



Revista de Estilos de Aprendizaje / Journal of Learning Styles
ISSN: 1988-8996 / ISSN: 2332-8533

Anonimización facial con Inteligencia Artificial en Musicoterapia: estudio piloto con AKOOL Face Swap y Syntonym

Beatriz Amorós Sánchez

Universidad Internacional de La Rioja, UNIR, España

beatriz.amoros@unir.net

ORCID: <https://orcid.org/0000-0001-9652-7138>

David J. Gamella Gonzalez

Universidad Internacional de La Rioja, UNIR, España

david.gamella@unir.net

ORCID: <https://orcid.org/0000-0001-9834-954X>

Received: 10 October 2025 / Accepted: 30 March 2026

Resumen

En musicoterapia, compartir materiales clínicos obliga a buscar un equilibrio entre la privacidad y la información emocional, ya que los métodos tradicionales como el desenfoco pueden ocultar las expresiones del rostro y reducir el valor clínico del material. Este estudio piloto incluyó a 20 musicoterapeutas que evaluaron tres fragmentos de vídeo presentados en cuatro versiones: original, desenfoco y sustitución facial mediante dos aplicaciones de inteligencia artificial (AKOOL Face Swap y Syntonym). Las valoraciones se recogieron mediante escalas tipo Likert y se analizaron con procedimientos estadísticos, junto con estimaciones de fiabilidad y de concordancia entre evaluadores. Los resultados mostraron diferencias significativas ($\chi^2(2)=13.30$; $p=0.0013$), con mejoras relevantes en utilidad clínica y expresividad en las versiones con inteligencia artificial frente al desenfoco (≈ 0.8 – 0.9 puntos; $p<0.01$), sin diferencias entre AKOOL y Syntonym. La fiabilidad fue de aceptable a excelente ($\alpha=0.61$ – 0.88 ; $\omega=0.79$ – 0.91) y la concordancia entre evaluadores fue alta al considerar el promedio de las valoraciones ($ICC=0.86$). La sustitución facial con inteligencia artificial preserva mejor las señales expresivas y ofrece un equilibrio más favorable entre privacidad y utilidad que el desenfoco, lo que la hace viable para investigación y difusión pública bajo condiciones éticas y regulatorias adecuadas. No obstante, los materiales originales continúan siendo más apropiados para la docencia y la supervisión en contextos cerrados y con el consentimiento correspondiente.

Palabras clave: Musicoterapia; anonimización facial; inteligencia artificial; expresión facial; privacidad.

[en] Filogenia Facial anonymization with Artificial Intelligence in Music Therapy: a pilot study with AKOOL Face Swap and Syntonym

Abstract

In music therapy, sharing clinical materials requires balancing privacy protection with the preservation of emotional information, as traditional methods such as blurring can hide facial expressions and

reduce the clinical value of the material. This pilot study involved 20 music therapists who evaluated three video excerpts presented in four versions: original, blurred, and face substitution using two artificial intelligence applications (AKOOL Face Swap and Syntonym). Ratings were collected using Likert-type scales and analyzed with statistical procedures, together with estimates of reliability and inter-rater agreement. The results showed significant differences ($\chi^2(2)=13.30$; $p=0.0013$), with meaningful improvements in clinical utility and expressivity for the AI-based versions compared to blurring ($\approx 0.8-0.9$ points; $p<0.01$), and no differences between AKOOL and Syntonym. Reliability ranged from acceptable to excellent ($\alpha=0.61-0.88$; $\omega=0.79-0.91$), and inter-rater agreement was high when averaged across raters ($ICC=0.86$). AI-based face substitution better preserves expressive signals and provides a more favorable balance between privacy and utility than blurring, making it suitable for research and public dissemination under appropriate ethical and regulatory safeguards. However, original materials remain more appropriate for teaching and closed supervision settings with the required consent.

Keywords: Music therapy; facial anonymization; artificial intelligence; facial expression; privacy.

Sumario: 1.Introducción 1.1.Propósito y contribución del artículo 1.2.Fundamentos técnicos 1.3.Ejes regulatorios 1.4.Problema aplicado: visibilidad clínica y salvaguarda de identidad 1.5.Justificación clínica y objetivo del estudio 2.Objetivos 2.1.Objetivo general 2.2.Objetivos específicos 2.3.Hipótesis 3.Metodología 3.1.Diseño 3.2.Participantes y criterios 3.3.Materiales y herramientas de anonimización 4.3.2.Herramientas seleccionadas 3.4.Procedimiento del estudio 3.5.Análisis de datos 3.6.Chequeo de sesgo de orden 3.7.Fiabilidad interna 3.8.Consideraciones éticas y de seguridad 4.Desarrollo 4.1.Resultados 4.2.Fiabilidad interna 5.Discusión 6.Conclusiones. Referencias bibliográficas. Anexos

1. Introducción

1.1. Propósito y contribución del artículo

La musicoterapia se ha consolidado como una herramienta eficaz para abordar afecciones psicológicas, emocionales y sociales, promoviendo el bienestar mediante la música como medio expresivo y terapéutico (Bruscia, 2014; Bunt & Stige, 2014). Este artículo afronta el reto de documentar y compartir práctica clínica sin vulnerar la identidad de los pacientes, integrando un encuadre técnico, basado en inteligencia artificial y medios sintéticos, con un enfoque ético, regulatorio y operativo aplicable a contextos clínicos de musicoterapia. Se destacan las consideraciones éticas que orientan la toma de decisiones y la relación terapéutica (Dileo, 2007; Fattorini & Gamella, 2022). En el contexto actual, proteger la privacidad resulta crucial ante la expansión de redes sociales y de sistemas de difusión digital que incrementan el riesgo de exposición y pérdida de control sobre los datos personales (Sæle & Gilbertson, 2024).

1.2. Fundamentos técnicos

El desarrollo de la inteligencia artificial ha generalizado técnicas capaces de generar o transformar imágenes y audio, con impacto directo en la anonimización audiovisual en contextos clínicos (Agarwal et al., 2024; Goodfellow et al., 2014; Johnson et al., 2016; Karras et al., 2019; Liu et al., 2015; Yim et al., 2024). De forma complementaria, Flores Cruz (2025) señala que las tecnologías inmersivas, como la realidad virtual, pueden potenciar la interacción en contextos formativos si se aplican con criterios éticos y atención a la diversidad cognitiva, una idea alineada con el uso de sustitución facial para equilibrar privacidad y comunicación expresiva. En este marco, los *deepfakes*, contenidos generados o modificados mediante redes neuronales profundas, plantean un debate abierto sobre riesgos y oportunidades (Hamilton, 2025; Kern, 2025; Tariq et al., 2025; Westerlund, 2019).

1.3. Ejes éticos y regulatorios

Desde la perspectiva regulatoria, toda aproximación técnica debe ajustarse a los principios del Reglamento General de Protección de Datos (RGPD): base legal, minimización, limitación de propósito, seguridad, retención y evaluación de impacto cuando proceda (Parlamento Europeo & Consejo, 2016). La sustitución facial y la conversión de voz, se consideran formas de seudonimización reforzada, no de anonimización absoluta, dada la posibilidad de reidentificación multimodal y de

alteraciones en la carga afectiva (Abdulaziz & Bondarev, 2025; Chesney & Citron, 2019; Hamilton, 2025; Tariq et al., 2025). Por ello, el cumplimiento normativo debe acompañarse de criterios de equidad, gobernanza y evaluación de impacto en IA biomédica, junto con pautas sobre consentimiento granular y revisable, señalización visible de modificaciones, políticas de retención y borrado, y mecanismos de trazabilidad y auditoría (AMTA, s. f.; Baumgartner et al., 2023; Behrens, 2020; Clements-Cortés et al., 2025; Conduah et al., 2025; Fisher, 2021; Hewson et al., 2022; NHS Lothian, 2023; Qureshi & Khan, 2025). Además, la disponibilidad de manipulaciones visuales de bajo coste refuerza la necesidad de políticas explícitas y alfabetización crítica, especialmente en contextos docentes (Navarro Martínez et al., 2024; Paris & Donovan, 2019). En esta línea, Pérez y García (2025) subrayan la importancia de una formación tecnológica ética que integre gobernanza y pensamiento crítico.

1.4. Problema aplicado: visibilidad clínica y salvaguarda de identidad

Se define la visibilidad clínica como la posibilidad de que una persona sea visible o audible y, por ello, potencialmente identificable en registros o plataformas. Dicho concepto plantea dilemas sobre consentimiento informado, privacidad y agencia. La literatura desaconseja publicar artefactos terapéuticos en redes sociales por riesgos legales y daños emocionales derivados de borrados, réplicas no autorizadas o usos fuera del encuadre terapéutico, con impacto en pacientes y familias (Reid & Miño, 2021). Los consentimientos para fines secundarios —docencia, difusión o investigación— requieren matices adicionales cuando involucran clientes actuales o pasados y posibles asimetrías de poder, así como coherencia con los códigos deontológicos (American Music Therapy Association [AMTA], s. f.; Behrens, 2020). En telemusicoterapia, pese al aumento de accesibilidad, surgen preocupaciones sobre presencia de terceros, grabación, almacenamiento y eliminación de archivos, lo que exige guías claras y políticas operativas (Clements-Cortés et al., 2025; NHS Lothian, 2023).

Proteger la identidad mediante anonimización audiovisual implica equilibrar la preservación de expresividad, prosodia y lenguaje corporal con la confidencialidad del paciente (Franzreb et al., 2024). El desenfoque y la pixelización pueden degradar la calidad clínica y ser parcialmente reversibles, afectando señales afectivas y no verbales relevantes (Shuvi et al., 2020; Todt et al., 2024). En campos clínicos afines, se han explorado salvaguardas de privacidad en vídeo que ofrecen aprendizajes útiles (Lavanchy et al., 2023). No obstante, persisten desafíos éticos para equilibrar privacidad y utilidad clínica en un ecosistema tecnológico en rápida expansión que demanda criterios éticos y de calidad para su adopción (Fernández-Company et al., 2024; Kern, 2025; Amorós-Sánchez et al., 2024).

1.5. Justificación clínica y objetivo del estudio

Las propuestas recientes indican que, si están bien diseñadas, la sustitución facial y la conversión de voz pueden preservar la integridad del contenido clínico y reducir la identificación directa, siempre que mantengan microseñales de prosodia, sincronía e iniciación de la interacción (Agarwal, 2024; Franzreb et al., 2024; Gupta et al., 2019). La preservación de estas microseñales es esencial, ya que la evidencia clínica en musicoterapia muestra mejoras en comunicación social y participación terapéutica (Geretsegger et al., 2014). Además, la inteligencia artificial empieza a incorporarse en procesos creativos e interventivos, lo que refuerza la necesidad de evaluar usos responsables que equilibren calidad clínica y protección de identidad (Gamella-González, 2025). Sobre esta base, el presente trabajo desarrolla un estudio comparativo de sustitución facial con IA en musicoterapia, articulado en los ejes técnico y ético-regulatorio descritos.

2. Objetivos

Este estudio aborda vacíos en la anonimización audiovisual en contextos terapéuticos, especialmente, la pérdida de riqueza emocional asociada a métodos tradicionales. Se llevó a cabo la evaluación de dos aplicaciones, de sustitución facial, AKOOL Face Swap y Syntonym, como herramientas de anonimización en sesiones de musicoterapia. Se analizó su viabilidad técnica, el impacto en la calidad clínica —preservación prosódica, codificación afectiva y comprensión de la interacción— y las implicaciones éticas relativas a consentimiento, expectativas de uso y gobernanza. Además, se integró literatura reciente sobre beneficios clínicos, riesgos de reidentificación, sesgos y degradación de la utilidad del material. En los apartados siguientes se describen los objetivos del estudio, el diseño metodológico, las herramientas empleadas y el procedimiento de análisis de datos.

2.1. Objetivo general

Evaluar, desde el juicio profesional de musicoterapeutas, el equilibrio entre utilidad y expresividad frente a protección de identidad en tres estrategias de anonimización comparadas con el original: desenfoque, sustitución facial IA-A e IA-B, así como su aceptabilidad ética.

2.2. Objetivos específicos

Para dar cumplimiento al objetivo general, se establecieron los siguientes objetivos específicos, orientados a desglosar y operacionalizar los distintos componentes del análisis:

- Comparar la preservación de la expresividad facial y no verbal entre original, desenfoque, IA-A e IA-B.
- Valorar la utilidad clínica percibida, incluida la interpretabilidad del caso y su valor docente y supervisivo.
- Estimar la protección de identidad percibida en desenfoque, IA-A e IA-B, tomando el original como referencia.
- Evaluar la aceptabilidad ética de cada condición y describir preferencias de uso profesional en relación con consentimiento, señalización y gestión de materiales.
- Analizar preferencias de formato según el objetivo de uso.

2.3. Hipótesis

A continuación, se presentan las posibles hipótesis que serán resueltas al finalizar el artículo:

- H1. La sustitución facial con IA (IA-A e IA-B) preserva mejor la expresividad que el desenfoque.
- H2. El desenfoque y la sustitución facial con IA incrementan la protección de identidad respecto al original.
- H3. La aceptabilidad ética será mayor en la condición que logre el mejor equilibrio entre utilidad, expresividad y privacidad.

3. Metodología

3.1. Diseño

El estudio adoptó un diseño mixto secuencial explicativo, caracterizado por la recogida y el análisis inicial de datos cuantitativos, seguidos de una fase cualitativa orientada a explicar, ampliar y contextualizar dichos resultados (Creswell & Plano Clark, 2014). En la primera fase, cuasiexperimental intra-sujetos, se evaluó la viabilidad técnica y el desempeño de dos herramientas de sustitución facial con inteligencia artificial frente al desenfoque y al material original. La segunda fase, de análisis temático reflexivo (Braun & Clarke, 2019), interpretó los resultados cuantitativos desde la experiencia clínica de musicoterapeutas.

Participaron profesionales en ejercicio, centrados en la utilidad clínica y la aceptabilidad ética del material anonimizado. Se aplicaron fases abreviadas: familiarización, codificación inductivo-deductiva, generación y revisión de temas y síntesis reflexiva. Dos analistas analizaron por separado el 25 % del material para definir las primeras categorías. Luego compararon sus análisis y resolvieron las diferencias mediante acuerdo, aplicando después ese sistema de categorías al resto del material. Además, registraron observaciones sobre la utilidad, la expresividad y la aceptabilidad ética del material.

La integración de las fases se realizó mediante triangulación e interpretación conjunta de los resultados cuantitativos y cualitativos, con el fin de comprender beneficios y riesgos del uso de IA en grabaciones de musicoterapia (Creswell & Plano Clark, 2014; Pink et al., 2016). Las respuestas abiertas se recogieron tras cada bloque para reducir sesgos de recuerdo, y el meta-informe vinculó patrones cuantitativos con temas explicativos.

Para cada fragmento se generaron cuatro versiones (original, desenfoque, sustitución facial A y B) presentadas en secuencia controlada y con contrabalanceo parcial, es decir, alternando el orden de los vídeos entre los participantes para evitar que la posición o el cansancio influyeran en las valoraciones (Creswell & Plano Clark, 2014). Los evaluadores valoraron la calidad técnica (nitidez, sincronía audio-labial, coherencia temporal), la expresividad facial y no verbal, la protección de identidad, la utilidad clínica y la aceptabilidad ética.

La aceptabilidad ética se analizó conforme al RGPD, considerando consentimiento, minimización,

residencia y retención. Este diseño permitió relacionar el rendimiento técnico con la relevancia clínica y la adecuación ética de cada formato. Al tratarse de un estudio piloto, el objetivo fue generar señales explicativas para investigaciones confirmatorias.

3.2. Participantes y criterios

La muestra se estructuró en dos niveles complementarios: los vídeos clínicos y los musicoterapeutas evaluadores.

- Vídeos: se seleccionaron tres extractos públicos de sesiones de musicoterapia en YouTube, priorizando diversidad de rostros y planos (frontal, lateral y con movimiento) para probar distintas condiciones de anonimización. Cada fragmento (<3 min) incluyó actividades típicas —improvisación, vocalización y canto guiado— relevantes para generar respuestas emocionales y expresivas (Bunt & Stige, 2014). Se verificaron licencias y permisos para uso académico.

Las identidades sintéticas proceden de bibliotecas con licencia y no corresponden a personas reales. La participación fue anónima con consentimiento informado vía correo electrónico. Al no implicar intervención clínica ni tratamiento de datos de pacientes, no se requirió auditoría ética externa. Se registraron metadatos técnicos (resolución, FPS, tipo de plano) para controlar factores que podían influir en los resultados.

- Musicoterapeutas: en el presente estudio participaron 20 profesionales (12 mujeres, 8 hombres) reclutados por conveniencia en asociaciones de España, Ecuador, México, Irlanda y Colombia. Distribución: 60 % mujeres y 40 % hombres; edad y experiencia se describieron como media (DE) y rango. Los criterios de inclusión de la muestra que se utilizaron fueron: más de un año de experiencia clínica, familiaridad con principios éticos de confidencialidad, disponibilidad para completar visionado y encuesta. Se excluyeron quienes no cumplieron la experiencia mínima, no completaron las cuatro versiones o no otorgaron consentimiento. Todas las personas firmaron consentimiento informado (Anexo 1), y el estudio se ajustó al RGPD (Parlamento Europeo & Consejo, 2016). La limitada diversidad demográfica se justifica por el carácter piloto y el muestreo por conveniencia; el estudio confirmatorio prevé muestreo estratificado por género, experiencia y edad.

3.3. Materiales y herramientas de anonimización

3.3.1. Cribado y criterios de selección

Se llevó a cabo un mapeo sistemático del panorama actual de aplicaciones (ver tabla1) de sustitución facial con IA, identificándose inicialmente 15 herramientas.

Tabla 1

Clasificación de las herramientas digitales de anonimización con IA

Herramienta	Plataformas	Requiere programación	Interfaz estándar	Precio	Conserva expresiones	Garantiza privacidad
Syntonym	iOS, Android, escritorio	No	Sí	Pago	Sí	No / Sí opción <i>offline</i>
Egonym	Web/API (Win / mac / Linux vía navegador)	Sí	Sí	Pago	Sí	No
SwapMe AI	iOS, Android	No	Sí	Pago	Sí	No
Video Anonymizer	iOS	No	Sí	Gratis	No (pixelado)	Sí
PutMask	Android	No	Sí	Gratis	No (desenfocado)	Sí
ObscuraCam	Android	No	Sí	Gratis	No (desenfocado)	Sí
Reface	iOS, Android	No	Sí	Pago	Sí	No
DeepFaceLab	Windows, Linux	Sí	No	Gratis	Sí	No/ Sí opción <i>offline</i>
FaceSwap	Win, macOS, Linux	No	Sí	Gratis	Sí	Sí

DeepPrivacy 2	Win/Linux	Sí. (Python)	No	Gratis	Sí	Sí
IDDiffuse	Win/Linux	Sí (Python)	No	Gratis	Sí	Sí
AKOOL Face Swap	Web/API (Win / mac / Linux vía navegador)	No	Sí	Pago/ Gratis	Sí	Sí
DeepSwap	Web/API (Win / mac / Linux vía navegador)	No	Sí	Pago	Sí	No
Mango AI	Web/API (Win / mac / Linux vía navegador)	No	Sí	Pago	Sí	No
Pixanova	Web/API (Win / mac / Linux vía navegador)	No	Sí	Pago/ Gratis	Sí	No

Los criterios de cribado fueron:

- Disponibilidad multiplataforma: iOS, Android, macOS, Windows y web, con soporte activo.
- Uso sin programación mediante interfaz gráfica y tiempo de procesamiento competitivo.
- Calidad percibida basada en la conservación de expresividad facial (mirada, gestos y movimientos faciales) y coherencia temporal.
- Privacidad y seguridad mediante uso local u offline o, en caso de nube, contrato DPA, residencia de datos en la UE, retención nula y ausencia de analítica o terceros.
- Coste o licencias compatibles con fines académicos y clínicos.

Tras aplicar estos criterios, se seleccionaron AKOOL Face Swap y Syntonym, que cumplieron los requisitos y mostraron el mejor desempeño preliminar.

3.3.2 Herramientas seleccionadas

Una vez realizada la preselección, se contactó con los proveedores de ambas aplicaciones, que manifestaron su interés en participar:

-AKOOL Face Swap (IA-A) es una plataforma abierta en la que cualquier usuario puede crear una cuenta o contratar una suscripción; para este estudio se habilitaron temporalmente funciones equivalentes a un plan superior (Akool, 2025).

-Syntonym (IA-B) requiere acceso concedido por el proveedor, quien facilitó acceso temporal con fines de investigación y colaboro en la conversión de los vídeos (Syntonym, 2025).

Ni la provisión de acceso ni el contacto con los equipos técnicos influyeron en el diseño, parametrización, análisis o reporte del estudio. Como control tradicional se incluyó el desenfoque gaussiano aplicado a la región facial.

3.4. Procedimiento del estudio

El procedimiento se estructuró en tres etapas: preparación, anonimización y evaluación. A continuación, se detalla cada una de ellas.

- Preparación: se seleccionaron las sesiones descritas en el apartado 3.2 y se cribaron las herramientas según 3.3.1, eligiendo AKOOL Face Swap y Syntonym como condiciones experimentales y el desenfoque gaussiano (técnica que difumina la región facial mediante un filtro matemático para ocultar la identidad) como control. Para su posible uso clínico se documentó la configuración de cumplimiento RGPD, incluyendo contrato DPA, residencia de datos en la UE, retención cero y ausencia de terceros.

- Anonimización: para cada fragmento se generaron cuatro versiones: 1. vídeo original sin modificar, 2. desenfoque gaussiano en la región facial, 3. sustitución facial A (AKOOL) y 4. sustitución facial B (Syntonym). Todas se distribuyeron en una carpeta compartida para el acceso de los evaluadores.

- Evaluación: veinte musicoterapeutas (ver 3.2) visionaron tres extractos en las cuatro versiones y completaron un cuestionario anónimo en Microsoft (Anexo 2) con ítems tipo Likert de 1 a 5, donde 1 correspondía a “muy deficiente”, 2 a “deficiente”, 3 a “aceptable”, 4 a “bueno” y 5 a “excelente”,

además de preguntas abiertas. Los ítems valoraron, por condición, la preservación de la expresividad facial y no verbal, la utilidad clínica, el valor para docencia y supervisión, la protección de identidad percibida y la calidad técnica (nitidez, sincronía audio-labial y coherencia temporal). Los estímulos se estandarizaron en parámetros audiovisuales, manteniendo el audio original. Se recogieron además juicios sobre aceptabilidad ética y preferencias de uso y formato. La participación fue voluntaria y anónima, y el procedimiento se ajustó al RGPD, sin requerir auditoría ética externa por su carácter no intervencional.

3.5. Análisis de datos

Para el análisis de datos se consideraron dos tipos de información: datos cuantitativos, correspondientes a las puntuaciones obtenidas en las escalas tipo Likert, y datos cualitativos, procedentes de las respuestas abiertas de los participantes. Ambos conjuntos se analizaron de forma complementaria para ofrecer una visión integrada de los resultados.

- Cuantitativo: las puntuaciones de calidad y utilidad clínica se analizaron mediante Friedman para comparar la IA y los videos de desenfoque, con descriptivos (media, desviación estándar) y post-hoc (Wilcoxon) para diferencias específicas (Creswell & Plano Clark, 2014).
- Cualitativo: las entrevistas se transcribieron íntegramente y se sometieron a análisis temático, identificando categorías emergentes sobre utilidad clínica, aceptabilidad ética y desafíos. Se aplicaron las directrices de Braun & Clarke (2019), con triangulación entre múltiples codificadores y soporte en Atlas.ti para trazabilidad y consenso.
- Análisis ético: los registros de auditoría se examinaron cualitativamente para verificar alineamiento con GDPR y principios éticos de la disciplina, identificando riesgos y medidas de mitigación (Fisher, 2021).

3.6. Chequeo de sesgo de orden

Dado que el orden de presentación fue fijo (desenfocado → Akool → Syntonym) y no se dispone de la condición “Original” en la matriz de calificaciones final, evaluamos una tendencia monótona por posición con la prueba de Page (N=20, k=3; alternativa unidireccional “no disminuye con la posición”). El p-valor se obtuvo por permutación Monte Carlo (100 000 repeticiones), reordenando los rangos dentro de cada participante y manteniendo los posibles empates.

De forma complementaria, contrastamos también la tendencia decreciente (fatiga) invirtiendo el orden, y realizamos una comparación exploratoria posición 1 vs. 3 (desenfocado vs. Syntonym) mediante Wilcoxon pareado con aproximación normal y corrección por continuidad.

3.7. Fiabilidad interna

Se estimaron el α de Cronbach y el ω total (modelo unifactorial) del compuesto por condición, obteniéndose valores de fiabilidad aceptables en desenfoque ($\alpha = .61$; $\omega = .79$) y elevados en las condiciones con IA (AKOOL: $\alpha = .81$; $\omega = .87$; Syntonym: $\alpha = .88$; $\omega = .91$).

3.8. Consideraciones éticas y de seguridad

El estudio no hubo auditoría ética externa al no trabajar directamente con pacientes y solamente usar un cuestionario anonimizado, por lo que cumplió los principios de confidencialidad, consentimiento informado y protección de datos (Warrier et al., 2023). Los datos originales se almacenaron en servidores seguros con cifrado AES-256; el acceso quedó restringido al equipo investigador. Los musicoterapeutas recibieron información clara sobre objetivos, procedimientos e implicaciones del uso de IA y participaron de forma voluntaria, indicando en el propio cuestionario que estaban conformes con la información recibida (Parlamento Europeo & Consejo, 2016).

4. Desarrollo

4.1 Resultados

En primera instancia, y dado que el tamaño muestral es inferior a 50, se evaluó la normalidad mediante Shapiro–Wilk sobre las puntuaciones medias por participante en cada formato. Empleamos pruebas no paramétricas por la escala ordinal y el diseño intra-sujeto, con o sin normalidad. Aunque la mayoría de las distribuciones no vulneraron la normalidad (Shapiro–Wilk, $p > .05$), se emplearon Friedman/Wilcoxon por la escala ordinal y el diseño intra-sujeto, salvo en la normalidad de la

diferencia Syntonym–Desenfocado ($p \approx 0.022$), considerada para elegir pruebas pareadas conservadoras (ver tabla 2). Las tablas de resultados pueden consultarse en el Anexo 3.

Tabla 2

Pruebas de normalidad (Shapiro–Wilk) por formato ($p > 0.05$ indica normalidad).

Formato	W	p
Desenfocado	0.971	0.778
AKOOL	0.954	0.426
Syntonym	0.957	0.491

Posteriormente, se empleó un contraste intra-sujeto de Friedman ($k=3$) y comparaciones pareadas con la prueba de Wilcoxon, aplicando corrección de Benjamini–Hochberg (BH): Prueba de Friedman: $\chi^2(2) = 13.30$, $p = 0.0013$; W de Kendall = 0.333. Se presentan estadísticos descriptivos (Tabla 3 y diagramas de cajas de la Figura 2) y se observa que los valores medios de las aplicaciones IA ($M = 3.79$, $SD = 0.65$ y $M = 3.73$, $SD = 0.85$) son superiores que los del método desenfocado ($M = 2.93$, $SD = 0.75$). Para una visualización de las medias ($\pm DE$ = desviación estándar) por condición véase la Figura 1, que facilita la comparación directa entre desenfoco y las dos opciones de sustitución facial con IA.

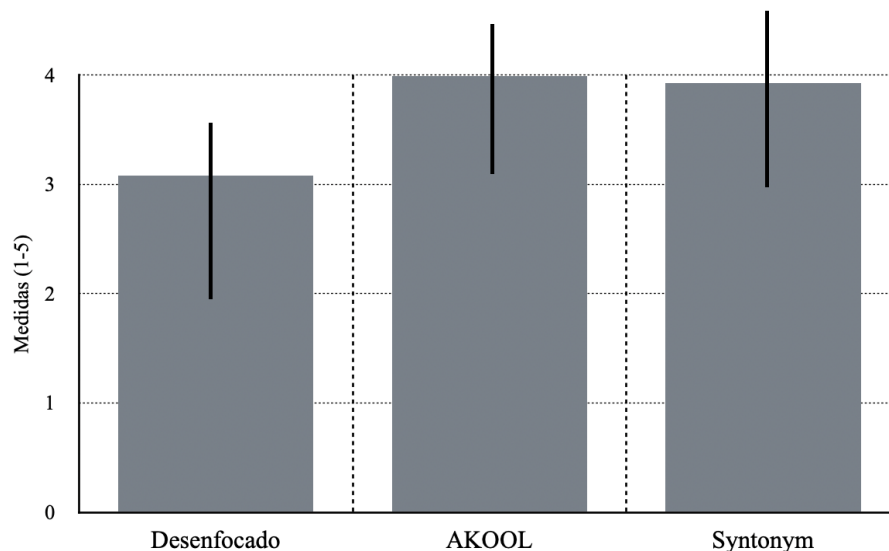
Tabla 3

Estadísticos descriptivos por formato (promedio por participante).

Formato	N	Media	DE	Mediana	Min	Max	IC95_L	IC95_U
Desenfocado	20	2.93	0.75	2.81	1.50	4.25	2.60	3.26
AKOOL	20	3.79	0.65	3.71	2.67	4.90	3.51	4.08
Syntonym	20	3.73	0.85	3.67	1.89	5.00	3.35	4.10

Figura 1

Puntuaciones medias ($\pm DE$) por condición (Desenfocado, AKOOL, Syntonym).



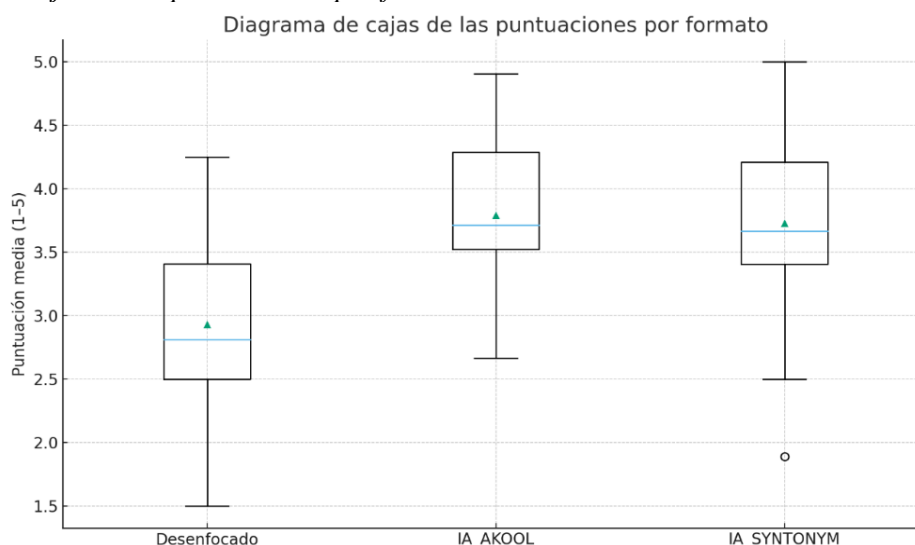
Nota: Barras de error: desviación estándar. Escala 1–5; mayores valores indican mayor utilidad/expresividad clínica.

La Figura 2 muestra la distribución por condición mediante boxplots con puntos individuales por participante, destacando la mayor centralidad y menor dispersión de las puntuaciones en AKOOL y

Syntonym frente al desenfoque.

Figura 2

Diagrama de cajas de las puntuaciones por formato.



Complementariamente, las comparaciones pareadas (Tabla 4) confirman el patrón observado en las figuras descriptivas. Tanto AKOOL como Syntonym superaron al desenfoque ($p_{BH} < 0.01$; $r \approx 0.66$ – 0.68 , efectos grandes), sin diferencias significativas entre ambas herramientas de IA ($p_{BH} \approx 0.84$; $r \approx 0.05$).

Tabla 4

Comparaciones pareadas (Wilcoxon, dos colas) con corrección BH; se informa r como tamaño de efecto.

Pareja	W	z aprox.	p	p _{BH}	R
AKOOL vs Syntonym	98.50	-0.24	0.8408	0.8408	0.05
AKOOL vs Desenfocado	23.00	-3.06	0.0012	0.0036	0.68
Syntonym vs Desenfocado	26.50	-2.93	0.0023	0.0035	0.66

Nota. W = estadístico de la prueba de Wilcoxon; z = valor estandarizado; p = nivel de significación; p_{BH} = valor de p corregido mediante el procedimiento de Benjamini–Hochberg; r = tamaño del efecto.

Los análisis respetan la naturaleza ordinal de las escalas y el diseño intra-sujeto. El contraste global de Friedman fue significativo ($\chi^2(2) = 13.30$, $p = .0013$; $W = 0.333$, efecto moderado), evidenciando diferencias entre formatos. En las comparaciones pareadas, AKOOL y Syntonym superaron al desenfoque tras corrección BH ($p < .01$; $r \approx 0.66$ – 0.68 , efectos grandes), sin diferencias significativas entre ambas ($p_{BH} \approx 0.84$). Este resultado no implica equivalencia salvo verificación mediante TOST con márgenes ± 0.30 y potencia adecuada.

Los intervalos de confianza (IC95%) indican el margen dentro del cual es probable que se encuentre el valor real de la media. Estos resultados sitúan las medias de IA por encima del desenfoque (AKOOL ≈ 3.79 , Syntonym ≈ 3.73 vs. desenfoque ≈ 2.93). En una escala de 1–5, la mejora de 0.8–0.9 puntos es clínicamente relevante, al reflejar mejor conservación de señales no verbales y prosódicas.

El efecto de posición mostró una tendencia creciente significativa (Page: $L = 259.0$, $p_{1-cola} = 0.00143$, permutación Monte Carlo 100 000 réplicas) y ninguna tendencia decreciente asociada a fatiga ($L' = 221.0$, $p_{1-cola} = 0.999$). La comparación exploratoria entre la posición 1 y 3 (desenfocado vs. Syntonym) fue significativa (Wilcoxon: $W = 22.5$, $z = -3.06$, $p = 0.0022$, bilateral).

La Figura 4 resume los tamaños de efecto (r): grandes para AKOOL y Syntonym frente al desenfoque

($r \approx 0.66-0.68$) y despreciable entre ambas herramientas ($r \approx 0.05$).

Figura 3

Preferencias por formato según finalidad (frecuencia).

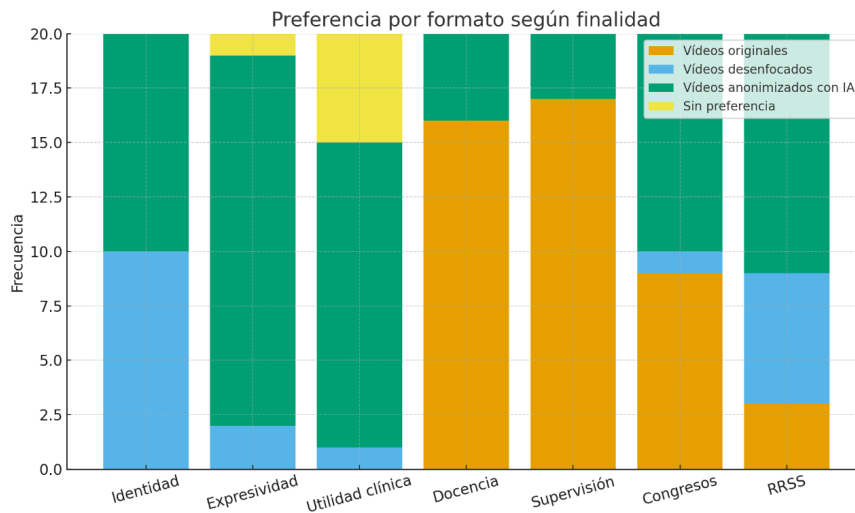
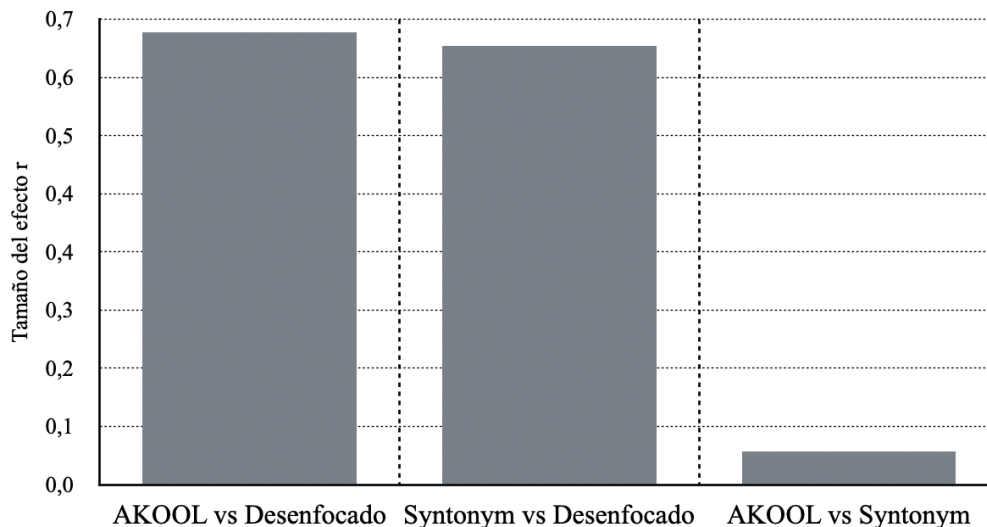


Figura 3 muestra la preferencia por IA en fines de identidad, divulgación y congresos, y la preferencia por originales en docencia y supervisión.

Figura 4

Tamaños de efecto (r) de Wilcoxon para comparaciones pareadas: AKOOL vs Desenfocado; Syntonym vs Desenfocado; AKOOL vs Syntonym



5.2. Fiabilidad interna

El análisis cuantitativo se realizó sobre tres formatos (desenfoco, AKOOL, Syntonym; $k=3$); el original se utilizó para preferencias por finalidad y discusión contextual. Estimamos α de Cronbach y ω total (modelo unifactorial) del compuesto por condición, promediando por participante las puntuaciones de los ítems disponibles en los 3 vídeos (DESENFOCADO: ítems 1–3; AKOOL: 1–7; Syntonym: 1–6). Intervalos al 95% por bootstrap (1.500 réplicas sobre participantes).

Tabla 5

Fiabilidad interna del compuesto (medias de ítems por condición)

Condición	k (ítems)	α (IC95%)	ω (IC95%)
-----------	-----------	------------------	------------------

Desenfocado	3	0.611 (0.325–0.802)	0.786 (0.655–0.888)
Akool	7	0.814 (0.652–0.881)	0.870 (0.755–0.922)
Syntonym	6	0.879 (0.709–0.947)	0.913 (0.789–0.963)

Por otro lado, para evaluar el grado de acuerdo entre los evaluadores, se calculó el coeficiente de correlación intraclase (ICC) para el conjunto de combinaciones vídeo × condición. Se obtuvieron dos estimaciones: una para un único evaluador (ICC(2,1)) y otra para el promedio de los 20 evaluadores (ICC(2,k)). Los resultados mostraron un acuerdo bajo cuando se considera a un evaluador individual (ICC(2,1) = 0.241), pero un acuerdo alto cuando se promedian las valoraciones de todos los participantes (ICC(2,k) = 0.864), lo que indica una buena consistencia global entre evaluadores.

Del análisis de las respuestas abiertas emergieron varios temas principales. En primer lugar, los participantes destacaron que las versiones con sustitución facial mediante inteligencia artificial conservaban mejor la expresividad y facilitaban la interpretación clínica en comparación con el desenfoque. En segundo lugar, se señaló la importancia de la señalización visible de las modificaciones para evitar confusión o malentendidos en el uso de los materiales. Asimismo, algunos musicoterapeutas expresaron preocupación por el posible riesgo de reidentificación y por el uso de estos vídeos fuera de contextos profesionales controlados. Finalmente, se observó una preferencia por emplear las versiones con sustitución facial para fines de investigación y difusión, mientras que los materiales originales se reservarían para docencia y supervisión en entornos cerrados.

5. Discusión

Este estudio examinó, con juicio profesional de musicoterapeutas, el equilibrio entre utilidad, expresividad y protección de identidad en cuatro condiciones: original, desenfoque y sustitución facial IA-A e IA-B. Se evaluó su aceptabilidad ética y las preferencias de uso. Las hipótesis fueron: H1, la sustitución facial preserva mejor la expresividad que el desenfoque; H2, ambas técnicas incrementan la protección frente al original; y H3, la aceptabilidad aumenta al equilibrar utilidad y privacidad.

El marco teórico sitúa el trabajo entre la protección de la identidad en entornos hiperexpuestos (AMTA, s. f.; Behrens, 2020; Sæle & Gilbertson, 2024) y la necesidad de mantener señales clínicas esenciales como prosodia, gestualidad y sincronía (Bruscia, 2014; Bunt & Stige, 2014). La literatura advierte que el desenfoque degrada la información afectiva, mientras que las técnicas generativas pueden conservarla (Shuvi et al., 2020; Todt et al., 2024; Westerlund, 2019).

Los resultados obtenidos respaldan este planteamiento: hubo diferencias globales ($\chi^2(2)=13.30$, $p=0.0013$; $W=0.33$). IA-A e IA-B superaron al desenfoque (efectos grandes) sin diferir entre sí, apoyando H1 y estudios previos (Franzreb et al., 2024; Todt et al., 2024). En H2, la protección percibida fue mayor con desenfoque e IA que con el original, en línea con AMTA y Behrens (s. f.; 2020). En H3, la aceptabilidad dependió del equilibrio privacidad-utilidad: las versiones con IA fueron preferidas para difusión e investigación, y los originales, para docencia y supervisión con consentimiento reforzado (Bruscia, 2014; Bunt & Stige, 2014; Gold, 2016; Kim & Lee, 2018).

La literatura en salud mental refuerza esta interpretación: las técnicas tipo deepfake ofrecen potencial clínico, pero requieren salvaguardas éticas y transparencia (Hoek, 2025). En formación sanitaria, se observan valor educativo y preocupaciones por privacidad (Martínez et al., 2024). Las transformaciones deben calibrarse para no distorsionar información paralingüística (Zhu et al., 2024). Si la anonimización conserva señales no verbales y prosodia, la utilidad clínica aumenta, siempre que existan barreras frente a reidentificación o desinformación.

En el plano ético-jurídico, el RGPD considera la sustitución facial y la conversión de voz como seudonimización reforzada, con riesgos residuales de reidentificación (Abdulaziz & Bondarev, 2025; Chesney & Citron, 2019; Parlamento Europeo & Consejo, 2016). Se requieren gobernanza, transparencia y auditoría (Baumgartner et al., 2023; Conduah et al., 2025; Clements-Cortés et al., 2025; NHS Lothian, 2023).

La calidad de medida fue adecuada (α y ω en rangos esperables, alta concordancia interevaluador ICC(2,k)). Aunque el orden fijo pudo inducir sesgos, la prueba de Page ($L=259.0$, $p\approx 0.0014$) mostró una tendencia creciente sin afectar conclusiones. Entre las limitaciones figuran el tamaño muestral reducido ($n=20$), el muestreo por conveniencia y el uso de extractos de YouTube. Se recomienda incluir subescalas validadas de expresividad, sincronía e identidad y replicar en grabaciones clínicas

reales (Khalid et al., 2021; 2023).

En telesalud, donde el vídeo es ubicuo, los datos respaldan sustituir el desenfoque por seudonimización más fiel, con consentimiento granular y señalización explícita. Esto coincide con guías de telepsicología (APA, 2024) y marcos éticos digitales que exigen transparencia y trazabilidad (Tiribelli et al., 2023).

Líneas de acción: en docencia cerrada, usar originales con consentimiento reforzado; en investigación y difusión, preferir IA con señalización visible, control de retención y evaluación de impacto. Desarrollar subescalas validadas, integrar conversión de voz y sustitución facial y vincular rendimiento técnico con valor clínico (Todt et al., 2024; Sun et al., 2024).

Como pautas prácticas, se proponen consentimiento continuo y revisitable, señalización visible, auditorías periódicas de riesgo de reidentificación y evaluación de impacto para usos docentes o públicos. Estas medidas refuerzan el equilibrio privacidad-utilidad observado.

En síntesis, la sustitución facial con IA preserva mejor las señales expresivas que el desenfoque y ofrece un balance privacidad-utilidad más favorable para investigación y difusión. El original mantiene un valor formativo superior. La decisión de uso debe ser contextual y regirse por consentimiento, transparencia y trazabilidad, pilares de una telesalud segura y eficaz en musicoterapia.

6. Conclusiones

Los resultados permiten dar respuesta a los objetivos planteados, al mostrar que la sustitución facial con inteligencia artificial ofrece un equilibrio más favorable entre privacidad y utilidad clínica que el desenfoque tradicional. En línea con la literatura que señala las limitaciones del desenfoque y las oportunidades, aunque con riesgos, de los medios sintéticos (Franzreb et al., 2024; Todt et al., 2024; Westerlund, 2019), se confirma que esta técnica permite compartir material clínico conservando señales expresivas esenciales, siempre que se aplique bajo criterios de gobernanza ética y controles de calidad adecuados.

Los hallazgos respaldan su uso en investigación y comunicación externa cuando existen salvaguardas como el consentimiento informado, la minimización de datos, la señalización visible, y políticas claras de retención, borrado, auditoría y trazabilidad. No obstante, el material original continúa siendo preferible para la docencia y la supervisión en entornos cerrados y con consentimiento reforzado.

Se podrían indicar como líneas futuras de investigación las siguientes: ampliar la diversidad de casos mediante estudios confirmatorios con muestras mayores ($n = 50-100$) y muestreo estratificado por género, experiencia y edad, así como utilizar grabaciones clínicas diseñadas y recogidas expresamente para la investigación con consentimiento informado; y explorar la combinación de rostro sintético y conversión de voz, incorporando métricas de inteligibilidad, congruencia afectiva y riesgo de reidentificación que complementen el juicio experto, junto con la validación del cuestionario en población de musicoterapeutas. Este trabajo tiene un alcance metodológico y no sustituye el asesoramiento ético o legal profesional.

Referencias

- Abdulaziz, S., & Bondarev, E. (2025). *Unmasking performance gaps: A comparative study of human anonymization and its effects on video anomaly detection* (arXiv:2507.14083). arXiv. <https://doi.org/10.48550/arXiv.2507.14083>
- Agarwal, S., Peta, S., & Panyam, S. (2024). Deepfakes in healthcare: Reviewing the transformation potential and its challenges. *International Journal of Intelligent Systems and Applications in Engineering*, 12(4), 3965–3970.
- Akool. (n.d.). *Best face swap for photos & videos online with free trial*. Akool. <https://akool.com/apps/faceswap>
- Amorós-Sánchez, B., Gamella-González, D. J., Cisneros-Álvarez, P., & Gisbert-Caudeli, V. (2024). A systematic review of the technology available for data collection and assessment in music therapy. En A. L. Brooks (Ed.), *ArtsIT 2023... Proceedings, Part I* (pp. 41–54). Springer. https://doi.org/10.1007/978-3-031-55319-6_4
- American Music Therapy Association. (n.d.). *Code of ethics*. <https://www.musictherapy.org/about/ethics/>

- American Psychological Association (2024). Proposed Revision of Guidelines for the Practice of Telepsychology. Tomado de: <https://www.apa.org/practice/guidelines/telepsychology-revisions.pdf>
- Baumgartner, R., Arora, P., Bath, C., Williams, R. (2023). Fair and equitable AI in biomedical research and healthcare: Social science perspectives. *Artificial Intelligence in Medicine*, 144, 102658. <https://doi.org/10.1016/j.artmed.2023.102658>
- Behrens, G. A. (2020). Considerations when writing and presenting consent forms for clients. *Music Therapy Perspectives*, 38(1), 38–41. <https://doi.org/10.1093/mtp/miz029>
- Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, 11(4), 589–597. <https://doi.org/10.1080/2159676X.2019.1628806>
- Bruscia, K. E. (2014). *Defining music therapy* (4th ed.). Barcelona Publishers.
- Bunt, L., & Stige, B. (2014). *Music therapy: An art beyond words* (2.^a ed.). Routledge.
- Chesney, R., & Citron, D. (2019). Deepfakes: A looming challenge for privacy, democracy, and national security. *California Law Review*, 107, 1753–1819. <https://doi.org/10.2139/ssrn.3213954>
- Clements-Cortés, A., Fuller, A., Kelly, L., Pranjić, M., Selvarajah, I., Mercadal-Brotons, M., & Bridi, N. (2025). Music therapists' global perspectives on telehealth music therapy: A qualitative interview inquiry. *Music Therapy Perspectives*. Advance online publication. <https://doi.org/10.1093/mtp/miae030>
- Conduah, A. K., Ofoe, S., & Siaw-Marfo, D. (2025). Data privacy in healthcare: Global challenges and solutions. *Digital Health*, 11, 20552076251343959. <https://doi.org/10.1177/20552076251343959>
- Creswell, J. W., & Plano Clark, V. L. (2014). *Designing and conducting mixed methods research* (2.^a ed.). SAGE.
- Dileo, C. (2007). *Ethical thinking in music therapy*. Jeffrey Books.
- Fattorini Vaca, A., & Gamella-González, D. J. (2022). Ética profesional en musicoterapia desde la perspectiva del musicoterapeuta. En S. Olivero Guidobono (Coord.), *Artes y humanidades en el centro de los conocimientos: Miradas sobre el patrimonio, la cultura, la historia, la antropología y la demografía* (pp. 637–670). Dykinson. <https://www.dykinson.com/libros/artes-y-humanidades-en-el-centro-de-los-conocimientos-miradas-sobre-el-patrimonio-la-cultura-la-historia-la-antropologia-y-la-demografia/9788413779263/>
- Fernández-Company, J. F., Gamella-González, D. J., & García-Rodríguez, M. (2024). Autoevaluación de la práctica en musicoterapia para el crecimiento profesional. En M. del M. Simón Márquez, P. Molina Moreno, J. J. Gázquez Linares, & S. Fernández Gea (Coords.), *Innovación en salud: Estrategias emergentes para la docencia y la investigación* (pp. 27–33). ASUNIVEP. <https://dialnet.unirioja.es/servlet/libro?codigo=974646>
- Fisher, C. B. (2021). *Decoding the ethics code: A practical guide for psychologists* (5th ed.). SAGE.
- Flores Cruz, R. (2025). *Uso de la realidad virtual para afrontar obstáculos didácticos, atendiendo los principales estilos de aprendizaje*. Revista de Estilos de Aprendizaje, 18(35). Recuperado de <https://revistaestilosdeaprendizaje.com/index>
- Franzreb, C., Das, A., Gieseler, H., Jahn, E. C., Polzehl, T., & Möller, S. (2024). Towards audiovisual anonymization for remote psychotherapy: A subjective evaluation. En *Proceedings of SPSC 2024* (pp. 102–110). ISCA. <https://doi.org/10.21437/SPSC.2024-17>
- Gamella-González, D. J. (2025). Proyecto MUTia: Inteligencia artificial aplicada al songwriting en musicoterapia. En D. De la Rosa & P. Renés (Eds.), *Políticas educativas y nuevos modelos de enseñanza-aprendizaje en educación superior y formación profesional* (pp. 229–252). Tirant lo Blanch.
- Geretsegger, M., Elefant, C., Mössler, K. A., & Gold, C. (2014). Music therapy for people with autism spectrum disorder. *Cochrane Database of Systematic Reviews*, 2014(6), CD004381. <https://doi.org/10.1002/14651858.CD004381.pub3>
- Gold, C. (2016). Abstracts of the 10th European Music Therapy Conference. *Nordic Journal of Music Therapy*, 25(sup1), 1–156. <https://doi.org/10.1080/08098131.2016.11783620>
- Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial networks. *arXiv*. <https://doi.org/10.48550/arXiv.1406.2661>

- Gupta, A., Malik, P., Pathania, M., & Rathaur, V. K. (2019). Overview of artificial intelligence in medicine. *Journal of Family Medicine and Primary Care*, 8(7), 2328–2331. <https://doi.org/10.4103/jfmpe.jfmpe.440.19>
- Hamilton, L. (2025). *The future of music therapy: An exploration of music therapists' perceptions of artificial intelligence and its ethical implications* [Trabajo fin de grado, Edinburgh College of Art, Universidad de Edimburgo]. Edinburgh College of Art Graduate Show. <https://www.2025.graduateshow.eca.ed.ac.uk/sites/default/files/2025-05/LH%20Dissertation%20ECA%20Graduate%20Show%202025.pdf>
- Hewson, T., Abraham, S., Randles, N., Akinola, A., Cliff, R., Byrne, P., & Ramkisson, R. (2022). The recording of mental health consultations by patients: Clinical, ethical and legal considerations. *BJPsych Bulletin*, 46(3), 133–137. <https://doi.org/10.1192/bjb.2021.89>
- Hoek S, Metselaar S, Ploem C. & Bak, C. (2025). Promising for patients or deeply disturbing? The ethical and legal aspects of deepfake therapy. *Journal of Medical Ethics* 51:481-486. <https://doi.org/10.1136/jme-2024-109985>
- Johnson, J., Alahi, A., & Fei-Fei, L. (2016). Perceptual losses for real-time style transfer and super-resolution. En *ECCV 2016* (pp. 694–711). Springer. https://doi.org/10.1007/978-3-319-46475-6_43
- Karras, T., Laine, S., & Aila, T. (2019). A style-based generator architecture for generative adversarial networks. En *CVPR 2019* (pp. 4401–4410). IEEE. <https://doi.org/10.1109/CVPR.2019.00453>
- Kern, P. (2025). Artificial intelligence in music therapy: A new era of personalized care and scalable impacts. *MiSOSTENiDO*, 5(9), 8–16. <https://doi.org/10.59028/misostenido.2025.02>
- Khalid, H., Kim, M., Tariq, S., & Woo, S. S. (2021). Evaluation of an audio-video multimodal deepfake dataset using unimodal and multimodal detectors. En *ADGD '21, ACM MM Workshops* (pp. 7–15). ACM. <https://doi.org/10.1145/3476099.3484315>
- Khalid, N., Qayyum, A., Bilal, M., Al-Fuqaha, A., & Qadir, J. (2023). Privacy-preserving artificial intelligence in healthcare: Techniques and applications. *Computers in Biology and Medicine*, 158, 106848. <https://doi.org/10.1016/j.compbiomed.2023.106848>
- Kim, J., & Lee, S. (2018). Music therapy for depression: A meta-analysis. *The Arts in Psychotherapy*, 59, 103–112. <https://doi.org/10.1016/j.aip.2018.03.004>
- Lavanchy, J. L., Vardazaryan, A., Mascagni, P., et al. (2023). Preserving privacy in surgical video analysis using a deep learning classifier to identify out-of-body scenes in endoscopic videos. *Scientific Reports*, 13, 9235. <https://doi.org/10.1038/s41598-023-36453-1>
- Liu, Z., Luo, P., Wang, X., & Tang, X. (2015). Deep learning face attributes in the wild. En *ICCV 2015* (pp. 3730–3738). IEEE. <https://doi.org/10.1109/ICCV.2015.425>
- Navarro Martínez, O., Fernández-García, D., Cuartero Monteagudo, N., & Forero-Rincón, O. (2024). Possible health benefits and risks of deepfake videos: A qualitative study in nursing students. *Nursing Reports*, 14(4), 2746–2757. <https://doi.org/10.3390/nursrep14040203>
- NHS Lothian. (2023). *Audio and video recordings of psychological therapy sessions policy*. <https://policyonline.nhslothian.scot/wp-content/uploads/2023/03/Audio-and-Video-Recordings-of-Psychological-Therapy-Sessions-Policy.pdf>
- Paris, B., & Donovan, J. (2019). *Deepfakes and cheap fakes: The manipulation of audio and visual evidence*. Data & Society. <https://datasociety.net/library/deepfakes-and-cheap-fakes/>
- Parlamento Europeo & Consejo de la Unión Europea. (2016). Regulation (EU) 2016/679 (GDPR). *Official Journal of the European Union*, L119, 1–88. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32016R0679>
- Pérez, M., & García, J. (2025). *Adaptación y validación de la escala TPACK en la formación de profesores para la educación a distancia*. *Revista de Estilos de Aprendizaje*, 18(35). Recuperado de <https://revistaestilosdeaprendizaje.com/index>
- Pink, S., Horst, H., Postill, J., Hjorth, L., Lewis, T., & Tacchi, J. (2016). *Digital ethnography: Principles and practice*. SAGE.
- Qureshi, J., & Khan, S. (2025). Artificial intelligence deepfakes in healthcare systems: A double-edged sword? *International Journal of Data Science and Big Data Analytics*, 5(1), 84–93. <https://doi.org/10.51483/IJDSBDA.5.1.2025.84-93>
- Reid, A., & Miño, P. (2021). When therapy goes public: Copyright gatekeepers and sharing therapeutic artifacts on social media. *International Journal of Communication*, 15, 950–969. <https://ijoc.org/index.php/ijoc/article/view/1652>

- Sæle, R. G., & Gilbertson, S. (2024). Ethical reflections on social media use within music therapy: A conversational focus group study. *Nordic Journal of Music Therapy*, 33(5), 409–425. <https://doi.org/10.1080/08098131.2024.2350939>
- Shuvi, M., Fish, N., Aberman, K., Shamir, A., & Cohen-Or, D. (2020). *Neural alignment for face de-pixelization* (arXiv:2009.13856). arXiv. <https://doi.org/10.48550/arXiv.2009.13856>
- Sun, J., Yang, J., Zhou, G., Jin, Y., & Gong, J. (2024). Understanding human-AI collaboration in music therapy through co-design with therapists. En *CHI 2024 Proceedings* (Art. 704, pp. 1–21). ACM. <https://doi.org/10.1145/3613904.3642764>
- Syntonym. (n.d.). *Generative AI for privacy! Hyper realistic face anonymization for videos & images*. Syntonym. <https://www.Syntonym.com/>
- Tariq, S., Woo, S. S., Singh, P., Irmalasari, I., Gupta, S., & Gupta, D. (2025). *From prediction to explanation: Multimodal, explainable, and interactive deepfake detection framework for non-expert users* (arXiv:2508.07596) <https://arxiv.org/abs/2508.07596>
- Tiribelli, S., Monnot, A., Shah, S. F. H., Arora, A., Toong, P. J., & Kong, S. (2023). Ethics Principles for Artificial Intelligence-Based Telemedicine for Public Health. *American journal of public health*, 113(5), 577–584. <https://doi.org/10.2105/AJPH.2023.307225>
- Todt, J., Hanisch, S., & Strufe, T. (2024). Fantômas: Understanding face anonymization reversibility. *Proceedings on Privacy Enhancing Technologies*, 2024(4), 24–43. <https://doi.org/10.56553/POPETS-2024-0105>
- Warrier, U., Warrier, A., & Khandelwal, K. (2023). Ethical considerations in the use of artificial intelligence in mental health. *The Egyptian Journal of Neurology, Psychiatry and Neurosurgery*, 59, 139. <https://doi.org/10.1186/s41983-023-00735-2>
- Westerlund, M. (2019). The emergence of deepfake technology: A review. *Technology Innovation Management Review*, 9(11), 39–52. <https://doi.org/10.22215/timreview/1282>
- Yim, D., Khuntia, J., Parameswaran, V., & Meyers, A. (2024). Preliminary evidence of the use of generative AI in health care clinical services: Systematic narrative review. *JMIR Medical Informatics*, 12, e52073. <https://doi.org/10.2196/52073>
- Zhu, Y., Imoussaïne-Aïkous, M., Côté-Lussier, C., and H. Falk, T. (2024). On the Impact of Voice Anonymization on Speech Diagnostic Applications: A Case Study on COVID-19 Detection. *Trans. Info. For. Sec.* 19 5151–5165. <https://doi.org/10.1109/TIFS.2024.3390990>

Financiación

El estudio se realizó con recursos propios y con la colaboración de las plataformas AKOOL Face Swap y Syntonym, que apoyaron la elaboración de los vídeos utilizados.

Agradecimientos

Se agradece la colaboración técnica de AKOOL Face Swap y Syntonym, así como la participación voluntaria de los musicoterapeutas que contribuyeron al desarrollo del estudio piloto.

Conflicto de intereses

Los autores declaran no tener conflicto de intereses con los contenidos de este artículo.

Contribución de autores

La autoría fue compartida equitativamente al 50% cada uno. Ambos participaron en el diseño metodológico, la revisión bibliográfica, la elaboración del material experimental, el análisis de resultados y la redacción final.



© 2026 by the authors. Submitted for open access publication under the terms and conditions of the Creative Commons