

Diseño y validación de un instrumento para evaluar la calidad de las apps gamificadas que promueven hábitos saludables en el alumnado de educación primaria y secundaria

Design and validation of an instrument to evaluate the quality of gamified apps that promote healthy habits in primary and secondary school students

Claudio Delgado-Morales
Universidad de Huelva. Huelva, España
claudio.delgado@dedu.uhu.es

Ana Duarte-Hueros
Universidad de Huelva. Huelva, España
duarte@dedu.uhu.es

María de los Ángeles Merino-Godoy
Universidad de Huelva. Huelva, España
angeles.merino@denf.uhu.es

Sara Conde-Vélez
Universidad de Huelva. Huelva, España
sara.conde@dedu.uhu.es

Resumen

La promoción de la salud en edades tempranas a través de la tecnología gamificada ha emergido en los últimos años como una estrategia innovadora con grandes posibilidades educativas para mejorar áreas específicas de los hábitos saludables como la higiene, la alimentación, la actividad física o la educación digital entre otras, siendo necesaria la implicación de todos los ámbitos que rodean a niños, niñas y adolescentes, especialmente del contexto familiar, con el fin de alcanzar un bienestar integral. En concreto, destacan las aplicaciones móviles (apps) que incorporan contenidos relacionados con la salud y están dirigidas a niños, niñas y adolescentes. Sin embargo, la gran mayoría no cuentan con respaldo científico basado en conocimiento experto. Este estudio busca crear un instrumento sólido y fiable para que las familias evalúen la calidad de este software, asegurando su uso educativo y seguro. Se trata de un cuestionario de 39 ítems organizados en seis dimensiones. Tras una prueba piloto con 33 familias, los resultados de los análisis estadísticos realizados han revelado unas propiedades psicométricas muy favorables, sugiriendo la utilidad del instrumento para apoyar a las familias en la selección de este tipo de apps educativas.

Palabras clave: validación, instrumento, aplicaciones móviles, hábitos saludables, menores.

Abstract

The promotion of health in early ages through gamified technology has emerged in recent years as an innovative strategy with great educational potential to improve specific areas of healthy habits such as hygiene, nutrition, physical activity, or digital education, among others, requiring the involvement of all environments surrounding children and adolescents, especially the family context, in order to achieve overall well-being. Specifically, mobile applications (apps) that incorporate health-related content and target children and adolescents are notable. However, the vast majority lack scientific support based on expert knowledge. This study aims to create a robust and reliable instrument for families to assess the quality of this kind of software,

ensuring its educational and safe use. The instrument consists of a questionnaire with 39 items organized into six dimensions. Following a pilot test with 33 families, the results of the statistical analyses conducted have revealed highly favorable psychometric properties, suggesting the utility of the instrument in supporting families in selecting such educational apps.

Key words: validation, instrument, mobile applications, healthy habits, minors.

1. Introducción

Los hábitos saludables, entendidos como un fenómeno social, adquieren importancia a lo largo de todas las etapas de la vida. Estos tienen su origen en una declaración de la Organización Mundial de la Salud [OMS] (2004) en la cual, mediante la *Estrategia mundial sobre régimen alimentario, actividad física y salud*, se destacó la necesidad de hacer frente a la alimentación poco saludable y el sedentarismo, identificados como factores de riesgo en niños, niñas y adolescentes, con el propósito de prevenir afecciones o enfermedades. Al respecto, los hábitos saludables se definen como el conjunto de comportamientos tanto individuales como colectivos orientados a alcanzar una calidad de vida óptima y una salud integral, caracterizada por ausencia total de enfermedades (Salas-Cabrera, 2015).

Las áreas clave de los hábitos saludables incluyen higiene, alimentación, actividad física, prevención de adicciones, rutinas de descanso, educación sexual, educación emocional, educación vial, educación en valores, educación ambiental, educación digital, así como aprovechamiento del ocio y tiempo libre (Delgado-Fernández y Tercedor-Sánchez, 2021; Ramos-Gordillo, 2003; Sánchez-Álvarez, 2007).

En este sentido, cada vez es mayor el uso de la tecnología, especialmente de los dispositivos móviles como *smartphones* o *tablets*, por parte de todos los sectores de la población, para gestionar diversos aspectos de la vida, como soporte laboral, comunicarse, para entretenimiento, en el contexto educativo, etc.

En concreto, en los últimos años se han incrementado exponencialmente el desarrollo y la difusión de *apps* que incorporan contenidos sobre salud y son pertenecientes al ámbito de la salud móvil (del término en inglés *mobile health; mHealth*). Este concepto fue adoptado por la OMS (2011) y tiene que ver con la práctica médica y de salud pública apoyada por dispositivos móviles como *smartphones*, *tablets* o relojes inteligentes entre otros.

Considerando las aportaciones de autores y autoras como Alexandrou et al. (2023), Thornton et al. (2021) y Zahid et al. (2020), las *apps* ofrecen posibilidades educativas muy interesantes para motivar a niños, niñas y adolescentes a adoptar comportamientos saludables. Asimismo, en muchas de ellas se integran técnicas de juego que combinan el entretenimiento y el aprendizaje a través de elementos gamificados, lo que da lugar a grandes ventajas para la educación (Brown et al., 2022).

Zichermann y Cunningham (2011) explican que en la gamificación se recurre al juego y a ciertas características vinculadas con la acción de jugar para lograr unos determinados resultados de aprendizaje y una modificación de la conducta en situaciones que por sí mismas no son lúdicas.

Recientemente, se han intensificado los proyectos de diseño y desarrollo de *apps* gamificadas centradas en la promoción de hábitos saludables en niños, niñas y adolescentes (Brown et al., 2020; Campos et al., 2019; Fuemmeler et al., 2020; Kim et al., 2021; Martin et al., 2020; Mehta et al., 2021; Schmidt et al., 2020).

Sobre este tipo de *apps*, Chan et al. (2017) y Jeminiwa et al. (2019) ponen de relieve que la población más joven prefiere utilizar *apps* pertenecientes al ámbito de la *mHealth* porque mantienen el compromiso y la participación por medio de la gamificación. De hecho, la incorporación de elementos gamificados en estas *apps* no solo aumenta la motivación y la implicación, sino que también genera conductas saludables positivas (Jamaludin et al., 2021; Sousa et al., 2020).

Si bien hay una gran proliferación de *apps* de salud disponibles actualmente en el mercado, destacando entre ellas las dirigidas a niños, niñas y adolescentes, sumado a las oportunidades educativas que presenta la gamificación para la adquisición de hábitos saludables en estas edades, hay una falta de información considerable acerca de la calidad y la fiabilidad de este tipo de software, ya que son muy pocas las *apps* que están reguladas y respaldadas por evidencia científica con respecto a sus contenidos, características y funcionalidades (Cheng et al., 2020; Martín-Fernández et al., 2020).

En consecuencia, la desinformación y la adopción de conductas negativas para la salud se han ido extendido en la infancia y la adolescencia al utilizar *apps* de baja calidad con información no verificada (Frontini et al., 2020); de ahí la difícil tarea de localizar aquellas *apps* de mayor calidad, seguras y con un diseño ajustado a la edad de los usuarios y las usuarias finales basado en conocimiento experto (Albrecht et al., 2018; Expósito-López et al., 2023). Además, hay una falta de estudios empíricos sobre la evaluación de la calidad de este tipo de *apps* destinadas a niños, niñas y adolescentes (Jeminiwa et al., 2019).

Por lo tanto, a fin de poder guiar y ayudar a los agentes educativos a seleccionar adecuadamente tecnología educativa de calidad, el objetivo principal de este estudio es diseñar y desarrollar un instrumento de medida dirigido a las familias sólido, fiable, con ítems que se ajusten y representen bien a los constructos, y con unas óptimas propiedades psicométricas para evaluar la calidad de las *apps* gamificadas orientadas a la promoción de hábitos de vida saludable en el alumnado de educación primaria y secundaria.

2. Método

2.1. Enfoque metodológico

Este estudio parte de un diseño de investigación de corte cuantitativo basado en los principios de un enfoque descriptivo y de correlación, donde se ha realizado una prueba piloto en forma de intervención, la cual tiene que ver con administrar el instrumento a sujetos con características similares a la muestra objetivo a efectos de ajustar ítems susceptibles de modificación (operatividad, redacción, etc.) y mejorar la propia apariencia del instrumento (Hernández-Sampieri et al., 2006).

2.2. Participantes

Se ha llevado a cabo un muestreo no probabilístico, por cuotas, de manera deliberada, por conveniencia y accesibilidad. Autores como McMillan y Schumacher (2005) exponen que el muestreo no probabilístico es la forma de muestreo más común en la investigación educativa, a través del cual se selecciona a un grupo de personas por la accesibilidad en lugar de recurrir a la aleatoriedad. Por lo tanto, la selección de sujetos se basa en características de interés para el estudio, quedando las cuotas definidas previamente para

garantizar la representatividad de la muestra en función de la población (Hernández-Sampieri et al., 2006; Hernández-Sampieri y Mendoza-Torres, 2018).

En este sentido, la muestra ha quedado constituida por un total de 33 sujetos (padres, madres, tutores legales, etc.) con hijos, hijas o menores a cargo escolarizados entre el 3.^{er} curso de educación primaria y el 2.^o curso de educación secundaria obligatoria (ESO) de siete centros educativos diferentes de Huelva capital y provincia.

Según McMillan y Schumacher (2005), dado que un estudio piloto con al menos diez individuos puede revelar necesidades de ajuste en un cuestionario, se ha conseguido un tamaño muestral adecuado.

En cuanto a las características sociodemográficas de las familias participantes, que han englobado factores como el género y la edad, han participado 27 mujeres (81,8%) y seis hombres (18,2%). El rango de edad ha abarcado desde los 31 hasta los 52 años, con una media que se ha situado en los 44,73 años.

2.3. Instrumento

Las dimensiones e ítems del cuestionario propuesto provienen de un estudio de tesis que incluyó una revisión exhaustiva de la literatura sobre instrumentos de evaluación de *apps* de salud e indicadores de gamificación en el diseño de *apps* destinadas al fomento de hábitos saludables en niños, niñas y adolescentes (Delgado-Morales y Duarte-Hueros, 2022; 2023), teniendo como objetivo principal desarrollar un marco de evaluación de la calidad de este tipo de software dirigido al alumnado de educación primaria y ESO.

Posteriormente, las dimensiones y los indicadores iniciales se han validado teóricamente a través de dos técnicas: (1) grupo de discusión con cinco personas expertas en métodos de investigación, tecnología educativa y promoción de la salud, y (2) entrevista semiestructurada con un sujeto experto en metodologías de investigación.

Una vez reagrupados los ítems en subdimensiones, se ha acudido a un juicio de 20 expertos y expertas en educación, tecnología educativa, salud, métodos de investigación, informática y desarrollo de software. Este número de sujetos participantes se ha determinado conforme a Rubio et al. (2003), quienes recomiendan seleccionar al menos tres personas por grupo de profesionales.

Tras calcular el índice de validez de contenido de un ítem (*I-CVI*), de la escala (*S-CVI*) y el estadístico Kappa de Cohen modificado (k^*), presentado por Polit et al. (2007), se han encontrado valores excelentes según Lynn (1986) y Polit et al. (2007), con una concordancia casi perfecta entre jueces.

Finalmente, se ha construido un cuestionario de 39 ítems de escala tipo Likert 1-6 puntos (1=Muy en desacuerdo; 6=Muy de acuerdo) y tres ítems de respuesta abierta. Los ítems están asociados a seis dimensiones: “Diseño gráfico y visual de la interfaz”, “Usabilidad de la *app*”, “Información”, “Protección de datos”, “Gamificación” y “Satisfacción”. Puede consultar la propuesta de cuestionario completa en el anexo.

2.4. Procedimiento

Con el propósito de solicitar la colaboración de las familias, se ha contactado vía telefónica con un grupo de docentes mentores (tutores y tutoras) de educación primaria y ESO que trabajan en centros educativos de Huelva capital y provincia.

Tras obtener el compromiso de participación de las familias, a través del correo electrónico se ha remitido a dicho grupo de docentes mentores el enlace al cuestionario

en línea, junto con los detalles de la prueba piloto, para su difusión entre padres, madres, tutores legales, etc.

La prueba piloto ha consistido en probar *apps* asignadas al azar y valorarlas con la versión inicial del cuestionario. Estas han sido “Healthy Jeart” y “Con la comida SI se juega”, dos *apps* dirigidas a niños, niñas y adolescentes de entre 8 y 16 años, gratuitas y monolingües (en idioma español), creadas para fomentar la adopción de conductas saludables de manera didáctica y lúdica.

Las familias han probado de manera independiente una de las anteriores *apps* durante un periodo de prueba de dos semanas, con la posibilidad de hacerlo junto a sus hijos, hijas o menores a cargo.

Este estudio ha contado con la aprobación del Comité de Ética de la Investigación Provincial de Huelva con código interno 0611-N-23.

2.5. Análisis de datos

Los datos recopilados han sido almacenados y procesados utilizando el programa de análisis estadístico *IBM SPSS Statistics* (versión 27.0.1.0). Asimismo, los datos han sido previamente organizados y depurados en el programa *Microsoft Excel* (versión 16.69.1). Se ha calculado para cada escala del cuestionario los estadísticos descriptivos de la muestra, examinando diferentes propiedades para observar si los datos tienen una distribución normal: medias, desviaciones típicas, asimetría y curtosis.

Además, con el objetivo de explorar la relación entre las variables de cada escala, se han analizado las correlaciones no paramétricas de los ítems por medio del coeficiente de correlación de Spearman dado que la muestra es inferior a 50 sujetos (Razali y Wah, 2011).

Posteriormente, se ha analizado la consistencia interna de las escalas calculando el coeficiente Alfa de Cronbach a fin de identificar ítems poco consistentes (se han excluido ítems con valores $\alpha < 0,3$).

Para determinar la estructura factorial del cuestionario y las asociaciones entre variables, se ha realizado un análisis factorial exploratorio (AFE) empleando la extracción de máxima verosimilitud y la rotación Varimax. Siguiendo a Nunnally y Bernstein (1995), se han obviado los valores inferiores a 0,30 y se han aceptado cargas factoriales superiores a 0,50.

En las escalas con más de un factor, se ha llevado a cabo un análisis adicional de la fiabilidad y las correlaciones ítem-total corregido. También se ha evaluado el ajuste del modelo mediante el análisis de ecuaciones estructurales, con el software *IBM SPSS Amos* (versión 18), según los criterios de Byrne (2010), Hu y Bentler (1999) y Kline (2010): CMIN/DF entre 2 y 5; CFI, IFI y NFI $> 0,90$; RMSEA $< 0,06$; y HOELTER > 200 .

3. Resultados

3.1. Dimensión 1: Escala “Diseño gráfico y visual de la interfaz”

Índices descriptivos

En la Tabla 1 se muestran los datos descriptivos y los índices de asimetría y curtosis de la escala “Diseño gráfico y visual de la interfaz”.

Tabla 1.

Estadísticos descriptivos e índices de asimetría y curtosis de la escala "Diseño gráfico y visual de la interfaz"

Variables	Mínimo	Máximo	Media	Desviación	Asimetría	Curtosis
V1a	3	6	5,30	,883	-1,236	,970
V1b	3	6	5,18	,950	-,851	-,337
V1c	2	6	5,30	,984	-1,715	3,143
V2	3	6	5,30	,810	-1,002	,519

En la primera escala de cuatro ítems, el valor máximo ha sido 6 y el mínimo, mayoritariamente, entre 2 y 3. La media ha sido superior a 5, indicando altas puntuaciones de los sujetos. Asimismo, las desviaciones típicas han mostrado menor dispersión y variabilidad, aunque se da un alto grado de incidencia de las variables en el diseño de una *app* de este tipo. La distribución de los valores ha presentado una asimetría negativa, con puntuaciones más extremas situadas por debajo de la media. La curtosis ha sido generalmente positiva, mostrando una distribución leptocúrtica. De acuerdo con Kline (2005), todos los valores se han tomado como aceptables.

Correlaciones ítem-total

A continuación, las correlaciones ítem-total han sido todas significativas en el nivel ,01 (bilateral) y en el nivel ,05 (bilateral) (Tabla 2); lo que indica que puede haber un factor latente que las origina.

Tabla 2.

Correlaciones ítem-total de la escala "Diseño gráfico y visual de la interfaz"

	V1a	V1b	V1c	V2
V1a	1			
V1b	,587**	1		
V1c	,420*	,699**	1	
V2	,601**	,618**	,660**	1

Notas:

** La correlación es significativa en el nivel ,01 (bilateral).

* La correlación es significativa en el nivel ,05 (bilateral).

Fiabilidad

En esta primera escala de "Diseño gráfico y visual de la interfaz", formada por cuatro ítems, se ha obtenido un índice de fiabilidad de ,855; lo que se traduce en una consistencia interna bastante satisfactoria. Además, los valores parciales de cada ítem han sido todos superiores a 0,6 con un alto grado de correlación (Tabla 3).

Tabla 3.

Índices de fiabilidad de la escala "Diseño gráfico y visual de la interfaz"

Alfa de Cronbach total de la escala ,855		
Variables	CIT-C	Alfa de Cronbach si el elemento se ha suprimido
V1a	,655	,832
V1b	,790	,773
V1c	,616	,853
V2	,748	,799

Nota: CIT-C: Correlación ítem-total corregida.

Dimensionalidad

Seguidamente, se ha realizado un AFE a fin de comprobar la estructura factorial de la escala mediante el método de extracción de máxima verosimilitud y el método de rotación Varimax. En la Tabla 4 se puede observar que se ha extraído un único factor que explica el 61,15% de la varianza, y cuyas variables han tenido todas cargas factoriales superiores a 0,60. Por estas razones, se intuye que los ítems reflejan bien el constructo.

Tabla 4.

Estructura factorial exploratoria de la escala "Diseño gráfico y visual de la interfaz"

KMO=,738; Bartlett, $\chi^2=60,508$; gl=6; P<,000 F1	
Variables	Carga del factor
V1b	,856
V2	,814
V1a	,757
V1c	,692
Varianza explicada (%)	61,152

Dado que se ha extraído un único factor, se ha comprobado la unidimensionalidad siguiendo el criterio de Kaiser (1960); es decir, el autovalor del primer componente tiene que ser >1. La regla se cumple y, por lo tanto, se trata de una escala unifactorial (Tabla 5).

Tabla 5.

Varianza total explicada de la escala "Diseño gráfico y visual de la interfaz"

Factor	Varianza total explicada					
	Autovalores iniciales		Sumas de cargas al cuadrado de la extracción			
	Total	% de varianza	% acumulado	Total	% de varianza	% acumulado
1	2,814	70,340	70,340	2,446	61,152	61,152
2	,639	15,978	86,318			
3	,315	7,874	94,192			
4	,232	5,808	100,000			

Nota: Método de extracción: máxima verosimilitud.

3.2. Dimensión 2: Escala “Usabilidad de la app”

Índices descriptivos

Tal como se expone en la Tabla 6, los estadísticos descriptivos y los índices de asimetría y curtosis en la segunda escala de siete ítems, como en la escala anterior, han sido aceptables. Se destaca que el valor mínimo ha sido más elevado. La curtosis ha sido mayoritariamente negativa, indicando una distribución platicúrtica.

Tabla 6.

Estadísticos descriptivos e índices de asimetría y curtosis de la escala "Usabilidad de la app"

Variables	Mínimo	Máximo	Media	Desviación	Asimetría	Curtosis
V3	4	6	5,42	,663	-,733	-,446
V4	4	6	5,42	,561	-,250	-,895
V5a	4	6	5,52	,667	-1,067	,050
V5b	4	6	5,21	,740	-,366	-1,039
V5c	4	6	5,30	,770	-,595	-1,033
V6	3	6	5,27	,839	-,907	,064
V7	4	6	5,36	,699	-,647	-,669

Correlaciones ítem-total

Si bien en esta segunda escala “Usabilidad de la app” predominan correlaciones significativas en el nivel ,01 (bilateral) y en el nivel ,05 (bilateral), hay ciertas variables que no han correlacionado bien con el resto y/o han tenido una baja puntuación (Tabla 7). Es el caso de las variables V3 “Pienso que podría resultar fácil para los usuarios y las usuarias aprender a usar las funciones de la app”; V4 “Pienso que podría resultar fácil para los usuarios y las usuarias aprender a navegar a través de la app (p.ej., la información se localiza allí donde se espera encontrar sin necesidad de hacer esfuerzos)”; y V7 “En la app aparecen todos los enlaces de pantalla necesarios para acceder a las ubicaciones deseadas (p.ej., volver al menú principal, salir del juego, etc.)”.

Tabla 7.

Correlaciones ítem-total de la escala "Usabilidad de la app"

	V3	V4	V5a	V5b	V5c	V6	V7
V3	1						
V4	,537**	1					
V5a	,246	,535**	1				
V5b	,352*	,593**	,770**	1			
V5c	,100	,448**	,757**	,800**	1		
V6	,099	,315	,603**	,586**	,656**	1	
V7	,328	,327	,347*	,391*	,301	,387*	1

Notas:

** La correlación es significativa en el nivel ,01 (bilateral).

* La correlación es significativa en el nivel ,05 (bilateral).

Fiabilidad

En esta segunda escala de “Usabilidad de la *app*”, integrada por siete ítems, se ha alcanzado un índice de fiabilidad de ,871. De este modo, la consistencia interna, al igual que en la escala anterior, ha sido bastante adecuada (Tabla 8). No obstante, se ha decidido suprimir finalmente la variable V3 antes mencionada, pues la literatura recomienda excluir los ítems con valores por debajo de 0,45.

Tabla 8.
Índices de fiabilidad de la escala "Usabilidad de la app"

Variables	CIT-C	Alfa de Cronbach total de la escala
		,871
V3	,366	,886
V4	,645	,856
V5a	,778	,836
V5b	,847	,824
V5c	,738	,840
V6	,680	,850
V7	,528	,868

Nota: CIT-C: Correlación ítem-total corregida.

Dimensionalidad

Tras un análisis iterativo, se han eliminado las variables V4 y V7 anteriormente descritas al tener cargas factoriales débiles <0,50 y/o saturar en varios factores. Estas dos variables, junto con la variable V3 excluida tras el análisis de la fiabilidad, no han correlacionado bien con el resto.

Por consiguiente, puede apreciarse en esta segunda escala que se ha extraído un único factor que explica el 73,38% de la varianza, teniendo las cuatro variables que lo componen cargas factoriales superiores a 0,70 (Tabla 9). En base a estos valores, se intuye que los ítems parecen medir el mismo constructo.

Tabla 9.
Estructura factorial exploratoria de la escala "Usabilidad de la app"

Variables	KMO=,846; Bartlett, $\chi^2=87,729$; gl=6; P<,000	
	F1	
V5c	,910	,852
V5b	,899	,837
V5a	,864	,810
V6	,743	,713
Alfa de Cronbach total	,910	
Varianza explicada (%)	73,385	

Nota: CIT-C: Correlación ítem-total corregida.

De nuevo, se ha analizado la unidimensionalidad al haberse extraído un único factor, confirmando que se trata de una escala unifactorial (Tabla 10).

Tabla 10.

Varianza total explicada de la escala "Usabilidad de la app"

Factor	Varianza total explicada					
	Autovalores iniciales			Sumas de cargas al cuadrado de la extracción		
	Total	% de varianza	% acumulado	Total	% de varianza	% acumulado
1	3,188	79,699	79,699	2,935	73,385	73,385
2	,405	10,134	89,834			
3	,225	5,621	95,454			
4	,182	4,546	100,000			

Nota: Método de extracción: máxima verosimilitud.

3.3. Dimensión 3: Escala "Información"

Índices descriptivos

En la Tabla 11 se observa que en esta tercera escala de 13 ítems, una vez más, los estadísticos descriptivos y los índices de asimetría y curtosis se han tomado como aceptables. A diferencia de las escalas anteriores, el valor mínimo se ha situado en 1 y 2 en la mayoría de las variables. La media, entre 4,97 y 5,52, ha revelado puntuaciones elevadas de los sujetos. Nuevamente, la curtosis ha sido positiva en su totalidad.

Tabla 11.

Estadísticos descriptivos e índices de asimetría y curtosis de la escala "Información"

Variables	Mínimo	Máximo	Media	Desviación	Asimetría	Curtosis
V8a	4	6	5,52	,712	-1,160	,034
V8b	2	6	5,45	,869	-2,283	6,835
V9a	2	6	5,12	,053	-1,280	1,308
V9b	1	6	4,97	1,159	-1,477	2,936
V9c	1	6	5,12	1,139	-1,872	4,372
V10a	2	6	5,45	,833	-2,441	8,395
V10b	1	6	5,42	1,001	-2,963	11,491
V11	3	6	5,42	,867	-1,589	2,018
V12a	1	6	5,15	1,278	-1,832	3,082
V12b	1	6	5,33	1,080	-2,470	7,567
V12c	1	6	5,18	1,236	-1,846	3,226
V13a	1	6	5	1,414	-1,764	2,616
V13b	2	6	5,27	1,069	-1,569	1,938

Correlaciones ítem-total

En cuanto a las correlaciones ítem-total, han sido todas significativas en el nivel ,01 (bilateral) y en el nivel ,05 (bilateral) (Tabla 12).

Tabla 12.
Correlaciones ítem-total de la escala "Información"

	V8a	V8b	V9a	V9b	V9c	V10a	V10b	V11	V12a	V12b	V12c	V13a	V13b
V8a	1												
V8b	,660**	1											
V9a	,486**	,609**	1										
V9b	,423*	,447**	,771**	1									
V9c	,417*	,549**	,536**	,686**	1								
V10a	,614**	,436*	,420*	,385*	,410*	1							
V10b	,676**	,667**	,556**	,562**	,549**	,762**	1						
V11	,595**	,435*	,420*	,453**	,474**	,470**	,679**	1					
V12a	,693**	,589**	,591**	,536**	,454**	,508**	,802**	,650**	1				
V12b	,520**	,454**	,442*	,485**	,471**	,445**	,739**	,635**	,780**	1			
V12c	,598**	,618**	,713**	,631**	,522**	,547**	,798**	,563**	,838**	,780**	1		
V13a	,488**	,545**	,621**	,631**	,532**	,608**	,825**	,613**	,786**	,729**	,764**	1	
V13b	,493**	,528**	,590**	,661**	,536**	,632**	,829**	,497**	,685**	,727**	,822**	,857**	1

Notas:

** La correlación es significativa en el nivel ,01 (bilateral).

* La correlación es significativa en el nivel ,05 (bilateral).

Fiabilidad

Como se expone en la Tabla 13, en esta tercera escala de “Información”, formada por 13 ítems, se ha conseguido un índice de fiabilidad de ,935; es decir, la consistencia interna ha sido excelente. Las variables menos consistentes han sido V8a “El contenido de la *app* está bien escrito y es correcto” y V10a “Teniendo en cuenta la edad de los usuarios y las usuarias, la explicación visual de los elementos de la *app* a través de gráficos, imágenes, vídeos, etc. es clara”, pero no se han excluido dado que los valores han sido próximos al valor recomendado.

Tabla 13.
Índices de fiabilidad de la escala "Información"

	Alfa de Cronbach total de la escala ,935	
Variables	CIT-C	Alfa de Cronbach si el elemento se ha suprimido
V8a	,490	,936
V8b	,764	,929
V9a	,764	,928
V9b	,637	,932
V9c	,558	,935
V10a	,450	,937
V10b	,840	,926
V11	,517	,935
V12a	,810	,926
V12b	,772	,928

V12c	,881	,923
V13a	,848	,925
V13b	,766	,928

Nota: CIT-C: Correlación ítem-total corregida.

Dimensionalidad

De la misma manera que la escala que precede, se han excluido variables con cargas factoriales débiles <0,50 y/o por saturar en varios factores después de llevar a cabo un proceso de análisis iterativo. Estas variables han sido V8a y V10a (anteriormente descritas); V9a “Para alcanzar los objetivos propuestos en la *app*, la información contenida es suficiente (p.ej., considere los enlaces a más información y recursos, no se incluyen datos sin importancia ni explicaciones excesivas, etc.)”; V9b “Para alcanzar los objetivos propuestos en la *app*, la información contenida es exhaustiva (p.ej., considere los enlaces a más información y recursos, no se incluyen datos sin importancia ni explicaciones excesivas, etc.)”; V9c “Para alcanzar los objetivos propuestos en la *app*, la información contenida es concisa (p.ej., considere los enlaces a más información y recursos, no se incluyen datos sin importancia ni explicaciones excesivas, etc.)”; V11 “La información de la *app* proviene de una fuente confiable (p.ej., considere si la información que se incluye es actual, se especifica claramente quién es el autor o la autora, etc.)”; V13a “Los objetivos específicos que se plantean en la *app* (detallados en la descripción o dentro de la propia *app*) para las actividades propuestas, los desafíos/retos, cada uno de los niveles, etc. son medibles”; y V13b “Los objetivos específicos que se plantean en la *app* (detallados en la descripción o dentro de la propia *app*) para las actividades propuestas, los desafíos/retos, cada uno de los niveles, etc. son alcanzables”.

Las variables V8a y V10a, a su vez, han arrojado bajos valores de consistencia interna en el análisis de la fiabilidad, anteriormente detallado.

Así pues, se observa en la Tabla 14 que, en esta tercera escala, la extracción final ha arrojado un único factor que explica el 80,99% de la varianza. Además, las cinco variables han tenido cargas factoriales por encima de 0,70 y se entiende, a priori, que reflejan adecuadamente el constructo.

Tabla 14.

Estructura factorial exploratoria de la escala "Información"

KMO=,831; Bartlett, $\chi^2=182,457$; gl=10; P<,000		
Variables	F1	CIT-C
V12c	,974	,935
V12a	,943	,893
V12b	,916	,901
V10b	,866	,881
V8b	,789	,791
Alfa de Cronbach total	,953	
Varianza explicada (%)	80,998	

Nota: CIT-C: Correlación ítem-total corregida.

Como en las demás escalas, se ha analizado la unidimensionalidad a partir del factor extraído, comprobando que esta escala es también unifactorial (Tabla 15).

Tabla 15.
Varianza total explicada de la escala "Información"

Factor	Varianza total explicada					
	Autovalores iniciales			Sumas de cargas al cuadrado de la extracción		
	Total	% de varianza	% acumulado	Total	% de varianza	% acumulado
1	4,261	85,226	85,226	4,050	80,998	80,998
2	,367	7,335	92,560			
3	,186	3,714	96,274			
4	,133	2,657	98,931			
5	,053	1,069	100,000			

Nota: Método de extracción: máxima verosimilitud.

3.4. Dimensión 4: Escala "Protección de datos"

Índices descriptivos

Como se muestra en la Tabla 16, los estadísticos descriptivos y los índices de asimetría y curtosis en esta cuarta escala de cinco ítems, con la misma tendencia que el resto de las escalas, han estado dentro de los valores aceptables. El valor mínimo, en esta ocasión, ha variado entre 1 y 3. La asimetría se ha mantenido negativa y la curtosis positiva.

Tabla 16.
Estadísticos descriptivos e índices de asimetría y curtosis de la escala "Protección de datos"

Variables	Mínimo	Máximo	Media	Desviación	Asimetría	Curtosis
V14	3	6	5,27	1,008	-1,177	,192
V15	3	6	5,36	,895	-1,371	1,165
V16	2	6	5,24	1,001	-1,519	2,367
V17	1	6	4,88	1,431	-1,409	1,446
V18	1	6	5,06	1,368	-1,519	1,510

Correlaciones ítem-total

Por su parte, las correlaciones han sido todas significativas en el nivel ,01 (bilateral) (Tabla 17).

Tabla 17.
Correlaciones ítem-total de la escala "Protección de datos"

	V14	V15	V16	V17	V18
V14	1				
V15	,773**	1			
V16	,469**	,631**	1		
V17	,521**	,500**	,733**	1	
V18	,567**	,623**	,734**	,846**	1

Nota: ** La correlación es significativa en el nivel ,01 (bilateral).

Fiabilidad

En esta cuarta escala de “Protección de datos”, constituida por cinco ítems, se ha arrojado un índice de fiabilidad de ,887; lo que quiere decir que esta escala posee un muy buen grado de consistencia interna. Algo que también se justifica con los valores parciales de cada ítem (Tabla 18).

Tabla 18.

Índices de fiabilidad de la escala "Protección de datos"

Alfa de Cronbach total de la escala ,887		
Variables	CIT-C	Alfa de Cronbach si el elemento se ha suprimido
V14	,599	,890
V15	,703	,873
V16	,842	,842
V17	,761	,861
V18	,821	,841

Nota: CIT-C: Correlación ítem-total corregida.

Dimensionalidad

Se ha optado, tras varios análisis, por eliminar finalmente la variable V14 “En la *app* se toman medidas para proteger a menores de acuerdo con las leyes actuales” al haber arrojado una carga factorial débil <0,50.

De este modo, tal como se expone en la Tabla 19, se puede apreciar en esta cuarta escala que se ha extraído un solo factor que explica el 68,27% de la varianza. De las cuatro variables que lo conforman, tres de ellas han tenido cargas factoriales >0,70. La única variable con una carga factorial inferior a 0,70 ha sido V15 “Los datos confidenciales de los usuarios y las usuarias están protegidos y anonimizados (se garantiza la privacidad de la información registrada en la *app*)”. Sin embargo, el valor es aceptable al ser >0,50.

Tabla 19.

Estructura factorial exploratoria de la escala "Protección de datos"

KMO=,695; Bartlett, $\chi^2=109,537$; gl=6; P<,000 F1		
Variables	Carga del factor	CIT-C
V18	,993	,887
V17	,919	,800
V16	,773	,835
V15	,550	,598
Alfa de Cronbach total	,890	
Varianza explicada (%)	68,274	

Nota: CIT-C: Correlación ítem-total corregida.

Como se ha extraído un único factor, se ha comprobado una vez más la unidimensionalidad, quedando confirmado que es una escala unifactorial (Tabla 20).

Tabla 20.

Varianza total explicada de la escala "Protección de datos"

Factor	Varianza total explicada					
	Autovalores iniciales			Sumas de cargas al cuadrado de la extracción		
	Total	% de varianza	% acumulado	Total	% de varianza	% acumulado
1	3,073	76,817	76,817	2,731	68,274	68,274
2	,694	17,357	94,173			
3	,157	3,931	98,104			
4	,076	1,896	100,000			

Nota: Método de extracción: máxima verosimilitud.

3.5. Dimensión 5: Escala “Gamificación”

Índices descriptivos

Se observa en la Tabla 21 que en esta quinta escala, formada por seis ítems, los estadísticos descriptivos y los índices de asimetría y curtosis, con valores muy similares sobre todo a la escala de “Información” y a la de “Protección de datos”, se han situado dentro de los valores aceptables.

Tabla 21.

Estadísticos descriptivos e índices de asimetría y curtosis de la escala "Gamificación"

Variables	Mínimo	Máximo	Media	Desviación	Asimetría	Curtosis
V19	1	6	5,06	1,413	-1,883	2,991
V20	1	6	5,30	1,212	-2,313	5,406
V21	2	6	5,21	1,111	-1,466	1,384
V22	2	6	5,39	,864	-2,128	6,376
V23	2	6	5,09	1,011	-1,349	1,838
V24	1	6	4,85	1,460	-1,387	1,313

Correlaciones ítem-total

Según se detalla en la Tabla 22, puede observarse en esta quinta escala que todas las correlaciones han sido significativas en el nivel ,01 (bilateral) y en el nivel ,05 (bilateral) a excepción de las correlaciones de la variable V22 “La *app* cuenta con un personaje principal o personajes principales. Indique en qué medida considera que este aspecto es importante y necesario para la *app*” con las variables V19 “La *app* tiene una narrativa claramente definida (p.ej., considere si los usuarios y las usuarias pueden tomar decisiones sobre aspectos de la historia en la que se basa la ambientación del juego, la posibilidad de explorar escenarios, elegir misiones, etc.)” y V21 “El contenido de la *app* está influenciado por los logros de los usuarios y las usuarias; esto es, se producen cambios en el contenido según el comportamiento y el progreso dentro de la *app* superando desafíos/retos, completando actividades, subiendo de nivel, etc.”, cuyas puntuaciones han sido más bajas.

Tabla 22.

Correlaciones ítem-total de la escala "Gamificación"

	V19	V20	V21	V22	V23	V24
V19	1					
V20	,475**	1				
V21	,706**	,796**	1			
V22	,094	,472**	,262	1		
V23	,383*	,406*	,546**	,375*	1	
V24	,522**	,535**	,641**	,397*	,726**	1

Notas:

** La correlación es significativa en el nivel ,01 (bilateral).

* La correlación es significativa en el nivel ,05 (bilateral).

Fiabilidad

En esta quinta escala de “Gamificación”, compuesta por seis ítems, se ha obtenido un índice de fiabilidad de ,835; lo que significa que la consistencia interna ha sido bastante satisfactoria. Cinco ítems han tenido valores parciales superiores a 0,6, mientras que la variable V22, antes mencionada, ha mostrado menor consistencia (Tabla 23). No obstante, el valor de Alfa de Cronbach no aumentaría de manera significativa si se eliminara dicha variable; de ahí que se haya decidido mantenerla para evaluar su comportamiento en el AFE, dada su pertinencia para el estudio.

Tabla 23.

Índices de fiabilidad de la escala "Gamificación"

	Alfa de Cronbach total de la escala ,835	
Variables	CIT-C	Alfa de Cronbach si el elemento se ha suprimido
V19	,627	,807
V20	,662	,797
V21	,772	,777
V22	,269	,860
V23	,632	,806
V24	,720	,785

Nota: CIT-C: Correlación ítem-total corregida.

Dimensionalidad

En esta quinta escala, la extracción ha arrojado dos factores que explican el 71,06% de la varianza (Tabla 24): un primer factor que reúne tres ítems sobre **elementos del juego y estrategias para aumentar el compromiso** (narrativa, actividades lúdicas destinadas a fomentar un cambio positivo de hábitos, progreso dentro de la *app*, etc.); y un segundo factor, también con tres ítems, referidos a **recursos de gamificación centrados en el personaje principal y los escenarios** (interacción, diálogos, personalización, etc.).

Tabla 24.

Estructura factorial exploratoria de la escala "Gamificación"

Variables	F1		F2	
	Carga del factor	CIT-C	Carga del factor	CIT-C
V21	,955	,886		
V20	,837	,781		
V19	,791	,756		
V23			,896	,792
V24			,756	,709
V22			,586	,509
Alfa de Cronbach total	,897		,789	
Varianza explicada (%)	40,937		30,127	

Estimación de parámetros y evaluación del ajuste

Para esta quinta escala de "Gamificación", se ha llevado a cabo un análisis factorial confirmatorio (AFC) con el objetivo de comprobar la adecuación de un modelo de dos factores.

Como se muestra en la Figura 1, la correlación entre los dos factores ha sido buena ($\beta=0,53$; $p<,001$). Además, los valores en los pesos de regresión y en las covarianzas han sido significativos.

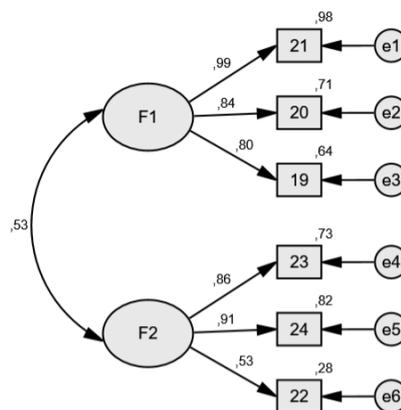


Figura 1. *Estimaciones del modelo para la escala "Gamificación"*

Como se detalla en la Tabla 25, los índices de ajuste han sido generalmente significativos excepto RMSEA ($,191>0,06$) y HOELTER ($29<200$), los cuales son sensibles al tamaño muestral. Sin embargo, los resultados no han sugerido una modificación de índices. Los valores de Chi-cuadrado ($X^2/g1=2,164$), CFI ($,921$), IFI ($,926$) y NFI ($,870$) han puesto de manifiesto que el ajuste de la escala es satisfactorio y el modelo se considera aceptable. Por todo ello, los tres ítems de cada factor hacen una aportación adecuada a la dimensión de procedencia al estar midiendo el mismo constructo.

Tabla 25.

Indicadores de bondad de ajuste del modelo. Escala "Gamificación"

X^2/gl	CFI	IFI	NFI	RMSEA	HOELTER
2,164	,921	,926	,870	,191	29

3.6. Dimensión 6: Escala "Satisfacción"

Índices descriptivos

Atendiendo a la Tabla 26, en esta sexta y última escala de cuatro ítems, los estadísticos descriptivos y los índices de asimetría y curtosis se han tomado nuevamente como aceptables, similares a los valores obtenidos en las escalas de "Información", "Protección de datos" y "Gamificación". En comparación con las demás escalas, el valor mínimo se ha situado en 1 en todas las variables.

Tabla 26.

Estadísticos descriptivos e índices de asimetría y curtosis de la escala "Satisfacción"

Variables	Mínimo	Máximo	Media	Desviación	Asimetría	Curtosis
V25	1	6	5,33	1,267	-2,252	4,605
V26	1	6	4,79	1,728	-1,393	,641
V27	1	6	5,06	1,391	-1,525	1,602
V28	1	6	4,82	1,467	-1,434	1,422

Correlaciones ítem-total

Una vez más, las correlaciones ítem-total han sido todas significativas en el nivel ,01 (bilateral) (Tabla 27).

Tabla 27.

Correlaciones ítem-total de la escala "Satisfacción"

	V25	V26	V27	V28
V25	1			
V26	,616**	1		
V27	,845**	,751**	1	
V28	,600**	,708**	,698**	1

Nota: ** La correlación es significativa en el nivel ,01 (bilateral).

Fiabilidad

En base a la información de la Tabla 28, en esta sexta escala de "Satisfacción", formada por cuatro ítems, se ha logrado un índice de fiabilidad de ,949 (excelente consistencia interna). Asimismo, esta ha sido la escala en la que se han obtenido mejores valores parciales de cada ítem, ya que todos han sido superiores a 0,8.

Tabla 28.
Índices de fiabilidad de la escala "Satisfacción"

Alfa de Cronbach total de la escala ,949		
Variables	CIT-C	Alfa de Cronbach si el elemento se ha suprimido
V25	,886	,935
V26	,845	,953
V27	,905	,926
V28	,917	,921

Nota: CIT-C: Correlación ítem-total corregida.

Dimensionalidad

Se puede observar en la Tabla 29 que en esta sexta escala se ha extraído un único factor que explica el 84,34% de la varianza. Las cuatro variables que lo componen, a su vez, han tenido cargas factoriales superiores a 0,80. Por lo tanto, se intuye que los ítems parecen medir el mismo constructo.

Tabla 29.
Estructura factorial exploratoria de la escala "Satisfacción"

KMO=,851; Bartlett, $\chi^2=134,682$; gl=6; P<,000		F1
Variables	Carga del factor	
V27	,946	
V28	,937	
V25	,926	
V26	,863	
Varianza explicada (%)	84,340	

Finalmente, dado que solo se ha extraído un factor, se ha comprobado la unidimensionalidad. De nuevo, se trata de una escala unifactorial (Tabla 30).

Tabla 30.
Varianza total explicada de la escala "Satisfacción"

Factor	Varianza total explicada					
	Autovalores iniciales		Sumas de cargas al cuadrado de la extracción			
	Total	% de varianza	% acumulado	Total	% de varianza	% acumulado
1	3,527	88,169	88,169	3,374	84,340	84,340
2	,249	6,231	94,400			
3	,121	3,037	97,436			
4	,103	2,564	100,000			

Nota: Método de extracción: máxima verosimilitud.

Obtenidos todos los valores, se ha vuelto a calcular el coeficiente Alfa de Cronbach de la escala global y de cada escala en particular. El nuevo valor arrojado ha sido de ,960 para la escala global y los valores en cada escala se han situado entre ,835 y ,953, revelando que el instrumento de medición posee una excelente consistencia interna y una alta confiabilidad.

4. Discusión

La falta de estudios empíricos relacionados con la evaluación de la calidad de este tipo de software (Jeminiwa et al., 2019) hace que esta investigación sea innovadora y relevante al tratarse del primer trabajo en el que se ha llevado a cabo, mediante una prueba piloto, una validación completa incluyendo un AFC.

Una vez depurados los datos, la propuesta de instrumento, denominado *Cuestionario MECAPPSalud-G dirigido a las Familias*, consta de seis escalas y 27 ítems que se adecuan y representan bien a los constructos: “Diseño gráfico y visual de la interfaz”, con cuatro ítems; “Usabilidad de la *app*”, con cuatro ítems; “Información”, con cinco ítems; “Protección de datos”, con cuatro ítems; “Gamificación”, con seis ítems; y “Satisfacción”, con cuatro ítems. Al respecto, se ha conseguido un gran equilibrio en el número de ítems de cada escala.

La fiabilidad de las subescalas del cuestionario ha sido comparable a la de otras escalas del ámbito de la *mHealth*. Se han observado coeficientes Alfa de Cronbach similares a los del instrumento *MARS* (Stoyanov et al., 2015) y a los de la herramienta *Enlight* (Baumel et al., 2017).

Estos resultados, a su vez, han sido similares a los reportados en los estudios de validación de la escala *MARS* en sus diversas traducciones, adaptaciones y versiones (Bardus et al., 2020; Böhme et al., 2019; Domnich et al., 2016; Martín-Payo et al., 2019; Messner et al., 2020; Saliassi et al., 2021; Stoyanov et al., 2016).

Seguidamente, el AFE de las escalas en común entre el cuestionario propuesto y el instrumento desarrollado por Stoyanov et al. (2015) ha puesto de manifiesto que la escala de “Diseño gráfico y visual de la interfaz” (equivalente a *Aesthetics*) ha sido unifactorial en ambos casos, aunque ha estado compuesta por tres ítems en el caso de *MARS*. Por su parte, si bien la escala de “Información” (comparable a *Information quality*) también ha sido unifactorial en los dos instrumentos, ha estado conformada por siete ítems en *MARS* (Terhorst et al., 2020).

En cambio, se desconoce hasta el momento si se ha evaluado la validez intraconstructos de la herramienta *Enlight* mediante AFE.

Por otro lado, los distintos análisis estadísticos efectuados han revelado robustez en la configuración de las dimensiones y las subdimensiones que subyacen al instrumento, con datos que se han distribuido con normalidad en todas las escalas. Los índices KMO han sido aceptables/buenos en el conjunto de dimensiones, con unos valores que han oscilado entre ,695 y ,851, lo que respalda la idoneidad del análisis factorial (González-Calatayud et al., 2022).

Además, la prueba de esfericidad de Bartlett ha arrojado índices de significación $<,001$ en todas las escalas. Por consiguiente, se puede rechazar la hipótesis nula de esfericidad y los análisis factoriales son totalmente aplicables para explicar los datos (Romero-Martínez y Ordóñez-Camacho, 2015).

En la misma línea, los índices derivados del AFC de la escala de “Gamificación” han sido aceptables en su mayor parte pese al tamaño de la muestra, que considerando los criterios planteados por Hernández-Sampieri et al. (2006), ha sido constituida por un buen número de familias.

5. Conclusiones

Con esta investigación se ofrece a las familias, como agentes educativos clave, un instrumento de medida de fácil aplicación y útil para poder utilizar en el hogar con sus hijos e hijas *apps* de calidad destinadas al fomento de hábitos de vida saludable a través de la integración de elementos gamificados. Por supuesto, el instrumento también resultará de interés tanto para la docencia como para la investigación, ya que ayudará a seleccionar la *app* de mayor calidad en relación con la adopción de hábitos de vida saludable en el alumnado.

La promoción de la salud en niños, niñas y adolescentes es muy necesaria hoy en día por parte de todos los ámbitos que les rodean y, en este sentido, educar en hábitos saludables desde el contexto familiar y educativo mediante este tipo de *apps* basadas en evidencia científica puede potenciar la adquisición y consolidación de conocimientos y hábitos beneficiosos para la salud en estas edades.

Por último, entre las limitaciones se destacan aquellas inherentes al instrumento, como la deseabilidad en las respuestas y la propia aplicación del cuestionario; se ha administrado en un único momento temporal y, por ende, no es posible evaluar la fiabilidad test-retest. Asimismo, se recomienda explorar sus propiedades psicométricas en un estudio a nivel nacional.

Presentación del artículo: 10 de octubre de 2024

Fecha de aprobación: 26 de diciembre de 2024

Fecha de publicación: 8 de enero de 2025

Delgado-Morales, C., Duarte-Hueros, A., Merino-Godoy, M. Á. y Conde-Vélez, S. (2025). Diseño y validación de un instrumento para evaluar la calidad de las *apps* gamificadas que promueven hábitos saludables en el alumnado de educación primaria y secundaria. *RED. Revista de Educación a Distancia*, 25(81). <http://dx.doi.org/10.6018/red.632871>

Declaración del autor o de los autores sobre el uso de LLM

Este artículo no ha utilizado para su redacción textos provenientes de un LLM (ChatGPT u otros).

Financiación

Este trabajo no ha recibido ninguna subvención específica de los organismos de financiación en los sectores públicos, comerciales o sin fines de lucro.

Referencias

- Albrecht, U., Hillebrand, U. y von Jan, U. (2018). Relevance of Trust Marks and CE Labels in German-Language Store Descriptions of Health Apps: Analysis. *JMIR Mhealth Uhealth*, 6(4), e10394. <https://doi.org/10.2196/10394>
- Alexandrou, C., Henriksson, H., Henström, M., Henriksson, P., Nyström, C. D., Bendtsen, M. y Löf, M. (2023). Effectiveness of a Smartphone App (MINISTOP 2.0) integrated in primary child health care to promote healthy diet and physical activity

- behaviors and prevent obesity in preschool-aged children: randomized controlled trial. *International Journal of Behavioral Nutrition and Physical Activity*, 20(22). <https://doi.org/10.1186/s12966-023-01405-5>
- Bardus, M., Awada, N., Ghandour, L. A., Fares, E., Gherbal, T., Al-Zanati, T. y Stoyanov, S. R. (2020). The Arabic Version of the Mobile App Rating Scale: Development and Validation Study. *JMIR Mhealth Uhealth*, 8(3), e16965. <https://doi.org/10.2196/16956>
- Baumel, A., Mathur, N., Kane, J. M. y Muench, F. (2017). Enlight: A Comprehensive Quality and Therapeutic Potential Evaluation Tool for Mobile and Web-Based eHealth Interventions. *Journal of Medical Internet Research*, 19(3), e82. <https://doi.org/10.2196/jmir.7270>
- Böhme, C., von Osthoff, M. B., Frey, K. y Hübner, J. (2019). Development of a Rating Tool for Mobile Cancer Apps: Information Analysis and Formal and Content-Related Evaluation of Selected Cancer Apps. *Journal of Cancer Education*, 34(1), 105-110. <https://doi.org/10.1007/s13187-017-1273-9>
- Brown, J. M., Franco-Arellano, B., Froome, H., Siddiqi, A., Mahmood, A. y Arcand, J. (2022). The Content, Quality, and Behavior Change Techniques in Nutrition-Themed Mobile Apps for Children in Canada: App Review and Evaluation Study. *JMIR Mhealth Uhealth*, 10(2), e31537. <https://doi.org/10.2196/31537>
- Brown, J. M., Savaglio, R., Watson, G., Kaplansky, A., LeSage, A., Hughes, J., Kapralos, B. y Arcand, J. (2020). Optimizing Child Nutrition Education With the Foodbot Factory Mobile Health App: Formative Evaluation and Analysis. *JMIR Formative Research*, 4(4), e15534. <https://doi.org/10.2196/15534>
- Byrne, B. M. (2010). *Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming*. (2nd ed.). Routledge.
- Campos, L. F., Cavalcante, J. P., Machado, D. P., Marçal, E., Silva, P. G. y Rolim, J. P. (2019). Development and evaluation of a mobile oral health application for preschoolers. *Telemedicine and e-Health*, 25(6), 492-498. <https://doi.org/10.1089/tmj.2018.0034>
- Chan, A., Kow, R. y Cheng, J. K. (2017). Adolescents' perceptions on smartphone applications (apps) for health management. *Journal of Mobile Technology in Medicine*, 6(2), 47-55. <http://dx.doi.org/10.7309/jmtm.6.2.6>
- Cheng, H., Tutt, A., Llewellyn, C., Size, D., Jones, J., Taki, S., Rossiter, C. y Denney-Wilson, E. (2020). Content and quality of infant feeding smartphone apps: five-year update on a systematic search and evaluation. *JMIR Mhealth Uhealth*, 8(5), e17300. <http://dx.doi.org/10.2196/17300>
- Delgado-Fernández, M. y Tercedor-Sánchez, P. (2021). *Estrategias de intervención en educación para la salud desde la Educación Física*. INDE.
- Delgado-Morales, C. y Duarte-Hueros, A. (16-18 Noviembre, 2022). *Hacia un marco de evaluación de la gamificación en Apps que promueven hábitos saludables en jóvenes [comunicación]*. XXV Congreso Internacional Educación y Tecnología EDUTECH 2022, Palma, España. <https://bit.ly/3PY8Yy1>

- Delgado-Morales, C. y Duarte-Hueros, A. (2023). Revisión sistemática de instrumentos que evalúan la calidad de aplicaciones móviles de salud. *Pixel-Bit. Revista De Medios Y Educación*, (67), 35-58. <https://doi.org/10.12795/pixelbit.97867>
- Domnich, A., Arata, L., Amicizia, D., Signori, A., Patrick, B., Stoyanov, S. R., Hides, L., Gasparini, R. y Panatto, D. (2016). Development and validation of the Italian version of the Mobile Application Rating Scale and its generalisability to apps targeting primary prevention. *BMC Medical Informatics and Decision Making*, 16(83). <https://doi.org/10.1186/s12911-016-0323-2>
- Expósito-López, J., Romero-Díaz de la Guardia, J. J. y Olmedo-Moreno, E. M. (2023). Diseño y validación de contenido de una rúbrica para medir el valor educativo de aplicaciones para dispositivos móviles. *Revista de Educación a Distancia*, 23(72). <https://doi.org/10.6018/red.54226>
- Frontini, R., Sousa, P., Dixe, M. A., Ferreira, R. y Figueiredo, M. C. (2020). Designing a mobile app to promote healthy behaviors and prevent obesity: analysis of adolescents' preferences. *Informatics for Health and Social Care*, 45(3), 327-341. <https://doi.org/10.1080/17538157.2020.1725766>
- Fuemmeler, B. F., Holzwarth, E., Sheng, Y., Do, E. K., Miller, C. A., Blatt, J., Rosoff, P. M. y Østbye, T. (2020). Mila blooms: a mobile phone application and behavioral intervention for promoting physical activity and a healthy diet among adolescent survivors of childhood cancer. *Games for Health Journal*, 9(4), 279-289. <https://doi.org/10.1089/g4h.2019.0060>
- González-Calatayud, V., Prendes-Espinosa, M. P. y Solano-Fernández, I. M. (2022). Instrumento de análisis de la competencia de emprendimiento digital en educación superior. *RELIEVE - Revista Electrónica De Investigación Y Evaluación Educativa*, 28(1). <https://doi.org/10.30827/relieve.v28i1.22831>
- Hernández-Sampieri, R. y Mendoza-Torres, C. P. (2018). *Metodología de la investigación: las rutas cuantitativa, cualitativa y mixta*. McGraw Hill Education.
- Hernández-Sampieri, R., Fernández-Collado, C. y Baptista-Lucio, P. (2006). *Metodología de la investigación*. (4.ª ed.). McGraw-Hill Interamericana.
- Hu, L. y Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternative. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55. <https://doi.org/10.1080/10705519909540118>
- Jamaludin, N. F., Tengku-Wook, T. S., Mat-Noor, S. F. y Qamar, F. (2021). Gamification Design Elements to Enhance Adolescent Motivation in Diagnosing Depression. *International Journal of Interactive Mobile Technologies*, 15(10), 154-172. <https://doi.org/10.3991/ijim.v15i10.21137>
- Jeminiwa, R. N., Hohmann, N. S. y Fox, B. I. (2019). Developing a Theoretical Framework for Evaluating the Quality of mHealth Apps for Adolescent Users: A Systematic Review. *The Journal of Pediatric Pharmacology and Therapeutics*, 24(4), 254-269. <https://doi.org/10.5863/1551-6776-24.4.254>
- Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20(1), 141-151. <https://doi.org/10.1177/001316446002000116>

- Kim, S., Park, Y. y Ackerman, M. K. (2021). Designing an Indoor Air Quality Monitoring App for Asthma Management in Children: User-Centered Design Approach. *JMIR Formative Research*, 5(9), e27447. <https://doi.org/10.2196/27447>
- Kline, R. B. (2005). *Principles and Practice of Structural Equation Modeling*. (2nd ed.). Guilford Press.
- Kline, R. B. (2010). *Principles and Practice of Structural Equation Modeling*. (3rd ed.). Guilford Press.
- Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research*, 35(6), 382-386.
- Martín-Fernández, A., Marco-Cuenca, G. y Salvador-Oliván, J. A. (2020). Evaluación y acreditación de las aplicaciones móviles relacionadas con la salud. *Revista Española de Salud Pública*, 94, e202008085.
- Martín-Payo, R., Fernández-Álvarez, M. M., Blanco-Díaz, M., Cuesta-Izquierdo, M., Stoyanov, S. R. y Llana-Suárez, E. (2019). Spanish adaptation and validation of the Mobile Application Rating Scale questionnaire. *International Journal of Medical Informatics*, 129, 95-99. <https://doi.org/10.1016/j.ijmedinf.2019.06.005>
- Martin, A., Caon, M., Adorni, F., Andreoni, G., Ascolese, A., Atkinson, S., Bul, K., Carrion, C., Castell, C., Ciociola, V., Condon, L., Espallargues, M., Hanley, J., Jesuthasan, N., Lafortuna, C. L., Lang, A., Prinelli, F., Puidomenech-Puig, E., Tabozzi, S. A. y McKinstry, B. (2020). A Mobile Phone Intervention to Improve Obesity-Related Health Behaviors of Adolescents Across Europe: Iterative Co-Design and Feasibility Study *JMIR Mhealth Uhealth*, 8(3), e14118. <https://doi.org/10.2196/14118>
- McMillan, J. H. y Schumacher, S. (2005). *Investigación educativa. Una introducción conceptual*. (5.ª ed.). Pearson Educación.
- Mehta, P., Moore, S. L., Bull, S. y Kwan, B. M. (2021). Building MedVenture – A mobile health application to improve adolescent medication adherence – Using a multidisciplinary approach and academic-industry collaboration. *Digital Health*, 7. <https://doi.org/10.1177%2F20552076211019877>
- Messner, E., Terhorst, Y., Barke, A., Baumeister, H., Stoyanov, S. R., Hides, L., Kavanagh, D. J., Pryss, R., Sander, L. y Probst, T. (2020). The German Version of the Mobile App Rating Scale (MARS-G): Development and Validation Study. *JMIR Mhealth Uhealth*, 8(3), e14479. <https://doi.org/10.2196/14479>
- Nunnally, J. y Bernstein, Y. (1995). *Teoría psicométrica*. McGrawHill.
- Organización Mundial de la Salud. (2004). 57.ª Asamblea Mundial de la Salud. <https://bit.ly/3AhCOWS>
- Organización Mundial de la Salud. (2011). *mHealth: new horizons for health through mobile technologies: second global survey on eHealth*. Global Observatory for eHealth. <https://bit.ly/3LV8iVu>
- Polit, D. F., Beck, C. T. y Owen, S. V. (2007). Is the CVI an Acceptable Indicator of Content Validity? Appraisal and Recommendations. *Research in Nursing & Health*, 30(4), 459-467. <https://doi.org/10.1002/nur.20199>

- Ramos-Gordillo, A. S. (2003). *Actividad Física e Higiene para la Salud*. Servicio de publicaciones y producción de la Universidad de Las Palmas de Gran Canaria.
- Razali, N. M. y Wah, Y. B. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics*, 2(1), 21-33.
- Romero-Martínez, S. J. y Ordóñez-Camacho, X. G. (2015). *Psicometría*. Centro de Estudios Financieros.
- Rubio, D. M., Berg-Weger, M., Tebb, S. S., Lee, E. S. y Rauch, S. (2003). Objectifying content validity: Conducting a content validity study in social work research. *Social Work Research*, 27(2), 94-104. <https://doi.org/10.1093/swr/27.2.94>
- Salas-Cabrera, J. (2015). Estilos de vida saludables: un derecho fundamental en la vida del ser humano. *Revista Latinoamericana de Derechos Humanos*, 26(2), 37-51. <https://doi.org/10.15359/rldh.26-2.2>
- Saliasi, I., Martinon, P., Darlington, E., Smentek, C., Tardivo, D., Bourgeois, D., Dussart, C., Carrouel, F. y Fraticelli, L. (2021). Promoting Health via mHealth Applications Using a French Version of the Mobile App Rating Scale: Adaptation and Validation Study. *JMIR Mhealth Uhealth*, 9(8), e30480. <https://doi.org/10.2196/30480>
- Sánchez-Álvarez, P. (Coord.). (2007). *Desarrollad hábitos de vida saludables*. Editorial Azarbe.
- Schmidt, M., Fisher, A. P., Sensenbaugh, J., Ling, B., Rietta, C., Babcock, L., Kurowski, B. G. y Wade, S. L. (2020). User experience (re) design and evaluation of a self-guided, mobile health app for adolescents with mild traumatic brain injury. *Journal of Formative Design in Learning*, 4(2), 51-64. <https://doi.org/10.1007/s41686-019-00038-x>
- Sousa, P., Martinho, R., Reis, C. I., Dias, S. S., Gaspar, P. J., Dixe, M. A., Luis, L. S. y Ferreira, R. (2020). Controlled trial of an mHealth intervention to promote healthy behaviours in adolescence (TeenPower): Effectiveness analysis. *Journal of Advanced Nursing*, 76(4), 1057-1068. <https://doi.org/10.1111/jan.14301>
- Stoyanov, S. R., Hides, L., Kavanagh, D. J. y Wilson, H. (2016). Development and Validation of the User Version of the Mobile Application Rating Scale (uMARS). *JMIR Mhealth Uhealth*, 4(2), e72. <https://doi.org/10.2196/mhealth.5849>
- Stoyanov, S. R., Hides, L., Kavanagh, D. J., Zelenko, O., Tjondronegoro, D. y Mani, M. (2015). Mobile App Rating Scale: A New Tool for Assessing the Quality of Health Mobile Apps. *JMIR Mhealth Uhealth*, 3(1), e27. <https://doi.org/10.2196/mhealth.3422>
- Terhorst, Y., Philippi, P., Sander, L. B., Schultchen, D., Paganini, S., Bardus, M., Santo, K., Knitza, J., Machado, G. C., Schoeppe, S., Bauereiß, N., Portenhauser, A., Domhardt, M., Walter, B., Krusche, M., Baumeister, H. y Messner, E. M. (2020). Validation of the Mobile Application Rating Scale (MARS). *PLoS ONE* 15(11), e0241480. <https://doi.org/10.1371/journal.pone.0241480>
- Thornton, L., Gardner, L. A., Osman, B., Green, O., Champion, K. E., Bryant, Z., Teesson, M., Kay-Lambkin, F. y Chapman, C. (2021). A Multiple Health Behavior Change, Self-Monitoring Mobile App for Adolescents: Development and Usability

Study of the Health4Life App. *JMIR Formative Research*, 5(4), e25513.
<https://doi.org/10.2196/25513>

Zahid, T., Alyafi, R., Bantanl, N., Alzahrani, R. y Elfirt, E. (2020). Comparison of Effectiveness of Mobile App versus Conventional Educational Lectures on Oral Hygiene Knowledge and Behavior of High School Students in Saudi Arabia. *Patient Preference and Adherence*, 14, 1901-1909. <https://doi.org/10.2147/PPA.S270215>

Zichermann, G. y Cunningham, C. (2011). *Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps*. O'Reilly.

Anexo. Propuesta de Cuestionario MECAPPSalud-G dirigido a las Familias

Cuestionario MECAPPSalud-G dirigido a las Familias

Valore, para cada una de las siguientes dimensiones, las afirmaciones/expresiones que se presentan a continuación. Posiciónese (marque con una X) del 1 al 6 en el grado de la escala, siendo 1-Muy en desacuerdo y 6-Muy de acuerdo.

D1. Diseño gráfico y visual de la interfaz (en cumplimiento del Real Decreto 1112/2018, de 7 de septiembre, sobre accesibilidad de los sitios web y aplicaciones para dispositivos móviles del sector público, así como de la norma europea EN 301 549 V3.2.1 (2021-03) y de la norma española UNE 71362:2020)

1 2 3 4 5 6

1) Se adecuan a la edad de los usuarios y las usuarias la disposición y el tamaño de:

Las fuentes (tipo de letra)

Los botones

Los menús

2) El diseño visual de la *app* es armonioso para los usuarios y las usuarias (p.ej., considere el contraste de colores proporcionado, la facilidad para poder leer el texto, etc.)

D2. Usabilidad de la *app* (en cumplimiento del Real Decreto 1112/2018, de 7 de septiembre, sobre accesibilidad de los sitios web y aplicaciones para dispositivos móviles del sector público, así como de la norma europea EN 301 549 V3.2.1 (2021-03) y de la norma española UNE 71362:2020)

1 2 3 4 5 6

3) Pienso que podría resultar fácil para los usuarios y las usuarias aprender a usar las funciones de la *app*

4) Pienso que podría resultar fácil para los usuarios y las usuarias aprender a navegar a través de la *app* (p.ej., la información se localiza allí donde se espera encontrar sin necesidad de hacer esfuerzos)

5) Moverse entre pantallas de la *app* se logra de manera:

Rápida y fluida (sin mucha demora en las acciones deseadas)

Lógica (sabiendo cómo acceder a cada parte de la *app*)

Precisa e ininterrumpida (sin problemas que me dificulten)

6) Las interacciones en la *app* son consistentes e intuitivas en todas las pantallas o los componentes (p.ej., considere las pulsaciones, los tecleos y los deslizamientos en los botones, los menús, etc.)

7) En la *app* aparecen todos los enlaces de pantalla necesarios para acceder a las ubicaciones deseadas (p.ej., volver al menú principal, salir del juego, etc.)

D3. Información

1 2 3 4 5 6

8) El contenido de la *app*:
 Está bien escrito y es correcto
 Es coherente con los objetivos y los temas que se tratan

9) Para alcanzar los objetivos propuestos en la *app*, la información contenida es (p.ej., considere los enlaces a más información y recursos, no se incluyen datos sin importancia ni explicaciones excesivas, etc.):

Suficiente
 Exhaustiva
 Concisa

10) Teniendo en cuenta la edad de los usuarios y las usuarias, la explicación visual de los elementos de la *app* a través de gráficos, imágenes, vídeos, etc., es:

Clara
 Adecuada

11) La información de la *app* proviene de una fuente confiable (p.ej., considere si la información que se incluye es actual, se especifica claramente quién es el autor o la autora, etc.)

12) La información que se ofrece acerca del propósito de la *app* es (p.ej., sobre los usuarios y las usuarias, el objetivo principal, los beneficios y las ventajas de usarla, las formas apropiadas de utilizarla, etc.):

Suficiente
 Clara
 Precisa

13) Los objetivos específicos que se plantean en la *app* (detallados en la descripción o dentro de la propia *app*) para las actividades propuestas, los desafíos/retos, cada uno de los niveles, etc. son:

Medibles
 Alcanzables

D4. Protección de datos (al amparo del Reglamento General de Protección de Datos y de la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales)

1 2 3 4 5 6

14) En la *app* se toman medidas para proteger a menores de acuerdo con las leyes actuales

15) Los datos confidenciales de los usuarios y las usuarias están protegidos y anonimizados (se garantiza la privacidad de la información registrada en la *app*)

- 16) Los servicios de la *app* utilizados en la nube cuentan con las medidas de seguridad necesarias (p.ej., los datos personales de los usuarios y las usuarias se mantendrán anónimos y protegidos en la nube de Google, de Amazon, etc., evitándose un acceso no autorizado y el robo de información)
- 17) La *app* detecta e identifica vulnerabilidades de ciberseguridad, posibles amenazas y riesgo de explotación/abuso (p.ej., una actualización no verificada)
- 18) La *app* aplica las medidas de seguridad adecuadas a las vulnerabilidades de la ciberseguridad ante posibles amenazas, con el fin de reducir el riesgo de explotación/abuso (p.ej., actualizaciones de seguridad frecuentes para evitar filtraciones de datos)

D5. Gamificación

1 2 3 4 5 6

- 19) La *app* tiene una narrativa claramente definida (p.ej., considere si los usuarios y las usuarias pueden tomar decisiones sobre aspectos de la historia en la que se basa la ambientación del juego, la posibilidad de explorar escenarios, elegir misiones, etc.)
- 20) La *app* tiene estímulos que motivan al cambio de hábitos beneficiosos para la salud; es decir, se proponen actividades lúdicas, importantes y realizables que animan a los usuarios y las usuarias a completarlas (p.ej., a caminar, a beber más agua, a establecer una rutina de sueño, etc.)
- 21) El contenido de la *app* está influenciado por los logros de los usuarios y las usuarias; esto es, se producen cambios en el contenido según el comportamiento y el progreso dentro de la *app* superando desafíos/retos, completando actividades, subiendo de nivel, etc.
- 22) La *app* cuenta con un personaje principal o personajes principales. Indique en qué medida considera que este aspecto es importante y necesario para la *app*
- 23) La *app* posibilita interactuar con el personaje principal o los personajes principales (p.ej., mediante su cuidado, diálogos, etc.). Indique en qué medida considera que este aspecto es importante y necesario para la *app*
- 24) La *app* permite la opción de personalización del personaje principal o los personajes principales, los decorados, etc. Indique en qué medida considera que este aspecto es importante y necesario para la *app*

D6. Satisfacción

1 2 3 4 5 6

- 25) Recomendaría la *app* a personas que pudieran obtener beneficio de ella

- 26) Seguiría usando la *app* en los próximos 12 meses si fuera necesario
- 27) Se han cumplido mis expectativas al usar la *app*
- 28) Considero que es una de las mejores *apps* que he usado sobre el fomento de hábitos saludables en la población infantil y adolescente a través de la gamificación

29) Explique los motivos por los que cree que esta *app* podría favorecer una modificación y/o adquisición de hábitos saludables en su hijo o hija:

30) Utilizaría la *app* en casa con mi hijo o hija para:

31) Observaciones finales sobre la *app*: