celina, H., & Campo-Arias, A. (2005). An approach to the use of Cronbach's Alfa. *Revista Colombiana de Psiquiatría*, 34(4), 572-580.
- Crescenzi-Lanna, L., Valente, R., & Suárez-Gómez, R. (2019). Aplicaciones educativas seguras e inclusivas: La protección digital desde una perspectiva ética y crítica. *Comunicar*, 61, 93-102. <https://doi.org/10.3916/C61-2019-08>.
- Díaz-Barahona, J. (2020). Retos y oportunidades de la tecnología móvil en la educación física. *Retos: nuevas tendencias en educación física, deporte y recreación*, (37),

- 763-773.
<https://doi.org/10.47197/retos.v37i37.68851>.
- Díaz-Barahona, J., Molina-García, J., & Monfort-Pañego, M. (2019). Estudio de las actitudes y el interés de los docentes de primaria de educación física por las TIC en la Comunidad Valenciana. *Retos: nuevas tendencias en educación física, deporte y recreación*, (35), 267-272.
<https://doi.org/10.47197/retos.v0i35.63355>.
- Díaz-Barahona, J., Valverde, T., & Moya, I. (2022). Estereotipos asociados al cuerpo humano: análisis de aplicaciones móviles usadas en la educación físicodeportiva. *Pixel-Bit, Revista de Medios y Educación*, (64), 79-103.
<https://doi.org/10.12795/pixelbit.90297>.
- Ditrendia (Ed) (2020). Informe Mobile en España y en Mundo. <https://bit.ly/3i4OOBJ>.
- Etxeazarra, I., Castellano, J., & Usabiaga, O. (2013). Aplicación de diferentes estrategias para el control de calidad del dato de una herramienta observacional en fútbol formación. *Revista Iberoamericana de Psicología del Ejercicio y el Deporte*, 8(2), 301-316.
- Fernández, P., Vallejo, G., Livacic-Rojas, P., & Tuero, E. (2014). Validez Estructurada para una investigación cuasiexperimental de calidad. Se cumplen 50 años de la presentación en sociedad de los diseños cuasiexperimentales. *Anales de Psicología*, 30(2), 756-771.
<https://doi.org/10.6018/analesps.30.2.166911>.
- Fleiss, J. L., & Cohen, J. (1973). The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educational and psychological measurement*, 33(3), 613-619.
<https://doi.org/10.1177/001316447303300309>.
- García-Santos, D., & Ibáñez, S. J. (2016). Diseño y validación de un instrumento de observación para la valoración de un árbitro de baloncesto (IOVAB). *SPORT TK-Revista EuroAmericana de Ciencias del Deporte*, 5(2), 15-26.
- Gil-Quintana, J., Felipe-Ruiz, R., & Moreno-Muro, M.A. (2022). Influencers deportivos y su repercusión en el consumo, la actividad física y su proyección en redes sociales por los adolescentes andaluces. *Retos: nuevas tendencias en educación física, deporte y recreación*, 43, 591-602.
- Goodyear, C.K., & Quennerstedt, M. (2019). Young people's uses of wearable healthy lifestyle technologies; surveillance, self-surveillance and resistance. *Sport, Education and Society*, 24(3), 212-225.
<https://doi.org/10.1080/13573322.2017.1375907>.
- Haegele, J., & Zhu, X. (2019). Body image and physical education: Reflections of individuals with visual impairments. *European Physical Education Review*, 25(4), 1002-1016.
<https://doi.org/10.1177/1356336X18789436>.
- Huertas-Delgado, F. J., Molina-García, J., Van Dyck, D., & Chillón, P. (2019). A questionnaire to assess parental perception of barriers towards active commuting to school (PABACS): Reliability and validity. *Journal of Transport & Health*, 12, 97-104.
<https://doi.org/10.1016/j.jth.2018.12.004>.
- Kerner, C., & Goodyear, V. A. (2017). The Motivational Impact of Wearable Healthy Lifestyle Technologies: A Self-determination Perspective on Fitbits With Adolescents. *American journal of health education*, 48(5), 287-297.
<https://doi.org/10.1080/19325037.2017.1343161>.
- 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.
<https://doi.org/10.1016/j.jcm.2016.02.012>.
- Korkmaz, H. E., & Erden, M. (2014). A Delphi Study: The Characteristics of Democratic Schools. *The Journal of Educational Research*, 107(5), 365-373.
<https://doi.org/10.1080/00220671.2013.823365>.
- Landeta, J. (1999). *El método Delphi: una técnica de previsión para la incertidumbre*. Barcelona: Ariel.
- Landeta, J. (2006). Current validity of the Delphi method in social sciences. *Technological Forecasting and Social Change*, 73(5), 467-482.
<https://doi.org/10.1016/j.techfore.2005.09.002>.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 159-174.
- López, E. (2018). El método Delphi en la investigación actual en educación: una revisión teórica y metodológica. *Educación XXI*, 21(1), 17-40, <https://doi.org/10.5944/educXX1.20169>.
- López-Manjón, A., & Postigo, Y. (2016). ¿Qué libro de texto elegir? La competencia visual en las actividades con imágenes. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias*, 13(1), 84-101.
http://dx.doi.org/10.25267/Rev_Eureka_ensen_divulg_cienc.2016.v13.i1.07.
- Lowe, R.K. (2007). *Educational illustrations*. Western Australia: Savant Publications.
- Lupton, D. (2014a). Critical perspectives on digital health technologies. *Sociology Compass*, 8(12), 1344-1359.
<https://doi.org/10.1111/soc4.12226>.
- Martínez-Bello, V. E., Cabrera Y., Díaz-Barahona, J., & Bernabé-Villodre, M. (2021) Bodies in the early childhood education classroom: a Bourdieusian analysis of curricular materials, *Sports, Education and Society*, 26(1), 29-44.
<https://doi.org/10.1080/13573322.2019.1690442>.
- McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica*, 22(3), 276-282.
- Menescardi, C., Suárez-Guerrero, C., & Lizandra, J. (2021). Formación del profesorado de educación física en el uso de aplicaciones tecnológicas. *Apunts. Educación física y deportes*, 2(144), 33-43.
[https://doi.org/10.5672/apunts.2014-0983.es.\(2021/2\).144.05](https://doi.org/10.5672/apunts.2014-0983.es.(2021/2).144.05).
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., & The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The

- PRISMA Statement. *PLoS Medicine*, 6(7), 1–6. <https://doi.org/10.1371/journal.pmed.1000097>.
- Moya, I. (2020). Los libros de texto de educación física en la etapa de educación primaria: análisis de imágenes y estereotipos. *Apuntes*, 141, 92-92.
- Moya-Mata, I., Martín, J., Ruiz, L., & Ros, C. (2018). Diseño, fiabilidad y validez de una herramienta para el análisis de las imágenes de los libros de texto de Educación Física. *Retos: nuevas tendencias en educación física, deporte y recreación*, (34), 240-246.
- Moya-Mata, I., Ros, C., & Peirats, J. (2018). ¿Qué representan las portadas de los libros de texto de Educación Física en Primaria? *Retos: Nuevas Tendencias en Educación Física, Deporte y Recreación*, (34), 295–299. <https://doi.org/10.47197/retos.v0i34.63412>.
- Moya-Mata, I., Ruiz, L., Martín, J., Pérez, P.M., & Ros, C. (2017). La representación de la discapacidad en las imágenes de los libros de texto de Educación Física: ¿inclusión o exclusión? *Retos: Nuevas Tendencias en Educación Física, Deporte y Recreación* (32), 88–95. <https://doi.org/10.47197/retos.v0i32.52191>.
- Moya-Mata, I., Ruiz-Sanchis, L., & Ros, C. (2016). Diseño y validación de un instrumento para analizar las “otras imágenes” en los libros de texto de Educación Física. *Rebescolar. Revista Brasileira de Educação Física Escolar*, 3, 8-22.
- Neuendorf, K. A. (2011). Content analysis. A methodological primer for gender research. *Sex roles*, 64(3), 276-289. <https://doi.org/10.1007/s11199-010-9893-0>.
- Okoli, C., & Pawlowski, S. D. (2004). The Delphi method as a research tool: an example, design considerations and applications. *Information & management*, 42(1), 15-29. <https://doi.org/10.1016/j.im.2003.11.002>.
- Pérez, L., Llorente, E., Gavidia, V., Caurín, C. y Martínez, M. J. (2015). ¿Qué enseñar en la educación obligatoria acerca de la alimentación y la actividad física?: Un estudio con expertos. *Enseñanza de las Ciencias*, 33(1), 85-100. <http://dx.doi.org/10.5565/rev/ensciencias.1355>.
- Pozo, M. T., Gutiérrez, J. y Rodríguez, C. (2007). El uso del método Delphi en la definición de los criterios para una formación de calidad en animación sociocultural y tiempo libre. *Revista de Investigación Educativa*, 25(2), 351-366.
- 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>.
- Salcines Talledo, I., & González Fernández, N. (2015). Diseño y Validación del Cuestionario “Smartphone y Universidad. Visión del Profesorado” (SUOL). *Revista Complutense de Educación*, 27(2), 603-632. https://doi.org/10.5209/rev_RCED.2016.v27.n2.46912.
- Santos, J. R. A. (1999). Cronbach’s alpha: A tool for assessing the reliability of scales. *Journal of extension*, 37(2), 1-5.
- Spittle, M., Petering, F., Kremer, P., & Spittle, S. (2012). Stereotypes and self-perceptions of physical education pre-service teachers. *Australian Journal of Teacher Education*, 37(1), 19-42. <http://dx.doi.org/10.14221/ajte.2012v37n1.5>.
- Stoyanov, S. R., Hides, L., Kavanagh, D. J., Zelenko, O., Tjondronegoro, D., & Mani, M. (2015). Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR mHealth and uHealth*, 3(1), 27. <https://doi.org/10.2196/mhealth.3422>.
- Tirado-Morueta, R., & Aguaded-Gómez, J. I. (2014). Influencias de las creencias del profesorado sobre el uso de la tecnología en el aula. *Revista de Educación*, (363), 230-255.
- Traxler, J. (2016). Inclusion in an age of mobility. *Research in Learning Technology*, (24), 1-13. <https://doi.org/10.3402/rlt.v24.31372>.
- Wilkinson, S. (1998). Focus group methodology: a review. *International journal of social research methodology*, 1(3), 181-203. <https://doi.org/10.1080/13645579.1998.10846874>.