

Consistencia interevaluador en la evaluación de una app móvil para Educación Musical Básica: estudio piloto

Inter-evaluator consistency in the evaluation of a mobile app for Basic Music Education: pilot study.

Antonio León-Garrido^{a,1} , Julio Manuel Barroso-Osuna^{b,1} , Carmen Llorente-Cejudo^{c,1} 

¹ Universidad de Sevilla, Facultad de Ciencias de la Educación, Departamento de Didáctica y Organización Educativa. Sevilla, España

✉ ^a Autor de correspondencia aleon@us.es

✉ ^b jbarroso@us.es

✉ ^c karen@us.es

Recibido: 11/01/2025; Aceptado: 26/01/2025

Resumen

El avance de la tecnología en los contextos educativos, especialmente a través del uso de dispositivos y las aplicaciones móviles, ha transformado significativamente el panorama educativo. Por estos motivos, se planteó como objetivo evaluar el nivel de acuerdo en la entre evaluadores respecto a una aplicación móvil diseñada para el aprendizaje de conceptos básicos de la música. Para ello, se empleó una metodología basada en el Coeficiente de Correlación Intraclase (CCI) y el análisis ANOVA con Friedman como técnicas principales de análisis. El estudio utilizó un instrumento específico de evaluación de apps móviles de Educación Musical. La muestra estuvo compuesta por 60 docentes especializados en música, de los cuales 28 eran docentes universitarios y 32 eran maestros de Educación Primaria. Los resultados evidenciaron un alto nivel de acuerdo entre todos los evaluadores, lo que demuestra una fuerte consistencia y confiabilidad en las evaluaciones realizadas. Esto hallazgos respaldan tanto la solidez de las valoraciones emitidas sobre la aplicación móvil como la alta fiabilidad del instrumento empleado para su evaluación.

Palabras clave: evaluación de apps; App móviles; Educación Musical; confiabilidad; TIC.

Abstract

The advancement of technology in educational contexts, especially through the use of mobile devices and applications, has significantly transformed the educational landscape. For these reasons, the objective was to evaluate the level of agreement among the evaluators on a mobile application designed to learn the basic concepts of music. To this end, a methodology based on the Intraclass Correlation Coefficient (ICC) and ANOVA analysis with Friedman as the main analysis techniques was used. The study used a specific instrument to evaluate mobile music education apps. The sample was composed of 60 music teachers, of which 28 were university professors and 32 were primary school teachers. The results showed a high level of agreement among all the assessors, demonstrating strong consistency and reliability in the evaluations carried out. These findings support both the robustness of the evaluations issued on the mobile application and the high reliability of the instrument used for its evaluation.

Keywords: evaluation of apps; mobile apps; Music Education; reliability; ICT.

1.-INTRODUCCIÓN

El avance de la tecnología en los contextos educativos, popularmente mediante los dispositivos y las aplicaciones móviles (apps), ha revolucionado significativamente la educación (Romero Rodríguez et al., 2023; Gutiérrez-Castillo et al., 2024). La interacción con estas tecnologías proporciona una amplia variedad de servicios y funcionalidades accesibles directamente a través los teléfonos inteligentes, también conocidos como smartphones (Martínez-Baquero y Rodríguez-Umaña, 2022; Gutiérrez-Castillo, et al., 2023; Delgado-Morales y Duarte-Hueros, 2023).

Sin embargo, la proliferación masiva de apps móviles y su continuo desarrollo han planteado nuevos desafíos en las investigaciones relacionadas con la telefonía móvil. Estas investigaciones se enfocan en analizar como la calidad, la seguridad, la usabilidad y la eficacia de estos recursos tecnológicos (Gómez Redondo et al., 2020; Quezada-Bolaños y Gallardo-López, 2023). En este contexto, los desarrolladores de apps enfrentan dificultades al abordar variables relacionadas con la usabilidad en los ámbitos educativos. Estos desafíos incluyen cumplir con los estándares de calidad y cantidad de contenidos, así como la estética y técnica que debe presentar estos recursos (Gómez Redondo et al., 2020).

Para evaluar estos aspectos, es fundamental desarrollar análisis evaluativos de las tecnologías empleadas en la educación (Cabero-Almenara et al., 2017; Sdravopoulou et al., 2021). En especial, las apps móviles deben ser sometidas a evaluaciones detalladas para determinar si cumplen con los estándares de calidad necesarios para implementarlos en el ámbito educativo (Sdravopoulou et al., 2021; Quezada-Bolaños y Gallardo-López, 2023). Estas evaluaciones son esenciales para que los desarrolladores pueden diseñar nuevas apps ajustadas a las necesidades reales de los usuarios. Sin embargo, con frecuencia se comienza a adoptar otros patrones existentes en el mercado sin tomar en cuenta las opiniones de quienes las utilizan. Esto resulta en una desconexión de gran importancia entre la oferta tecnológica y las expectativas de los usuarios. De hecho, solo el 0.01% total de aplicaciones disponibles logra satisfacer a los usuarios, mientras que el 99.99% restante no alcanza este objetivo (Joachims, et al., 2017; Quezada-Bolaños y Gallardo-López, 2023).

La percepción de un usuario sobre una app móvil se forma en los primeros 50 milisegundos de interacción. Este breve tiempo es suficiente para que el usuario evalúe las características fundamentales. Por lo tanto, resulta imprescindible estudiar su eficacia, efectividad y satisfacción derivada de su uso (Joachims, et al., 2017). Este enfoque permitirá incrementar el porcentaje de aceptación y adopción de las apps móviles en los contextos educativos (Quezada-Bolaños y Gallardo-López, 2023).

Según Yu et al. (2022), Colebaugh et al. (2023) y Domínguez-Lloria et al. (2024) es fundamental evaluar las apps diseñadas para el aprendizaje de la música. Estas apps ofrecen herramientas para la enseñanza y la práctica musical, facilitando el acceso a los recursos educativos de forma interactiva. Evaluar la calidad de las apps garantiza a los usuarios la posibilidad de aprovechar plenamente las diversas funcionalidades que se integran en cada de una de ellas. Entre los aspectos recomendados para evaluar se incluyen la calidad del sonido, la calidad del contenido, la capacidad de personalización del aprendizaje, ente otros elementos. Asimismo, estos autores destacaron que el proceso de evaluación asegura experiencias educativas enriquecedoras y efectiva, ayudando a los estudiantes a desarrollar habilidades musicales y tecnológicas de forma más eficiente.

La evaluación de estos recursos garantiza a los usuarios una mayor seguridad, privacidad, flexibilidad y dinamización del proceso de enseñanza-aprendizaje (Gallardo-Montes et al,

2021; Rodríguez-Riesco y Senín-Calderón, 2022; Uribe-Guerrero y Acosta-Mesa, 2023). Esto resulta relevante considerando la diversidad de apps disponibles, diseñadas para abordar distintos temas y finalidades (Gomes et al., 2019). Además, cualquier herramienta tecnológica contribuye al desarrollo del aprendizaje al fomentar la autoeficacia y la motivación del alumnado, quien asume un rol protagónico en su proceso educativo. Estas herramientas también facilitan la resolución colaborativa de retos, permitiendo a los estudiantes a desarrollar habilidades y destrezas específicas en relación con el contenido (Romero-Tena y Romero-González, 2020; y Barragán et al., 2023).

La evaluación de un objeto de estudio se puede realizar a través de diversos métodos. Uno de ellos consiste en analizar la fiabilidad entre los codificadores para garantizar la consistencia, el nivel de acuerdo y confiabilidad de las codificaciones realizadas por los diferentes evaluadores. Estos niveles de acuerdo se obtienen mediante pruebas piloto, en las que se presentan la herramienta a evaluar junto con el instrumento de evaluación. A partir de estas pruebas, es posible calcular los indicadores estadísticos como el alfa de Cronbach, el Coeficiente de Correlación Interclase (CCI) y el análisis ANOVA con prueba de Friedman (Martínez-Pérez y Pérez-Martín, 2022; Calderón et al., 2023).

Por estos motivos, se planteó como objetivo de investigación: evaluar el nivel de consistencia y confiabilidad en la evaluación de una app móvil para el aprendizaje de contenidos básico de la música, utilizando un instrumento específico de evaluación de apps de Educación Musical. Para alcanzar este objetivo, se calcularon el CCI y la prueba ANOVA con Friedman. Estos análisis permitieron obtener una comprensión científica más profunda sobre la evaluación del objeto del estudio, identificando patrones positivos o discrepancias entre los codificadores respecto a la misma evaluación. Los hallazgos contribuirán a una implementación más efectiva de la tecnología en los contextos educativos de la educación musical, optimizando su uso y adaptabilidad.

2.-METODOLOGÍA

2.1-Diseño y participantes

La investigación se llevó a cabo mediante un estudio con diseño transversal, con el propósito de recoger y analizar los datos en ese momento específico. El tipo de muestreo empleado fue no probabilístico y de conveniencia, seleccionando a los participantes en función de su disponibilidad y accesibilidad. La metodología utilizada fue de carácter cuantitativo.

Para la selección de los evaluadores (muestra de los codificadores), se establecieron los siguientes criterios de inclusión: ser docentes especialistas en el área de Educación Musical, estar en activo en universidades y/o centros de Educación Primaria, contar con al menos dos años de experiencia laboral, y poseer formación académica específica en el área indicada.

El tamaño de la muestra fue de $N = 60$ docentes especializados en el área de música. Del ese total, 28 eran docentes universitarios (46.7%) y 32 maestros en centros de Educación primaria (53.3%). En cuanto a la distribución de género se observa que fue desigual: en el ámbito universitario, el 71.42% (20 docentes) eran mujeres y el 28.58% (8 docentes) hombres. En centros de Educación Primaria, el 62.5% (20 docentes) eran hombres y el 37.5% (12 docentes) eran mujeres. En conjunto, la muestra estuvo compuesta por un 46.7% (28) de hombres y un 53.3% (32) mujeres. La edad promedio de la muestra fue de 38.27 años, con una desviación estándar de 6.563 y un rango de edades entre los 27 y los 47 años.

Para la selección del objeto del estudio (muestra de app móviles), se establecieron los siguientes criterios:

1. Apps destinadas a Android e IOS.
2. Apps destinada al aprendizaje o refuerzo de contenidos básicos de Educación Musical (aprendizaje de lectura de notas, formación rítmica o formación auditiva).
3. Apps gratuita.
4. Apps con una puntuación mínima de 3.5 estrellas en las evaluaciones de los usuarios
5. Preferentemente en español
6. Y con actualizaciones periódicas

El tamaño de la muestra fue de $N= 40$ apps móviles que cumplieran con los criterios establecidos y abarcaban al menor uno de los contenidos mencionados. De estas, se seleccionó de forma aleatoria y probabilística una app para la investigación: «Clefs», una app móvil que reúne estas características especificadas y enfocada al aprendizaje de la lectura de notas musicales.

2.2.-Instrumento de recogida de datos

Con el propósito conocer el índice de acuerdo entre los evaluadores respecto a una misma app móvil para el aprendizaje básico de la música, se utilizó el instrumento de evaluación diseñado por León-Garrido et al. (2024). Este instrumento de evaluación consta de 70 ítems estructurado:

1. Identificación (39 ítems): orientadas a recolectar información sobre las características de la muestra.
2. Dimensión técnica-estética (10 ítems): enfocada en evaluar la calidad grafica y el diseño visual de la app.
3. Dimensión pedagógica-funcional (12 ítems): centrada en los aspectos clave del proceso de enseñanza-aprendizaje.
4. Dimensión musical (9 ítems): diseñada para valorar la calidad, cantidad e integración de actividades orientadas al desarrollo de habilidades musicales básica.

Este instrumento fue validado mediante la metodología del juicio de expertos, en el cual, participaron 38 especialistas, obteniendo una calificación positiva que respalda su uso. Aunque, la herramienta de evaluación proporciona principalmente datos cuantitativos, la sección de identificación incluye algunos elementos cualitativos que permiten caracterizar la aplicación evaluada, como su nombre, idioma y año de creación.

El uso de este instrumento permitió determinar si todos los evaluadores ofrecían una valoración consistente y sólida para medir la calidad de una app móvil, basada en las percepciones de un docente tras utilizarla como recurso para el aprendizaje de la música.

2.3.-Procedimiento

En primer lugar, se estableció contacto con los docentes de las instituciones educativas: universidades públicas, colegios públicos y privados de Educación Primaria. Durante este proceso, se explicó detalladamente el propósito de la investigación y se invitó a 60 docentes a participar en una sesión teórica-práctica sobre el uso de las apps móviles en el contexto de Educación Musical. La sesión teórica-práctica fue programada de manera individual para cada docente participante. En ella, se abordaron temas relacionados con la importancia del uso de las apps móviles en la educación, los beneficios que ofrecen, y las oportunidades que generan para mejorar los procesos de enseñanza-aprendizaje y la evaluación de la Educación Musical.

Asimismo, se destacó la relevancia y la necesidad de evaluar apps móviles antes de ser integradas en el aula. Esta evaluación es esencial para garantizar la calidad y la adecuación de estas herramientas al contexto educativo, asegurando su eficacia y pertinencia en el proceso de enseñanza-aprendizaje.

Posteriormente, se presentó y detalló el instrumento de evaluación que los docentes utilizarían, junto con la app móvil seleccionada para la investigación: Clefs. Esta app, disponible para sistemas Android e IOS, está diseñada específicamente para el aprendizaje de la lectura de notas musicales. Se instruyó a los participantes sobre cómo interactuar con la app y evaluar su calidad de acuerdo con sus percepciones y criterios de exigencia, considerando su potencial uso en el aula.

Cada docente tuvo la oportunidad de interactuar con la app de forma individualizada. Después, procedieron a realizar la evaluación utilizando el instrumento facilitado, diseñado en formato cuestionario online. En la sección de identificación, se solicitó a los participantes que indicaran si la app presentaba una característica específica de forma explícita o no. Para ello, se asignó un valor de 1 a las características presentes y 0 a aquellas ausentes según la evaluación del docente. En cuanto a las dimensiones técnicas, pedagógicas y musicales, se utilizó una escala Likert del 1 al 5, donde 1 representaba un grado de satisfacción totalmente desfavorable y 5 un grado de satisfacción altamente favorable.

Una vez completadas las evaluaciones por parte de los docentes, los datos fueron recopilados y almacenados en el programa estadístico profesional SPSS v29. Posteriormente, se realizaron análisis estadísticos para identificar patrones y tendencias en las evaluaciones del recurso tecnológico. Este análisis incluyó: cálculo del alfa de Cronbach, para determinar la consistencia interna del instrumento; CCI, aplicado mediante el modelo de dos factores del tipo consistencia, para medir el nivel de acuerdo entre los evaluadores; prueba de Chi-cuadrado de Friedman, para analizar diferencias significativas en las evaluaciones; y, por último, análisis descriptivo, para resumir y describir las principales tendencias observadas en los datos recopilados.

El CCI es uno de los métodos más empleados para analizar el índice de acuerdo entre los evaluadores (Correa-Rojas, 2021). Este permite medir de forma general la concordancia entre varias mediciones y de carácter cuantitativo, facilitando la evaluación de si las respuestas de los usuarios presentan variabilidad en su dispersión. El CCI genera valores que oscilan entre 0 y 1, donde: un valor cercano a 0 indica la ausencia de concordancia y un valor cercano a 1 refleja una concordancia y fiabilidad total. Para interpretar el índice de concordancia, se utilizan los siguientes criterios: <0.40 concordancia pobre; entre 0.40 y 0.59 concordancia suficiente; entre 0.60 y 0.74 concordancia buena; y, de 0.75 a 1 concordancia excelente (Martínez-Pérez y Pérez-Martín, 2022). Este análisis es muy útil en estudios donde es fundamental evaluar la consistencia de las evaluaciones por diferentes codificadores para obtener una medida precisa del nivel de acuerdo y fiabilidad.

Asimismo, en la prueba de Chi-cuadrado de Friedman se aplicó el coeficiente de concordancia e Kendall, el cual mide el nivel de acuerdo entre varios evaluadores sobre un mismo objeto evaluado. Este método es relevante porque permite obtener un acuerdo consensuado a partir de las respuestas recolectadas. El coeficiente de concordancia de Kendall oscila entre 0 y 1, donde: 1 indica un acuerdo total entre los evaluadores y 0 refleja un desacuerdo total. Para que el nivel de confiabilidad sea aceptable, el coeficiente debe cumplir con la condición $W \geq 0.5$. En caso contrario, será necesario revisar los procedimientos de la investigación para identificar posibles inconsistencias o ajustes requeridos (Del Pozo Franco et al., 2021).

3.-RESULTADOS

Los resultados almacenados fueron analizados mediante diversas aproximaciones. En primer lugar, se calculó el índice de fiabilidad alfa de Cronbach para evaluar la consistencia interna del instrumento, así como el CCI para estudiar el nivel de acuerdo de las evaluaciones realizadas los docentes en cada sección del instrumento y de forma general. Posteriormente, se realizó un análisis descriptivo de la app móvil seleccionada.

3.1-Fiabilidad de la escala: alfa de Cronbach

La escala de fiabilidad (Tabla 1) se aplicó a todos los elementos del instrumento de evaluación, exceptuando a los descriptivos: nombre de la app, idioma y año de publicación o actualización, resultando en un análisis de 67 ítems. En cuanto a la función de la app y los destinatarios, la fiabilidad obtenida es moderada, lo que refleja una consistencia interna aceptable entre los elementos de estas categorías. Para la temática con la que se puede trabajar la app, el coeficiente alfa fue superior, indicando una fiabilidad moderadamente buena. En contraste, la dimensión correspondiente a la rueda de la pedagogía de Carrington obtuvo una fiabilidad moderada-baja. No obstante, en las dimensiones técnica-estética, pedagógica-funcional y musical presentaron una fiabilidad muy alta, con coeficientes superiores a 0.9, lo que las posiciona como altamente consistentes. Al calcular el alfa de Cronbach para el instrumento completo, se obtuvo un coeficiente de 0.959, calificándolo como excelente. Este resultado indica que todos los elementos del instrumento están altamente correlacionados, lo que asegura que es una herramienta muy confiable para evaluar apps móviles en el contexto de la Educación Musical.

Tabla 1. Estadística de fiabilidad

	Alfa de Cronbach	N de elementos
Función de la app	,527	8
Versión	1,000	3
Requisito informático	1,000	3
Destinatarios	,601	8
Temática	,668	8
Rueda de la Pedagogía de Carrington	,496	6
Dimensión técnica-estética	,888	10
Dimensión pedagógica-funcional	,947	12
Dimensión musical	,913	9

	Alfa de Cronbach	N de elementos
Instrumento al completo	,959	67

3.2.- Análisis CCI y Chi-cuadrado de Friedman

Para evaluar el índice de acuerdo entre los evaluadores sobre un mismo objeto de estudio, se aplicó el CCI (Tabla 2). Este procedimiento se realizó utilizando los mismos ítems analizados previamente en la escala de fiabilidad.

Tabla 2. Coeficiente de Correlación Intraclase (CCI)

		Correlación intraclase ^b	Intervalo de confianza al 95%		Prueba F con valor verdadero 0			
			Límite inferior	Límite superior	Valor	gl1	gl2	Sig.
Función de la app	Medidas únicas	,143 ^a	,067	,247	2,167	59	354	<,001
	Medidas promedio	,538 ^c	,335	,697	2,167	59	354	<,001
Destinatarios	Medidas únicas	,185 ^a	,102	,296	2,586	59	354	<,001
	Medidas promedio	,613 ^c	,443	,746	2,586	59	354	<,001
Temática	Medidas únicas	,281 ^a	,179	,407	3,344	59	295	<,001
	Medidas promedio	,701 ^c	,566	,805	3,344	59	295	<,001
Rueda de la Pedagogía de Carrington	Medidas únicas	,141 ^a	,058	,253	1,985	59	295	<,001
	Medidas promedio	,496 ^c	,269	,671	1,985	59	295	<,001
Dimensión técnica-estética	Medidas únicas	,443 ^a	,345	,556	8,942	59	531	<,001
	Medidas promedio	,888 ^c	,841	,926	8,942	59	531	<,001
Dimensión pedagógica-funcional	Medidas únicas	,599 ^a	,507	,697	18,947	59	649	<,001
	Medidas promedio	,947 ^c	,925	,965	18,947	59	649	<,001
Dimensión musical	Medidas únicas	,540 ^a	,441	,647	11,555	59	472	<,001
	Medidas promedio	,913 ^c	,876	,943	11,555	59	472	<,001
Instrumento al completo	Medidas únicas	,296 ^a	,229	,389	25,836	59	3422	<,001

	Correlación intraclase ^b	Intervalo de confianza al 95%		Prueba F con valor verdadero 0			
		Límite inferior	Límite superior	Valor	gl1	gl2	Sig.
Medidas promedio	,961 ^c	,946	,974	25,836	59	3422	<,001

Modelo de efectos aleatorio de dos factores donde tanto los efectos de personas como los efectos de medidas son aleatorios. a. El estimador es el mismo, esté presente o no el efecto de interacción. b. Coeficientes de correlaciones entre clases del tipo C que utilizan una definición de coherencia. La varianza de medida intermedia se excluye de la varianza del denominador. c. Esta estimación se calcula suponiendo que el efecto de la interacción está ausente, porque de lo contrario no se puede estimar.

Las medidas únicas reflejan el acuerdo de una sola medición. En cuanto al CCI para estas medidas, los valores obtenidos son generalmente bajos, lo que indica una menor fiabilidad para una única validación. Sin embargo, las medias promedias muestran una mayor constancia cuando se combinan las variables. De hecho, se observa que los valores de las medidas promedio son más altos que los de las medidas únicas, lo que sugiere una mayor fiabilidad en la medición.

Los valores CCI de las medidas promedio (en comparación con las medidas únicas) sugieren que hay un acuerdo considerable entre todos los evaluadores. Los resultados evidenciaron que: la función de la app ha alcanzado el valor promedio de 0.538 y la rueda de pedagogía de Carrington 0.496, lo que indica un nivel suficiente. En cuanto a los destinatarios (0.613) y las temáticas (0.701), sugieren un nivel de acuerdo bueno. En lo que respecta a las dimensiones técnica-estética (0.888), pedagógica-funcional (0.947) y musical (0.913) muestran un nivel de acuerdo excelente. Además, en el instrumento de evaluación en su totalidad se obtuvo un CCI de 0.959, sugiriendo un nivel excelente de acuerdo.

Otro aspecto relevante es la prueba F y la significación ($p < 0.001$), ya que todas las correlaciones interclases son estadísticamente significativas, lo que respalda la fiabilidad de las mediciones realizadas.

Como se ha observado, la evaluación realizada por los codificadores tiende a coincidir más en sus evaluaciones cuando se agrupan y se calculan los promedios, lo que incrementa la fiabilidad consistente de los datos. Esto es crucial para determinar tanto la validez como la fiabilidad de la evaluación.

En la [Tabla 3](#) se presentan los resultados de un análisis de varianza (ANOVA) con la prueba de Friedman, utilizado para evaluar las posibles diferencias entre las medias de las medidas repetidas no paramétricas.

Tabla 3. ANOVA con prueba Friedman

ANOVA con prueba de Friedman						
		Suma de cuadrados	gl	Media cuadrática	Chi-cuadrado deFriedman	Sig.
Inter sujetos		749,966	59	12,711		
	Entre elementos	9720,705 ^a	58	167,598	2966,244	<,001
Intra sujetos	Residuo	1683,634	3422	,492		
	Total	11404,339	3480	3,277		
Total		12154,305	3539	3,434		

ANOVA con prueba de Friedman					
	Suma de cuadrados	gl	Media cuadrática	Chi-cuadrado deFriedman	Sig.
Media global = 2,17					
a. Coeficiente de concordancia de W = ,800.					

La suma de los cuadrados inter sujetos, con una media de 12.711, indica la variabilidad existente entre los codificadores. Por otro lado, la suma de los cuadrados de los intra sujetos, con una media de 167.598 y un chi cuadrado de Friedman de 2966.244 ($p < 0.0001$), refleja la variabilidad existente dentro de los sujetos en diferentes evaluaciones. Esto demuestra que, aunque existe un alto índice de acuerdo entre los evaluadores, no todos calificaron de manera uniforme todas las variables integradas, ya que observó variabilidad significativa, consistentes y no aleatorias.

Finalmente, el coeficiente de concordancia de $W = 0.800$ refuerza el alto grado de acuerdo entre los evaluadores, lo que respalda aún más los hallazgos obtenidos en el CCI, previamente mencionados. En otras palabras, ambas pruebas indican una alta fiabilidad del índice de acuerdo entre las evaluaciones realizada, especialmente cuando se consideran las medias promedios.

3.3.-Análisis descriptivos

A continuación, en la **Tabla 4** se presentan los datos descriptivos obtenidos de las evaluaciones realizadas sobre la app.

Tabla 4. Análisis descriptivos

	Mínimo Estadístico	Máximo Estadístico	Media Estadístico	Desv. estándar Estadístico	Asimetría Estadístico	Curtosis Estadístico
Función Informativa	0	1	,73	,446	-1,083	-,858
Función Evaluativa	0	1	,80	,403	-1,539	,379
Función Reproductora	0	1	,53	,503	-,137	-2,051
Función Exploradora	0	1	,13	,343	2,213	2,996
Función Guía	0	1	,53	,503	-,137	-2,051
Función Lúdica	0	1	,33	,475	,725	-1,526
Función Expresiva o Creativa	0	1	,13	,343	2,213	2,996
Función Ejercitación	1	1	1,00	,000	.	.
Versión Gratuita	0	1	1,00	,000	.	.
Versión Prueba	0	1	,00	,000	.	.
Versión Pago	0	0	,00	,000	.	.
Android	0	0	,00	,000	.	.
IOS	0	0	,00	,000	.	.

	Mínimo Estadístico	Máximo Estadístico	Media Estadístico	Desv. estándar Estadístico	Asimetría Estadístico	Curtosis Estadístico
Ambas	1	1	1,00	,000	.	.
Destinatario Educación Infantil	0	1	,13	,343	2,213	2,996
Destinatario Educación Primaria	0	1	,87	,343	-2,213	2,996
Destinatario Educación Secundaria	0	1	,93	,252	-3,564	11,071
Destinatario Bachillerato	0	1	,73	,446	-1,083	-,858
Destinatario Universidad	0	1	,67	,475	-,725	-1,526
Destinatario Conservatorios	0	1	,73	,446	-1,083	-,858
Destinatario Escuelas de Música	0	1	,80	,403	-1,539	,379
Destinatario Otros	0	0	,00	,000	.	.
Temática Edición Partituras	0	1	,13	,343	2,213	2,996
Temática Lenguaje y teoría Musical	1	1	1,00	,000	.	.
Temática Formación Instrumental	0	1	,13	,343	2,213	2,996
Temática Formación Vocal Auditiva	0	1	,13	,343	2,213	2,996
Temática Editor De Sonidos	0	0	,00	,000	.	.
Temática Formación Rítmica	0	1	,13	,343	2,213	2,996
Temática Creatividad Composición e Improvisación	0	1	,07	,252	3,564	11,071
Temática Cultura Musical	0	1	,33	,475	,725	-1,526
Carrington Recuerdo	0	1	,80	,403	-1,539	,379
Carrington Comprensión	0	1	,73	,446	-1,083	-,858
Carrington Aplicación	0	1	,73	,446	-1,083	-,858
Carrington Análisis	0	1	,33	,475	,725	-1,526
Carrington Evaluación	0	1	,47	,503	,137	-2,051
Carrington Creación	0	1	,07	,252	3,564	11,071
Dimensión técnica-estética						
Claridad de la Aplicación	3	5	4,07	,778	-,118	-1,323
Diseño de pantallas	2	5	3,87	,892	-,323	-,668
Calidad de sonido	2	5	3,93	1,071	-,548	-,996
Legibilidad de los contenidos	2	5	4,20	,917	-,961	,075
Estructura	3	5	4,00	,902	,000	-1,798
Navegación	3	5	4,13	,812	-,252	-1,438
Facilidad de Interacción	3	5	4,20	,840	-,398	-1,473
Integración de los medios	2	5	3,60	,887	-,314	-,551

	Mínimo Estadístico	Máximo Estadístico	Media Estadístico	Desv. estándar Estadístico	Asimetría Estadístico	Curtoisía Estadístico
Tiempo de instalación descarga	2	5	4,33	1,084	-1,370	,344
Actualizaciones	2	5	3,67	,951	,236	-1,158
Dimensión pedagógica-funcional						
Guía Didáctica	2	5	3,67	1,020	-,471	-,859
Capacidad de Motivación	1	5	3,67	1,145	-,706	-,023
Adaptación a distintitos ritmos	2	5	3,73	1,071	-,466	-,996
Niveles de dificultad	2	5	4,13	1,033	-1,040	-,057
Actividades en relación con los contenidos	1	5	3,67	1,361	-,531	-1,124
Fomento del aprendizaje y los conceptos	2	5	3,80	1,054	-,301	-1,155
Actividades previas	1	5	3,33	1,084	-,379	-,308
Evaluación del aprendizaje	1	5	3,67	1,145	-,706	-,023
Evaluación y grabación de los resultados	2	5	4,07	,936	-,648	-,554
Ayuda a cumplir los objetivos de aprendizajes, metas y resultados.	2	5	3,80	1,117	-,193	-1,447
Incita al trabajo cooperativo para alcanzar el objetivo de forma compartida	1	5	2,07	1,401	1,103	-,024
Presenta desafíos para resolver la actividad	2	5	3,40	1,153	-,022	-1,477
Dimensión Musical						
Calidad de las interpretaciones de los archivos multimedia	2	5	3,93	1,133	-,733	-,861
Adecuación del lenguaje y del contenido musical al nivel educativo	2	5	3,53	1,214	,037	-1,580
Secuenciación de los contenidos musicales	2	5	3,87	1,214	-,442	-1,442
Calidad de los contenidos musicales	1	5	3,33	1,311	-,090	-1,282
Calidad de los elementos visuales	1	5	3,13	1,371	,241	-1,454
Cantidad de los contenidos musicales	2	5	3,47	1,096	,247	-1,267
Integración de elementos teóricos musicales	1	5	3,40	1,153	-,297	-,559
Integración de elementos prácticos musicales	1	5	3,53	1,096	-,886	,063
Facilita el aprendizaje de la Educación Musical	2	5	3,67	1,145	-,145	-1,421

Como se observa en la [Tabla 4](#), la función predominante con la app móvil es la ejercitación (M= 1), seguida de la evaluativa (M=0.80) e informativa (M=0.73). A través de la app, se proporciona al alumno información sobre el contenido, se ejercitan las tareas a realizar y se lleva a cabo la evaluación. No obstante, también se presentan las demás funciones en menor medida. En cuanto a la versión utilizada, los usuarios la visualizan como una versión de gratuita (M=1) por unanimidad. En referencia a los destinatarios, se observa que la app pued

ser aplicada en diversos contextos educativos, predominando su utilización en Secundaria con una media de 0.93, seguido de Primaria con 0.87 y en Escuelas de Música con 0.80.

La temática general de la app es lenguaje y teoría musical ($M=1$), aunque algunos evaluadores también la han relacionado con la cultura musical ($M=0.33$), formación instrumental ($M=0.13$), vocal-auditiva ($M=0.13$), formación rítmica ($M=0.13$) y creatividad, composición e improvisación, ($M=0.07$), siendo estas en menor medida. Asimismo, según la rueda de la pedagogía de Carrington, esta herramienta incita al recuerdo ($M=0.80$), seguida de comprensión ($M=0.73$) y aplicación ($M=0.73$). Aunque, estas han sido las más valorada, también se ha mencionado otras.

En la dimensión técnica-estética, los ítems han sido valorados por los usuarios entre 2 y 5 puntos. La media se sitúa entre 3.6 y 4.33, lo que indica que, en promedio, la app se percibe como aceptable y buena. Por otro lado, la dimensión pedagógica-funcional, ha sido valorada por los evaluadores entre 1 y 5 puntos, lo que sugiere que existen diferentes percepciones sobre el uso de la app en relación con esta dimensión. La media oscila entre 2.07 y 4.13, indicando que, en general, la app se considera suficiente y buena, aunque solo ha recibido una valoración baja en el ítem "incita al trabajo colaborativo". Esto puede explicarse por el hecho de que este recurso móvil está diseñado exclusivamente para ser utilizado de manera individual, como un juego, y no en un contexto colaborativo, ya que se centra en el aprendizaje de contenidos musicales específicos. De manera similar, la dimensión musical ha sido valorada entre 1 y 5, lo que indica que algunos ítems presentan mayor dispersión que otros. La media obtenida en esta dimensión oscila entre 3.13 y 3.93, sugiriendo que el promedio de la app se considera aceptable.

Además, se observa que, de forma general, la desviación típica de encuentra entre 0.000 y 1.401, lo que indica que existe cierta variedad en los resultados, aunque esta variabilidad depende de cada ítem. No obstante, la mayoría se encuentra entre 0.3 y 1.3, lo que sugiere una variedad moderada. En cuanto a la asimetría, esta oscila entre -3.564 y 3.564, lo que indica que las distribuciones tienden a ser tanto simétricas como sesgadas, con algunas tendencias hacia valores altos o bajos. Por último, la curtosis se sitúa entre -2.051 y 11.071, sugiriendo que las distribuciones son más planas en los valores negativos, mientras que los valores positivos indican una distribución más puntiaguda; valores cercanos a 0 sugieren una distribución normal.

De manera consecutiva a estos datos, se obtiene la media generalizada para cada dimensión evaluada sobre la app móvil, así como la media total, con el fin de conocer el grado de percepción de este recurso tecnológico.

Tabla 5. Media general por dimensiones y por cómputo total.

	Mínimo	Máximo	Media	Desv. estándar	Asimetría	Curtosis
Dimensión técnica - estética	2,90	5,00	4,0000	,66548	-,190	-,1138
Dimensión pedagógica - funcional	1,75	5,00	3,5833	,92582	-,354	-,468
Dimensión musical	1,90	4,60	3,2867	,84926	,053	-,396
Cómputo total	2,28	4,87	3,6233	,76380	-,205	-,386

En la [Tabla 4](#), se observa que la media general obtenida en la dimensión técnica es de 4 puntos sobre 5, lo que se indica que la app ha tenido una valoración buena, sugiriendo que los usuarios están satisfechos con su rendimiento. La desviación estándar es baja, lo que sugiere

que las calificaciones tienden a concentrarse cerca de esa media. En cuanto a la asimetría y la curtosis, indican que la distribución está ligeramente sesgada hacia la izquierda y es achatada.

No obstante, en la dimensión pedagógica-funcional y en la dimensión musical, los valores obtenidos son neutrales, lo que sugiere que la app no se percibe con un fuerte impacto positivo ni negativo para el desarrollo del aprendizaje. En la dimensión pedagógica-funcional, se observa una mayor variabilidad en los resultados, con la asimetría y la curtosis sesgadas ligeramente hacia la izquierda y achatadas. Un patrón similar se presenta en la dimensión musical, donde las calificaciones son más moderadas en comparación con la anterior, sugiriendo una valoración menos destacada en estos aspectos.

De forma general, se observa que la app es un recurso que se encuentra entre un nivel neutral y aceptable, inclinándose más hacia este último. Por tanto, se podría sugerir que el recurso tecnológico evaluado por los especialistas sería aceptable para integrarse en el aula y fomentar el aprendizaje básico musical con la implementación de este recurso.gg

4.-DISCUSIÓN Y CONCLUSIÓN

La evaluación de un objeto tecnológico debe asegurar que los usuarios puedan valorar los aspectos clave como la calidad, la seguridad, la usabilidad y la eficacia del mismo (Cabero-Almenara et al., 2017; Quezada-Bolaños y Gallardo-López, 2023; Gutiérrez-Castillo et al., 2024). Este proceso es fundamental para identificar aquellas apps que siguen un patrón establecido, frente a aquellas que no lo hacen, lo que incrementa el porcentaje de apps que resultan satisfactorias para los usuarios (Joachims, et al., 2017; Quezada-Bolaños y Gallardo-López, 2023).

Las apps son consideradas como unos recursos tecnológicos que se pueden aplicar con mayor facilidad dentro del contexto educativo (Arts et al., 2021; Hernández-Lamas et al., 2021; Martín et al., 2021). De hecho, estos recursos se suelen utilizar con un promedio de 4,8 horas diarias (Ditrendia, 2022). Por tanto, deben emplearse como un enfoque pedagógico que fomente el aprendizaje de manera más autónoma, atractiva, motivadora, individualizada y adaptada al propio ritmo de cada estudiante. Esto ha sido respaldado en investigaciones como Liu y Shao (2022), Eusterbrock (2023), Yao y Li (2023), Barroso et al. (2024) quienes integraron el uso de las apps para formar estudiantes en el aula de música. Asimismo, el uso de estas apps contribuyó a incrementar el rendimiento académico de forma significativa en comparación con aquellos que no participaron en estudios centrados en apps móviles, favoreciendo así el desarrollo de competencias digitales (Gutiérrez-Castillo et al., 2023) y musicales (Cremades-Andreu y Campollo-Urkiza, 2022).

No obstante, a pesar de los beneficios mencionados, las apps de música deben ser evaluadas para conocer su calidad (Colebaugh et al., 2023; Domínguez-Lloria et al., 2024; Yu et al., 2024). Por estas razones, y en base al objetivo planteado, se ha evidenciado un alto nivel de acuerdo entre los 60 evaluadores sobre una misma app móvil, con un índice de fiabilidad de 0.959. Este nivel de confiabilidad ha alcanzado un valor de 0.961, con un nivel de confianza de 0.946 a 0.974 y con un nivel de significación inferior a 0.001. Según Martínez-Pérez y Pérez-Martín (2022), el índice de concordancia entre todos los evaluadores sobre un mismo objeto es excelente, dado que se encuentra en el rango entre 0.75 y 1, estando más cerca del valor máximo.

Además, en la prueba realizada prueba ANOVA con Friedman también se observó un alto índice de acuerdo, siendo $W = 800$, lo que refuerza los hallazgos obtenidos en el CCI. La media de la suma de los cuadrados intersujetos fue de 12.711, lo que indica una baja variabilidad

entre los codificadores. En cuanto a la suma de los cuadrados intrasujetos, la media fue de 167.598, y el valor del chi cuadrado de Friedman fue de 2966.244 ($p < 0.0001$), mostrando la variabilidad existente dentro de los sujetos en las diferentes evaluaciones, aunque sin ser excesiva.

La evaluación de esta app móvil ha permitido corroborar un alto índice de acuerdo entre todos los participantes, reflejado en las evaluaciones realizadas mediante el instrumento de evaluación proporcionado. Este elevado nivel de concordancia puede atribuirse a la comprensión total del instrumento de evaluación para apps en el contexto de la Educación Musical, así como a la capacidad de los docentes especialistas en esta área para identificar de manera individual las características, aspectos positivos y negativos de la app, evaluando sus dimensiones: técnica-estética, pedagógica-funcional y musical.

Además, independientemente del alto índice de acuerdo, es importante destacar que, aunque una app pueda tener diversas aplicaciones dentro de un contexto educativo, debe estar alineada con el contenido curricular de la materia en cuestión para ser integrada de manera adecuada. Esto permitirá a los docentes utilizar los recursos tecnológicos disponibles para desarrollar el aprendizaje musical de manera más enriquecedora, empleando diferentes estrategias, tal como sugirieron Liu y Shao (2022), Eusterbrock (2023), Yao y Li (2023) y Barroso et al. (2024).

En cuanto a la evaluación de la app, se ha observado que la dimensión técnica estética ha tenido una buena valoración, con 4 puntos sobre 5. Por otro lado, la dimensión pedagógica-funcional ha sido menor, con 3.58 puntos sobre 5 y la dimensión musical obtuvo 3.28 puntos sobre 5. Estos datos sugieren que, si bien la app presenta un buen diseño desde el punto de vista visual y de usabilidad, los contenidos integrados, así como los aspectos pedagógicos y funcionales, deben mejorarse progresivamente para alcanzar una mayor calidad en el recurso.

En pocas palabras, el nivel de acuerdo encontrado entre los evaluadores no solo respalda la fiabilidad del instrumento de evaluación utilizado, sino que también subraya la consistencia de la calidad de la app entre los docentes. Este nivel de concordancia, reflejado en los análisis estadísticos del CCI y el ANOVA con Friedman, demuestra que los docentes comparten una visión común sobre las fortalezas y debilidades de las apps móviles para el aprendizaje de la música. Este consenso es crucial para asegurar que las apps educativas implementadas en el aula sean de buena calidad, buscando así la efectividad del proceso de enseñanza-aprendizaje.

La alta consistencia de esta evaluación sugiere que, aunque la app tiene aspectos que deben mejorar, tiene un gran potencial para convertirse en un complemento valioso y adecuado para las clases de música, promoviendo así un aprendizaje más efectivo, atractivo e interdisciplinario.

FINANCIACIÓN

Proyecto subvencionado por el Ministerio de Ciencia, Innovación y Universidades, perteneciente a los subprogramas de Formación y Movilidad dentro del Programa Estatal de Promoción del Talento y su Empleabilidad en I+D+i.

CONTRIBUCIÓN DE LOS AUTORES

Antonio León-Garrido: Diseñador de la investigación. Conceptualización, Supervisión, Redacción – borrador original, Redacción – Revisión y edición.

Julio Manuel Barroso-Osuna: Diseñador de la investigación, Supervisión, Redacción – borrador original, Redacción – Revisión y edición.

Carmen Llorente-Cejudo: Diseñador de la investigación, Supervisión, Redacción – borrador original, Redacción – Revisión y edición.

REFERENCIAS

- Arts, I., Fischer, A., Duckett, D., y van der Wal, R. (2021). Information technology and the optimisation of experience – The role of mobile devices and social media in human-nature interactions. *Geoforum*, 122, 55–62. <https://doi.org/10.1016/j.geoforum.2021.03.009>
- Barragán-Sánchez, R., Romero-Tena, R., y García-López, M. (2022). Educational robotics to address behavioral problems in early childhood. *Education Sciences*, 13(1), 22. <https://doi.org/10.3390/educsci13010022>
- Barroso-Osuna, J., León-Garrido, A., Llorente-Cejudo, C., y Ruiz-Palmero, J. (2024). El impacto de las apps móviles y la carga del trabajo en la Educación Musical universitaria: Un estudio experimental. *Revista de Educación a Distancia (RED)*, 24(80). <https://doi.org/10.6018/red.602301>
- Cabero-Almenara, J., Llorente-Cejudo, C., y Gutiérrez-Castillo, J. J. (2017). Evaluación por y desde los usuarios: Objetos de aprendizaje con Realidad aumentada. *Revista de Educación a Distancia (RED)*, 53. <https://doi.org/10.6018/red/53/4>
- Calderón, J. C., Lorca, L. A., y Leao Ribeiro, I. (2023). Confiabilidad intra-evaluador relativa y absoluta en las evaluaciones de rango de movimiento de hombro y fuerza prensil en supervivientes de cáncer de mama: Estudio caso-control (Relative and absolute intrarater reliability in shoulder range of motion and handgrip strength assessments in breast cancer survivors: A case-control study). *Retos*, 48, 69–73. <https://doi.org/10.47197/retos.v48.95431>
- Colebaugh, C. A., Wilson, J. M., Flowers, K. M., Overstreet, D., Wang, D., Edwards, R. R., Chai, P. R., y Schreiber, K. L. (2023). The impact of varied music applications on pain perception and situational pain catastrophizing. *The Journal of Pain*, 24(7), 1181–1192. <https://doi.org/10.1016/j.jpain.2023.01.006>
- Correa-Rojas, J. (2021). Coeficiente de Correlación Intraclase: Aplicaciones para estimar la estabilidad temporal de un instrumento de medida. *Ciencias Psicológicas*. <https://doi.org/10.22235/cp.v15i2.2318>
- Creamades-Andreu, R., y Campollo-Urkiza, A. (2023). Alumnos y alumnas digitalmente competentes: Una aportación desde la educación musical. *ENSAYOS. Revista de La Facultad de Educación de Albacete*, 1–15. <https://doi.org/10.18239/ensayos.v38i1.3186>
- Del Pozo Franco, P. E., Peñafiel Palacios, A. J., y Cruz Piza, I. A. (2021). Estudio causal mediante Kendall y Pareto de la violencia contra la mujer en tiempos de confinamiento por COVID-19. *Dilemas Contemporáneos: Educación, Política y Valores*. <https://doi.org/10.46377/dilemas.v8i.2711>
- Delgado-Morales, C., y Duarte-Hueros, A. (2023). Una 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>
- Ditrendia (2022). “Ditrendia - Informe Mobile España y el Mundo”. Ditrendia. Digital Marketing Trends.

- Domínguez-Lloria, S., Oliveira, L. R., Diz-Otero, M., y Pino-Juste, M. (2023). Content evaluation of mobile device applications for teaching music in elementary education. *Multimedia Tools and Applications*, 83(16), 49673–49688. <https://doi.org/10.1007/s11042-023-17522-z>
- Eusterbrock, L. (2023). Mobile safe spaces and preset emotions: Making music with apps as a digital technology of the self. *Popular Music and Society*, 46(1), 50–69. <https://doi.org/10.1080/03007766.2022.2155029>
- Gallardo-Montes, C. D. P., Caurcel-Cara, M. J., y Rodríguez-Fuentes, A. (2021). Diseño de un sistema de indicadores para la evaluación y selección de aplicaciones para personas con Trastorno del Espectro Autista. *Revista Electrónica Educare*, 25(3), 1–24. <https://doi.org/10.15359/ree.25-3.18>
- Gomes, M. L. de S., Rodrigues, I. R., Moura, N. dos S., Bezerra, K. de C., Lopes, B. B., Teixeira, J. J. D., Vasconcelos, C. T. M., y Oriá, M. O. B. (2019). Avaliação de aplicativos móveis para promoção da saúde de gestantes com pré-eclâmpsia. *Acta Paulista de Enfermagem*, 32(3), 275–281. <https://doi.org/10.1590/1982-0194201900038>
- Gómez Redondo, C., Rico Rico, A., y Fontal Merillas, O. (2020). Diseño de una herramienta de evaluación de apps para la educación plástica. *Revista de Humanidades*, 39, 61. <https://doi.org/10.5944/rdh.39.2020.21814>
- Gutiérrez Castillo, J. J., León Garrido, A., y Barroso-Osuna, J. (2024). placa makey-makey en las aulas universitarias. *IJERI: International Journal of Educational Research and Innovation*, 22, 1–14. <https://doi.org/10.46661/ijeri.9621>
- Gutiérrez-Castillo, J. J., Palacios-Rodríguez, A., Martín-Párraga, L., y Serrano-Hidalgo, M. (2023). Development of digital teaching competence: Pilot experience and validation through expert judgment. *Education Sciences*, 13(1), 52. <https://doi.org/10.3390/educsci13010052>
- Hernández-Lamas, P., Cabau-Anchuelo, B., de Castro-Cuartero, Ó., y Bernabéu-Larena, J. (2021). Mobile applications, geolocation and information technologies for the study and communication of the heritage value of public works. *Sustainability*, 13(4), 2083. <https://doi.org/10.3390/su13042083>
- Joachims, T., Granka, L., Pan, B., Hembrooke, H., y Gay, G. (2017). Accurately interpreting clickthrough data as implicit feedback. *ACM SIGIR Forum*, 51(1), 4–11. <https://doi.org/10.1145/3130332.3130334>
- León-Garrido, A., Barroso-Osuna, J. M., y Llorente-Cejudo, M. C. (2024). Diseño y validación de un instrumento de evaluación para apps móviles musicales a través del juicio de expertos. *Aula Abierta*, 53(2), 169–179. <https://doi.org/10.17811/rifie.20620>
- Liu, X., y Shao, X. (2022). Modern mobile learning technologies in online piano education: Online educational course design and impact on learning. *Interactive Learning Environments*, 32(4), 1279–1290. <https://doi.org/10.1080/10494820.2022.2118787>
- Martin, A. J., Strnadová, I., Loblinzk, J., Danker, J. C., y Cumming, T. M. (2021). The role of mobile technology in promoting social inclusion among adults with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities*, 34(3), 840–851. <https://doi.org/10.1111/jar.12869>

- Martínez Pérez, J. A., y Pérez Martín, P. S. (2023). Coeficiente de correlación intraclase. *Medicina de Familia. SEMERGEN*, 49(3), 101907. <https://doi.org/10.1016/j.semerg.2022.101907>
- Martinez-Baquero, J. E., y Rodríguez-Umaña, L. A. (2022). Uso de aplicaciones móviles como herramienta de apoyo tecnológico para la enseñanza con metodología steam. *Revista Politécnica*, 18(36), 75–90. <https://doi.org/10.33571/rpolitec.v18n36a6>
- Quezada Bolaños, E. P., y Gallardo-López, L. (2023). TriSigEM: Un modelo semiótico para la evaluación de interfaces gráficas de usuario de apps. *Zincografía*. <https://doi.org/10.32870/zcr.v7i14.204>
- Rodríguez-Riesco, L., y Senín-Calderón, C. (2022). Aplicaciones móviles para evaluación e intervención en trastornos emocionales: Una revisión sistemática. *Terapia Psicológica*, 40(1), 131–151. <https://doi.org/10.4067/s0718-48082022000100131>
- Romero Rodríguez, J. M., Pombo, L., y Martínez Domingo, J. A. (2023). Implementación de los dispositivos móviles en educación superior: Casos de buenas prácticas docentes. *EDMETIC*, 12(1), 6. <https://doi.org/10.21071/edmetic.v12i1.15184>
- Romero-Tena, R., y Romero-González, A. (2020). Aprendizaje con robótica del patrón AB en niños de 3 años. *EduTEC. Revista Electrónica de Tecnología Educativa*, 72, 54–67. <https://doi.org/10.21556/edutec.2020.72.1579>
- Sdravopoulou, K., Castillo, J. J. G., y González, J. M. M. (2020). Naturalistic approaches applied to AR technology: An evaluation. *Education and Information Technologies*, 26(1), 683–697. <https://doi.org/10.1007/s10639-020-10283-4>
- Uribe-Guerrero, E. A., y Acosta-Mesa, D. R. (2023). Evaluación de Apps educativas relacionadas con el tema de balanceo de reacciones químicas. *Educación Química*, 34(2), 181–194. <https://doi.org/10.22201/fq.18708404e.2023.2.82606>
- Yao, B., y Li, W. (2023). The role of a teacher in modern music education: Can a student learn music with the help of modernized online educational technologies without teachers? *Education and Information Technologies*, 28(11), 14595–14610. <https://doi.org/10.1007/s10639-023-11786-6>
- Yu, G. Z., Liu, C. C., Piao, M., y Li, D. (2022). Teaching quality evaluation of China's network music education APP. *Applied Economics Letters*, 31(4), 265–269. <https://doi.org/10.1080/13504851.2022.2131708>