

NÚMERO 89 - Septiembre 2024

Sección especial:

INTELIGENCIA ARTIFICIAL EN LA EVALUACIÓN Y LA PERSONALIZACIÓN DEL APRENDIZAJE

Coordinada por:

José Luis Serrano. *Universidad de Murcia (España)*

Juan Moreno García. *Universitat de les Illes Balears (España)*

Publicado: 30-09-2024

DOI: <https://doi.org/10.21556/edutec.2024.89>



Inteligencia artificial y personalización del aprendizaje. Innovación educativa o promesas recicladas _____	2
Diseño y simulación de un modelo de predicción para la evaluación de la competencia digital docente usando técnicas de Machine Learning _____	19
Innovación educativa. Explorando el potencial de la Inteligencia Artificial Generativa en la construcción de esquemas cognitivos ____	45
La evaluación en educación matemática. Aportes de chatbots y futuros profesores de matemática _____	65
Un análisis experimental de la relación entre las evaluaciones proporcionadas por la IA y las proporcionadas por los docentes ____	85
Tutores de IA frente a instructores humanos. Percepciones de los estudiantes de educación superior en Hungría y España _____	106
Impacto de los sistemas de tutoría inteligente. Una revisión sistemática _____	122
Integración de la tecnología y la pedagogía en los sistemas de tutoría inteligente _____	145
Uso de las TIC por alumnado de aula hospitalaria. Un estudio de caso _____	157
Cómo influye el nivel socioeconómico en la autopercepción de la competencia digital del alumnado de educación secundaria en España _____	175
Redes sociales y aprendizaje móvil en Educación Física. Creación colaborativa de principios de diseño _____	191



Inteligencia artificial y personalización del aprendizaje: ¿innovación educativa o promesas recicladas?

Artificial Intelligence and Personalized Learning: Educational Innovation or Revived Promises?

 José Luis Serrano; j.l.serranosanchez@um.es; Universidad de Murcia (España)

 Juan Moreno-García; juan.moreno@uib.es; Universitat de les Illes Balears (España)

Resumen

Este artículo editorial introduce la sección especial titulada "Inteligencia artificial en la evaluación y la personalización del aprendizaje". Se presentan y contrastan las conclusiones de los siete estudios seleccionados en relación con investigaciones recientes. En este artículo se ofrecen cinco principales aportaciones. Primero, se muestran los avances en la integración del aprendizaje adaptativo con la inteligencia artificial generativa para la personalización del aprendizaje. A continuación, se explora el uso educativo de *chatbots*, destacando su capacidad para facilitar experiencias de aprendizaje más dinámicas y ajustadas a las necesidades de los estudiantes. En tercer lugar, se analiza el uso del aprendizaje automático para la creación de modelos predictivos que apoyen la toma de decisiones formativas. Posteriormente, se presentan los desafíos y las oportunidades que presentan los sistemas de tutoría inteligente para proporcionar retroalimentación inmediata y ofrecer recomendaciones para diseñar y ajustar itinerarios personalizados de aprendizaje. Finalmente, se comparten recomendaciones prácticas y reflexiones sobre los desafíos éticos y pedagógicos, la dependencia de la inteligencia artificial y algunos retos que enfrenta la investigación educativa.

Palabras clave: inteligencia artificial; personalización del aprendizaje; sistemas de tutoría inteligente; aprendizaje adaptativo.

Abstract

This editorial article introduces the special section titled "Artificial Intelligence in Assessment and Learning Personalization." The conclusions of the seven selected studies are presented and contrasted with recent research. This article offers five key contributions. First, it showcases advances in the integration of adaptive learning with generative artificial intelligence for learning personalization. Next, the educational use of chatbots is explored, highlighting their ability to facilitate more dynamic learning experiences tailored to students' needs. Third, the use of machine learning in creating predictive models that support formative decision-making is analyzed. Subsequently, the challenges and opportunities presented by intelligent tutoring systems are discussed, particularly in providing immediate feedback and offering recommendations for designing and adjusting personalized learning pathways. Finally, practical recommendations and reflections are shared regarding ethical and pedagogical challenges, the dependence on artificial intelligence, and some of the challenges faced by educational research.

Keywords: artificial intelligence; personalized learning; intelligent tutoring systems; adaptive learning.



1. INTRODUCCIÓN

Cuando en febrero de 2023 comenzamos a diseñar la convocatoria de esta sección especial, nos encontrábamos sumergidos en una avalancha de novedades, eventos, formaciones y publicaciones en medios de comunicación sobre la Inteligencia Artificial (IA), con especial énfasis en la IA Generativa (IAGen). Todo apuntaba que estábamos pasando por uno de esos momentos históricos que provocan cambios profundos en todas las dimensiones de la sociedad al nivel que lo hizo la electricidad a principios del siglo pasado. En el ámbito educativo, por ejemplo, la IA y la IAGen están destinadas a protagonizar la cuarta revolución educativa a través de las tecnologías (Prendes, 2023), con un impacto superior al de revoluciones anteriores, como las impulsadas por el vídeo, Internet y la Web 2.0

Las nuevas herramientas de IA, ya disponibles en ese momento, nos volvieron a invitar a centrarnos más en los medios que en los procesos. El gran error de la tecnología educativa desde sus inicios (Martínez, 2016). De nuevo, teníamos sobre la mesa la teoría de la disrupción, algo que sabemos que no funciona para el sector educativo. Sin embargo, en esta ocasión, además de sucumbir ante los increíbles avances de la IAGen, observamos que esta tecnología tenía grandes posibilidades de adoptarse debido a los cambios en las actitudes y estructuras sociales que la sociedad en su conjunto estaba provocando. Simplemente, era necesario observar cómo las universidades, centros educativos, docentes y la propia UNESCO (2019) expresaban preocupación al mismo tiempo que promovían la formación en el uso educativo de estas herramientas.

Con cautela, examinamos los estudios del momento para abordar los debates emergentes y los vacíos que la investigación educativa sobre IA no había sido capaz de llenar antes de esta nueva generación de IA, liderada por los grandes modelos de lenguaje (LLM por sus siglas en inglés) como ChatGPT de OpenAI. En 2020, Bates et al. concluyeron que, hasta esa fecha, el impacto de la IA en la educación era desconocido o incluso decepcionante, debido a los limitados avances logrados. En la revisión realizada por Zawacki-Richter et al. (2019) se detectó una baja participación de profesionales de la educación en esta línea de investigación. Además, en los estudios se descuidaban los aspectos pedagógicos y éticos de la aplicación de la IA en contextos educativos.

Aunque el número de estudios se ha multiplicado por ocho durante los últimos cinco años, Zawacki-Richter et al. (2024) vuelven a encontrar resultados similares. Las nuevas investigaciones están especialmente enfocadas en evaluar las capacidades de estas herramientas y sus impactos potenciales. Los autores insisten en que la investigación educativa todavía va muy por detrás del desarrollo técnico y metodológico de los algoritmos y de las aplicaciones de IA. De nuevo nos encontramos ante de una de las eternas asignaturas pendientes de la tecnología educativa: que la palabra “educativa” camine más cerca del término “tecnología”. Para lograr esto, necesitamos investigación que informe adecuadamente las prácticas educativas profesionales. La “carrera” entre el desarrollo tecnológico y la investigación en tecnología educativa está en permanente desigualdad.

Con los principios de la tecnología educativa muy presentes y los resultados de los estudios acumulados durante décadas, decimos acotar la sección especial a dos grandes temas sobre los que giraban la mayoría de las promesas con el uso educativo de la IA: la personalización del aprendizaje y la evaluación.

Nos marcamos como objetivo dar a conocer aquellas investigaciones que, desde la tecnología educativa, estudian como la personalización del aprendizaje guiada por la IA puede contribuir a diseñar modelos de enseñanza y situaciones de aprendizaje en contextos educativos formales y no formales. Nuestra perspectiva ha sido, y sigue siendo, apostar por una integración de la IA en la educación que asista y guíe tanto a estudiantes como a educadores en la toma de decisiones para diseñar itinerarios personalizados de aprendizaje (Salinas y De Benito, 2020), además de mejorar los modelos de evaluación (González-Calatayud et al., 2022; Ouyang et al., 2022), considerando la complejidad inherente a la educación. Las líneas de interés que formaron parte de la convocatoria de la sección especial fueron:

- Entornos de aprendizaje adaptativo potenciados por la tecnología.
- IA en la evaluación del aprendizaje.
- Influencia de la IA en el diseño de situaciones de enseñanza-aprendizaje.
- Itinerarios personalizados de aprendizaje asistidos por IA
- Los sistemas de tutoría inteligente.

Este artículo examina los resultados de los siete estudios seleccionados de entre los veinticuatro recibidos. Es importante reconocer que no todas las líneas de interés mencionadas previamente han sido abordadas de manera exhaustiva. Consideramos que esto se debe a la limitada disponibilidad de investigaciones recientes y al breve periodo transcurrido para desarrollar estudios rigurosos sobre los temas propuestos, especialmente en relación con la evaluación con el apoyo de IA. Esta última circunstancia es bien conocida en la investigación sobre tecnología educativa, un tema al que volveremos en la parte final de este artículo. No obstante, los estudios seleccionados representan contribuciones valiosas que proporcionan una base interesante para guiar investigaciones futuras y apoyar la toma de decisiones en contextos educativos.

2. INTELIGENCIA ARTIFICIAL Y PERSONALIZACIÓN DEL APRENDIZAJE

2.1. El primer paso en la personalización del aprendizaje: superar la Ley de la Polka

No está claro si las promesas del pasado de la IA en Educación se están cumpliendo actualmente, si están siendo simplemente recicladas o si nos encontramos ante una moda pasajera. La Ley de la Polka se trata de una de las metáforas que utilizamos en Prendes y Serrano (2016) para señalar los errores habituales en tecnología educativa. En este caso, se trata del error de creer que se descubre algo totalmente revolucionario para enseñar y aprender. Se introducen supuestas novedades que realmente vienen a repetir viejas ideas, obviando la base real y sin fundamento empírico. A pesar de la relevancia y actualidad del tema, la investigación en IA en educación tiene una larga trayectoria que se remonta a la década de los 70 con la promesa de que los tutores inteligentes tendrían el potencial de cambiar la educación gracias a la personalización (Woolf, 2007).

Junto con la evaluación de los aprendizajes en un mundo digital, la personalización del aprendizaje es una de las oportunidades de la IA en educación más señaladas (Li, y Wong, 2023; UNESCO, 2019; Zawacki-Richter et al., 2019). En el contexto educativo, se entiende como la adaptación que combina los objetivos, los intereses y las competencias de un estudiante con un proceso continuo de ajuste a medida que cambian las anteriores condiciones y el contexto (Bulger, 2016). Se trata de respetar los principios fundamentales de la enseñanza y el aprendizaje, facilitando que la mayoría del alumnado pueda seguir su propio itinerario de aprendizaje, considerando sus conocimientos previos, preferencias, motivaciones o metas.

Durante décadas, los sistemas tecnológicos de personalización del aprendizaje buscaron atender las diferencias individuales con bajo coste dentro de un sistema educativo masificado. A lo largo del siglo XX, surgieron teorías y tecnologías como la enseñanza programada y los sistemas ICAI (*Intelligent Computer Assisted Instruction*), que aplicaban IA para tomar decisiones basadas en grandes bases de datos que se enriquecen con el entrenamiento que el usuario hace tras usar la tecnología (Bartolomé y Lindín, 2019).

Hasta hace poco, estas plataformas generaban desconfianza al centrarse casi exclusivamente en mejorar el rendimiento académico. Dejaban de lado aspectos como la autorregulación, la autodeterminación, el bienestar emocional y la dimensión social del aprendizaje, aspectos que sabemos que mejoran la calidad del proceso de enseñanza y aprendizaje (Bulger, 2016).

En 2024, las instituciones educativas, especialmente en la educación superior y online, han comenzado a integrar tecnologías como la evaluación automatizada, reconocimiento facial, chatbots y herramientas predictivas que emplean algoritmos y machine learning. Entre las más prometedoras para la personalización del aprendizaje destacan los tutores y agentes inteligentes, sistemas adaptativos con gran potencial para crear itinerarios personalizados basados en los conocimientos previos, preferencias de los estudiantes y otras variables (Serrano y Sánchez, 2024).

2.2. La alianza entre el aprendizaje adaptativo y la inteligencia artificial generativa

Wang et al. (2023) definen el aprendizaje adaptativo como un enfoque que se ajusta al rendimiento del estudiante mediante el uso de algoritmos avanzados y análisis de datos. Este método modifica continuamente el contenido, la dificultad y los recursos según las interacciones y progresos del estudiante, garantizando -supuestamente- un aprendizaje personalizado que fomenta la eficacia y el desarrollo individual. Guettala et al. (2024) estudiaron los beneficios y desafíos asociados con la integración de la IAGen en los marcos de aprendizaje adaptativo. Antes de verlos, es necesario entender qué hay detrás de esta tecnología revolucionaria.

El desarrollo de las redes neuronales revolucionó la IA clásica al cambiar la forma en que las máquinas aprenden. Pasaron de ser programadas a entrenarse, descubriendo patrones y conexiones que optimizan su funcionamiento. La creación de unidades de procesamiento gráfico (GPU) capaces de realizar múltiples cálculos en paralelo fue un avance técnico crucial. La llegada de Internet ofreció la disponibilidad masiva de datos que la IA necesita para entrenarse de manera más efectiva (Sigman y Bilinkis, 2023). Este contexto nos sitúa ante un tipo de IA que ya hemos mencionado anteriormente: el aprendizaje automático (*machine learning*).

Los modelos generativos son un tipo de modelo dentro de la categoría de aprendizaje automático. Un modelo es un componente central del sistema IA. Es una representación matemática y algorítmica que determina el funcionamiento de este. Para generar un modelo de aprendizaje automático, la IA necesita datos en cantidad y calidad para alimentarse (Bilbao, 2024). De esta manera, tendrá experiencia para mejorar su rendimiento en tareas específicas sin ser explícitamente programada para cada una de esas tareas.

Una de las aplicaciones educativas del aprendizaje automático es el desarrollo de modelos predictivos. En el estudio publicado por Wiston Forero y Francisca Negre en esta sección especial con el título "[Diseño y simulación de un modelo de predicción para la evaluación de la competencia digital docente usando técnicas de Machine Learning](#)", se utiliza la herramienta Orange Data Mining. Los autores concluyen que el modelo es escalable y permite proponer itinerarios personalizados de aprendizaje. Según este trabajo, el modelo utilizado tiene capacidad para orientar la autoevaluación del docente y la evaluación externa, con lo que se guía la toma de decisiones formativas.

Profundizando en los modelos generativos, estos tienen como objetivo principal la creación de nuevos datos y la simulación de escenarios complejos. De esta manera, la IAGen emplea modelos avanzados para crear contenido original en formatos como texto, imágenes, audio y vídeo. A diferencia de los enfoques tradicionales, estos modelos no solo reconocen patrones o siguen reglas, sino que también replican la creatividad humana. Utilizan técnicas de aprendizaje automático -como el aprendizaje profundo y las redes neuronales- para analizar grandes volúmenes de datos. Mediante un entrenamiento intensivo, logran identificar patrones subyacentes y aplican modelos probabilísticos para generar contenido nuevo que sea coherente y adecuado al contexto (Lim et al., 2023; Yusuf et al., 2024).

Tras analizar 407 estudios, Yusuf et al. (2024) concluyeron que la IAGen se define de manera diferente en función desde la disciplina desde la que se estudie:

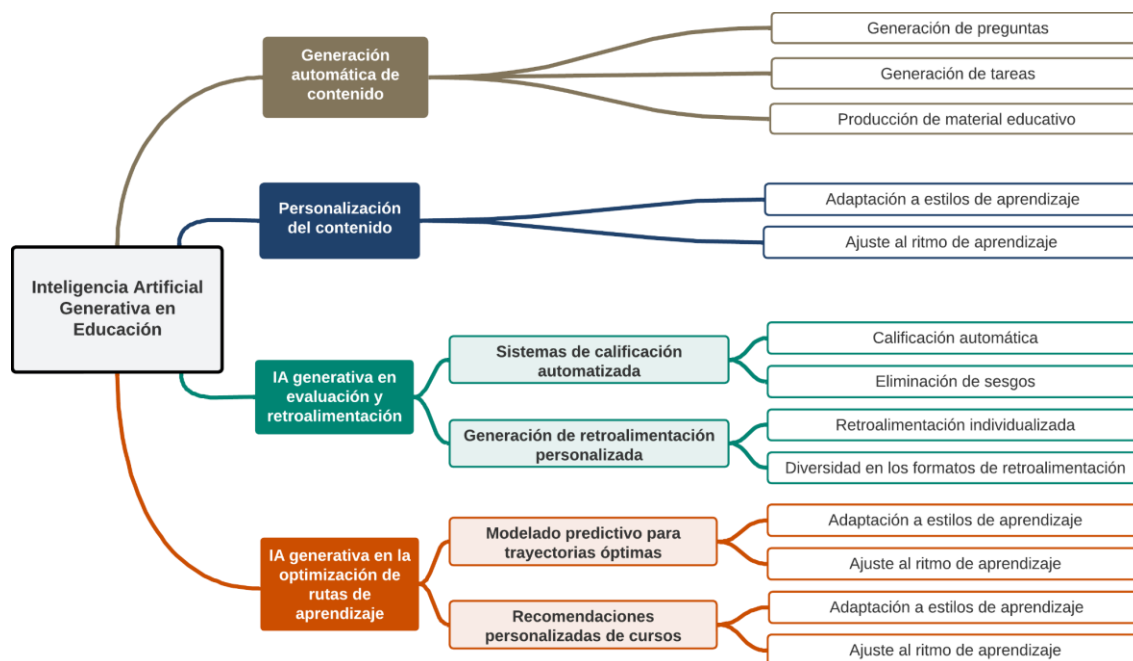
- Como herramienta para la mejora pedagógica. La transformación de las experiencias de enseñanza y aprendizaje es posible siguiendo enfoques personalizados, adaptativos y con contenido interactivo.
- Como herramienta para la formación y la práctica especializada para dar respuesta a necesidades únicas de disciplinas como la medicina o la enfermería mediante el uso de la simulación de casos reales y la toma de decisiones.
- Para la mejora de habilidades profesionales y la toma de decisiones en el campo de la educación financiera, la gestión de simulaciones de planificación estratégica y la formación en estadística y en ciencia de datos.
- Herramienta de escritura y de productividad, abordando necesidades concretas en el uso del lenguaje, la escritura académica o la programación.

Desde un punto de vista optimista, Guettala et al. (2024) revelan en su estudio los beneficios tangibles derivados de la integración de la IAGen: mayor compromiso de los estudiantes, mejores puntuaciones en las pruebas de contenido y un desarrollo acelerado de habilidades. Además, esta tecnología permite cambiar la forma en que se producen los materiales, así como

la evaluación y la entrega de *feedback* al automatizar, personalizar y adaptar en tiempo real los procesos implicados en estas tareas (Figura 1).

Figura 1

Posibilidades de la Inteligencia Artificial Generativa en la educación (Guettala et al., 2024).



La integración de la IAGen con otras plataformas y sistemas permite optimizar su funcionamiento y ofrecer nuevas funcionalidades. Plataformas como Duolingo¹ o Knewton² facilitan tareas administrativas y, sobre todo, se centran en la adaptación de contenidos y lecciones. Herramientas como SchoolAI³ y MagicSchool⁴ no solo cumplen con estas funciones, sino que también intentan ir un paso más allá al ofrecer apoyo emocional y fomentar la interacción humana. Por otro lado, Chatmind⁵ y Miro⁶ permiten la creación de representaciones visuales del conocimiento, lo que facilita la comprensión de conceptos complejos y fomenta la colaboración.

Esta evolución de las herramientas implica la combinación del aprendizaje adaptativo con la IAGen. Según Guettala et al. (2024), esto podría revolucionar la educación siempre y cuando se enfatice la importancia del aprendizaje socioemocional y la conexión humana, junto con la personalización. El objetivo debe ser integrar tecnologías avanzadas con un enfoque en el

¹ Duolingo: <https://es.duolingo.com/>

² Knewton: <https://www.knewton.com/>

³ SchoolAI: <https://schoolai.com/>

⁴ MagicSchool: <https://www.magicschool.ai/>

⁵ Chatmind: <https://chatmind.es/>

⁶ Miro: <https://miro.com/es/>

bienestar integral del estudiante, promoviendo un aprendizaje más inclusivo y humano, adaptado tanto a las capacidades cognitivas como emocionales de los estudiantes.

Aunque el número de estudios sobre IAGen en educación va en aumento, todavía no contamos con una investigación consolidada que nos permita apoyar con mayor seguridad la toma de decisiones en su integración educativa. Además, es realmente en educación superior donde más estudios podemos encontrar. En el otro extremo, la etapa de educación primaria ha recibido escasa atención en la investigación sobre IAGen. Problemas de infraestructura, accesibilidad, contenidos apropiados para la edad y retrasos en su integración pueden ser las causas según la revisión realizada por Yusuf et al. (2024).

Lo anterior no significa que dejemos de explorar las mejoras que esta tecnología puede aportar en la enseñanza y el aprendizaje, sino que debemos ser prudentes en su integración y evaluar su impacto en diferentes contextos.

2.3. Chatbots educativos: De ELIZA a ChatGPT

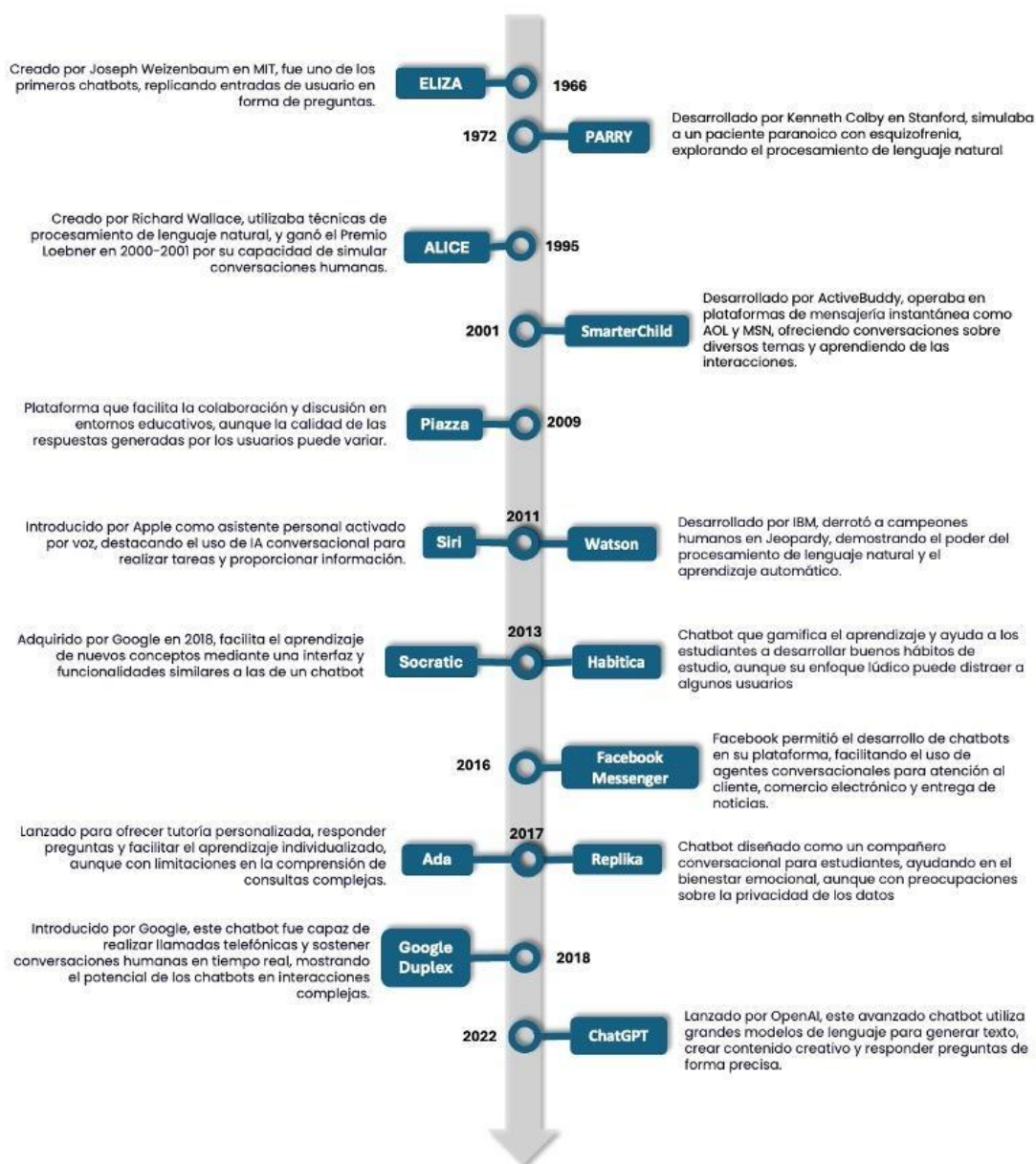
La aparición de los *chatbots* educativos se remonta a la década de los 60. ELIZA, desarrollado en 1966 por Joseph Weizenbaum en el MIT, es considerado uno de los primeros programas de IA. Era capaz de simular una conversación con un psicoterapeuta. Utilizaba técnicas simples de procesamiento de lenguaje natural para replicar entradas del usuario en forma de preguntas o respuestas predeterminadas. ELIZA demostró cómo un programa básico podía generar la ilusión de comprensión, influyendo en el desarrollo de la IA conversacional y en la percepción pública sobre la IA.

En las décadas siguientes, los *chatbots* avanzaron gracias a algoritmos más sofisticados y mayor capacidad computacional. En los años 90, proyectos como ALICE mejoraron la coherencia y relevancia de las respuestas, aunque aún carecían de comprensión profunda del contexto.

La transformación más significativa llegó con aparición de las redes neuronales profundas y los modelos de aprendizaje automático avanzados. ChatGPT, lanzado por OpenAI en 2022, representa un hito en esta evolución. Basado en modelos de lenguaje de gran escala entrenados con enormes conjuntos de datos textuales, ChatGPT puede comprender y generar texto de manera coherente y contextualizada. Esto le permite mantener conversaciones fluidas, responder preguntas complejas y adaptarse al contexto, ampliando significativamente sus aplicaciones en el ámbito educativo.

Figura 2

Evolución de Chatbots (1966-2022)



Nota. Adaptado de (Labadze et al., 2023).

El impacto potencial en la educación futura de los *chatbots* es sustancial. Pueden actuar como asistentes virtuales, mejorando la calidad de la enseñanza y ofreciendo experiencias de aprendizaje personalizadas. Así lo indican Labadze et al. en su revisión sistemática publicada en 2023. El uso de chatbots en educación ha demostrado mejorar los resultados de aprendizaje, razonamiento explícito y retención del conocimiento. Los *chatbots* ofrecen beneficios como asistencia inmediata, acceso rápido a información y una mejora en la experiencia educativa. Sin embargo, los efectos sobre la motivación, el compromiso y el pensamiento crítico generan resultados contradictorios.

Labadze et al. (2023) concluyen que la adopción de *chatbots* en educación genera opiniones divididas: los estudiantes los reciben con entusiasmo, mientras que los educadores son más críticos. Los *chatbots* ofrecen ventajas como el ahorro de tiempo en tareas administrativas, apoyo personalizado y permiten a los docentes centrarse en adaptar el contenido educativo. Para los estudiantes, proporcionan asistencia en tareas, retroalimentación en tiempo real y promueven el aprendizaje autónomo, desarrollando habilidades como la escritura y la resolución de problemas, mejorando así su motivación y compromiso.

También presentan capacidad para mejorar el desarrollo de esquemas cognitivos, fundamentales para que un aprendiz conecte la nueva información con los conocimientos previos. De esta manera se fomenta una comprensión más profunda, clave en el aprendizaje y en la retención de conocimientos. A estas conclusiones llegan los autores procedentes de tres instituciones ecuatorianas: Bernarda Salgado (Universidad Politécnica Salesiana), Yana Inzhivotkina (Universidad de Guayaquil), María Fernanda Ibáñez (Universidad Estatal Península de Santa Elena) y Jorge Gustavo Ugate (Universidad Ecotec). En este estudio titulado "[Educational Innovation: Exploring the Potential of Generative Artificial Intelligence in Cognitive Schema Building](#)", se indica que la IAGen es especialmente útil en la personalización del contenido educativo, la creación de entornos de aprendizaje inmersivos, y la automatización del feedback en procesos de evaluación. Esto se traduce en una mayor satisfacción y rendimiento cognitivo de los estudiantes, impulsando su motivación y participación en el proceso de aprendizaje. Sin embargo, el artículo también reconoce los desafíos que plantea la implementación de esta tecnología, como la generación de contenido de baja calidad, preocupaciones sobre la privacidad de los datos y la posibilidad de reforzar sesgos existentes en el material educativo.

En el trabajo de Guan et al. (2024) se concluye que los *chatbots* educativos ayudan a los estudiantes a identificar recursos de aprendizaje, implementar estrategias de aprendizaje adecuadas y monitorización de metas. En esta investigación se analizó cómo estas tecnologías apoyan el aprendizaje autorregulado. Sin embargo, los autores indicaron que ningún *chatbot* estudiado ofrece apoyo completo en las cuatro fases del modelo SRL de Winne y Hadwin (1998). Dicho apoyo se limita a tareas específicas, lo que reduce el impacto en el rendimiento académico.

ChatGPT de OpenAI es protagonista en múltiples estudios recientes adoptando el rol de *chatbot* educativo. En el estudio de Thorat et al (2024) se concluye que ChatGPT puede personalizar el aprendizaje al adaptar el contenido según las preferencias, estilos y ritmo de aprendizaje de cada estudiante. Ofrece experiencias autónomas, permitiendo que los estudiantes avancen a su propio ritmo, profundicen en temas de interés o revisen conceptos difíciles. Proporciona retroalimentación personalizada y recursos adicionales según las necesidades individuales. También puede realizar un seguimiento continuo del progreso del estudiante, ajustando dinámicamente los itinerarios personalizados de aprendizaje.

En la revisión realizada por Yusuf et al. (2024), se encontraron investigaciones que indican que ChatGPT era útil en la ayuda para el estudio o en la evaluación automatizada de la escritura, aunque en este último caso se identificaron diferencias significativas en la cantidad y el tipo de feedback proporcionado entre ChatGPT y las evaluaciones de los docentes.

En el estudio "[La evaluación en educación matemática: aportes de chatbots y futuros profesores de matemática](#)", presentado en esta sección especial firmado por tres autoras de NIEM- CONICET-UNCPBA (Argentina), Patricia Sureda, Ana Corica y Verónica Parra, y Daniela Gody y Silvia Schiaffino de ISITAN - UNCPBA/CONICET (Argentina), se concluye que los *chatbots* resultan asistentes valiosos para ofrecer diferentes tipos de evaluaciones en el área de matemáticas en Secundaria. Desde las más tradicionales, como una prueba escrita, a otras menos habituales, como un proyecto de investigación. Gemini es señalado en este estudio el *chatbot* más adecuado (por delante de ChatGPT y Copilot) como asistente en el proceso de diseño del proceso evaluativo.

El principal objetivo de otro trabajo presentado en esta sección especial titulado "[Un análisis experimental de la relación entre las evaluaciones proporcionadas por la inteligencia artificial y las proporcionadas por los docentes en formación](#)", es determinar si las evaluaciones proporcionadas generadas por la IA en trabajos escritos son coherentes con las realizadas por docentes en formación, y si factores como el género, el nivel de formación o el rendimiento académico influyen en dichas evaluaciones. Este estudio ha sido realizado por tres autores de la Universidad del País Vasco (España): Héctor Galindo, Nahia Delgado y Martín Sainz, junto con Ernesto Expósito, de la Université de Pau et des Pays de l'Adour (Francia). Los autores indican que ChatGPT es la herramienta que replica en gran medida los patrones de evaluación de los docentes en formación inicial. Otra conclusión interesante de esta investigación es que las evaluaciones realizadas por los docentes en formación con mejor desempeño estuvieron más alineadas con las proporcionadas por la IAGen en comparación con los estudiantes con menor desempeño.

Por último, señalar que en el estudio de Stadler et al. (2024) se analiza la carga cognitiva de los estudiantes al comparar el uso de ChatGPT con los motores de búsqueda tradicionales. Los autores concluyen que, aunque los LLM como ChatGPT ofrecen una forma eficaz de reducir la carga cognitiva intrínseca y externa, es posible que no siempre faciliten el aprendizaje en profundidad, necesario para tareas complejas de toma de decisiones. Los motores de búsqueda tradicionales, al requerir un compromiso más activo, pueden fomentar un aprendizaje de mayor calidad, lo que subraya la necesidad de prácticas educativas que fomenten el compromiso crítico con diversas fuentes de información.

2.4. Sistemas de tutoría inteligente: la gran promesa

Mediante el uso de algoritmos de aprendizaje automático para analizar gran cantidad de datos de los estudiantes, los sistemas de tutoría inteligente pretenden simular la tutoría humana. Estas tecnologías están evolucionando en los últimos años, pasando de solamente proporcionar retroalimentación a convertirse en entornos de aprendizaje complejos. Están llamados a ser los sistemas adaptativos que mayor potencial ofrecen para personalizar el aprendizaje y mejorar la eficacia del aprendizaje. Incluso, podrán convertirse en una guía de aprendizaje proactiva que no solo sirva para asistir, puesto que buscarán orientar e inspirar más allá de un árbol de decisiones que proporcione respuestas limitadas. Khanamigo⁷ es un ejemplo de tutor inteligente desarrollado por la Khan Academy y basado en GPT-4.

⁷ <https://www.khanmigo.ai/>

Existen indicios de que estos sistemas promoverán el diseño de itinerarios personalizados de aprendizaje teniendo en cuenta no solo los conocimientos previos de los estudiantes y sus preferencias, sino también un amplio conjunto de datos procedentes de distintas variables (Serrano y Sánchez, 2024). Sin embargo, los métodos de toma de decisiones de estas rutas de aprendizaje personalizadas siguen teniendo deficiencias en la predicción de la dificultad de los puntos de conocimiento, tal y como señala Zhang (2024). El autor, tras implementar un método de toma de decisiones de itinerarios de aprendizaje personalizados basado en el aprendizaje automático, concluye que en futuras investigaciones se debe mejorar tanto en aspectos tecnológicos como pedagógicos. Los tres grandes objetivos a perseguir son:

- Optimizar la estructura y el algoritmo del modelo para mejorar la eficiencia computacional con datos a gran escala.
- Considerar de manera integral los factores motivacionales de los estudiantes y los estados emocionales.
- Fortalecer las técnicas de preprocesamiento y limpieza de datos para garantizar la calidad y confiabilidad de los datos de entrenamiento.

Las emociones juegan un papel crucial en el proceso de aprendizaje. En un estudio reciente (Zheng et al., 2024) los autores examinaron la dinámica emocional durante las fases del aprendizaje autorregulado (planificación, ejecución y autorreflexión) utilizando un sistema de tutoría inteligente. Los investigadores señalan que los estudiantes de medicina con mejor rendimiento que participaron en el estudio demostraron emociones más estables en la fase de planificación. Y menos estables en la de autorreflexión, pero con una mayor excitación o activación emocional. Como vemos, esta tecnología también es utilizada como recurso e instrumento para investigar otros temas por su gran capacidad de recopilar y analizar datos.

La falta de empatía humana y apoyo emocional sigue siendo una preocupación importante. Así se señala también en el trabajo publicado en esta sección especial por Admad Hajeer, Árpád Papp-Váry y Éva Pólya de la Budapest Business University (Hungría) y que lleva por título "[AI Tutors vs. Human Instructors: Perceptions of Higher Education Students in Hungary and Spain](#)". Los autores indican que la familiaridad tecnológica, la utilidad percibida y la capacidad de brindar experiencias de aprendizaje personalizadas son factores que determinan las expectativas y percepciones de los estudiantes sobre los tutores de IA. Además, el estudio concluye que estas tecnologías destacan en proporcionar aprendizaje personalizado y adaptativo, pero, no brindan el apoyo emocional y empático que ofrecen los tutores humanos.

Retomando los desafíos, encontramos en esta sección especial dos trabajos que señalan aspectos pedagógicos pendientes de estos sistemas. En el trabajo "[Impacto de los sistemas de tutoría inteligente](#)" presentado por Noelia Carbonell (Universidad Internacional de La Rioja) y M^a Ángeles Hernández (Universidad de Murcia), se analiza el impacto de los sistemas de tutoría inteligente en el proceso de enseñanza-aprendizaje durante la última década. Las autoras señalan un crecimiento en el número de estudios, especialmente tras la pandemia de COVID-19. En esta revisión se concluye que, a pesar del potencial transformador educativo que tienen estos sistemas, es necesario que los docentes tengan una preparación más exhaustiva que permita poder realizar una reflexión crítica sobre su aplicabilidad.

Por su parte, el estudio "[Integración de la tecnología y la pedagogía en los sistemas de tutoría inteligente](#)" publicado por Geovanny Francisco Ruiz de la Universidad de Guayaquil, parte de la premisa de que la integración efectiva de estos sistemas sigue siendo un desafío. Coincidiendo con estudios previos, en este trabajo se subraya la importancia de desarrollar modelos pedagógicos basados en enfoques teóricos como el constructivismo, el aprendizaje por descubrimiento guiado y el aprendizaje situado. Además, el autor indica la relevancia que tiene proporcionar una retroalimentación oportuna, específica y adaptativa, además de proporcionar pistas y sugerencias para la guía del proceso de aprendizaje.

3. LA COMPLEJIDAD DE LA EDUCACIÓN, LÍMITES ÉTICOS Y EL EXCESO DE VELOCIDAD

Concluimos este artículo reflexionando sobre tres cuestiones. Mostramos también algunas recomendaciones prácticas e inquietudes que son de utilidad tanto para guiar futuros estudios como para apoyar la toma de decisiones durante la integración educativa de la IA.

La primera viene determinada por la naturaleza compleja de la educación. El proceso de enseñanza y aprendizaje es un sistema complejo compuesto por innumerables elementos, lo que hace que el conjunto sea mucho más que la simple suma de sus partes. Estos elementos se interrelacionan de manera constante, operando en conjunto para resolver problemas y alcanzar objetivos. La personalización del aprendizaje mediante IA promete ofrecer al estudiante lo que necesita en cada momento, considerando variables como sus intereses, conocimientos previos y su rendimiento anterior, o incluso comparándolo con el rendimiento de estudiantes anteriores. Pero ¿tenemos realmente claro qué datos necesitamos? ¿Sabemos con precisión qué variables influyen en el aprendizaje del estudiante en diferentes momentos? ¿Podemos cuantificarlas?

Serrano y Sánchez (2024) reflexionan sobre la simplificación de la educación al reducirla a los contenidos y al rendimiento académico. Ignorar la dimensión social del aprendizaje, el bienestar subjetivo o el estado emocional de los estudiantes es un grave error que nos lleva a medir lo irrelevante simplemente porque no podemos cuantificar lo que resulta difícil o imposible de medir (Bartolomé et al., 2020). Aunque la IA sigue sorprendiéndonos con su capacidad para recopilar y analizar datos, aún no podemos garantizar que sea capaz de identificar con precisión lo que cada estudiante necesita en cada momento. También es pertinente tener presente que ningún docente es capaz de alcanzar tal fin.

Estudios citados en este artículo (como el de Thorat et al., 2024) concluyen que, de entre las principales limitaciones de ChatGPT en la personalización del aprendizaje, se encuentra su incapacidad para entender plenamente las necesidades emocionales y preferencias individuales de cada estudiante. En cualquier caso, no deberíamos esperar que esto ocurra. Será más beneficioso que nos preparemos como formadores para interpretar los datos que la IA puede proporcionar, comprendiendo los patrones, relaciones y predicciones que pueda generar, pero siempre viéndola como una asistente. La decisión final sobre los itinerarios de aprendizaje personalizados debe recaer siempre en el docente y en colaboración con sus estudiantes.

En los entornos apoyados por IA, es crucial que los estudiantes fortalezcan su capacidad de autorregulación de aprendizaje, lo que les permitirá gestionar sus pensamientos, acciones, emociones y motivación para alcanzar los objetivos planteados (Panadero y Alonso-Tapia, 2014). Aunque la IA ofrecerá retroalimentación y personalización, la responsabilidad última del aprendizaje seguirá recayendo en el propio estudiante.

Un segundo punto recurrente en los estudios seleccionados para esta sección especial y una parte importante de otras investigaciones también incluidas en este artículo, tiene que ver con los desafíos éticos con la IA. Afrontar los desafíos pedagógicos y técnicos con un enfoque humano se ha convertido en una preocupación ética compartida entre educadores, tecnólogos educativos, políticos, filósofos, sociólogos... Se destaca la relevancia de emplear las herramientas de IA de manera responsable, respetando las normas de integridad académica y atendiendo a cuestiones como la privacidad de los datos, su fiabilidad, la transparencia o los derechos de autor.

Tal y como apunta Cordón (2023), la madurez de la IA como tecnología requiere marcos legales que garanticen su uso responsable, justo, inclusivo, confiable, seguro y transparente. Es esencial que la sociedad en su conjunto comprenda sus limitaciones y beneficios para asegurar que su desarrollo sea positivo. El autor señala que estamos viviendo un cambio profundo que nos llevará a convivir con sistemas inteligentes que ayudarán en la toma de decisiones y la automatización de tareas humanas. En la educación, se vislumbra un híbrido “humano + IA”, donde las personas aportamos pensamiento crítico y creatividad, mientras la IA apoya en el diseño del aprendizaje y la evaluación, ofreciendo un soporte clave en estos procesos.

Las instituciones educativas tienen la responsabilidad de crear marcos legales, conciencia entre los estudiantes y ofrecer formación de calidad al profesorado (Labadze et al., 2023). Estudios como el de Dúo et al. (2023) muestran una de las clásicas lecciones de la investigación en tecnología educativa ante la llegada de nuevas tecnologías: los docentes con mayor experiencia previa en su uso (en este caso, en la IA) valoran positivamente su integración en el aula. Una vez formados los docentes, estos pueden ayudar a los estudiantes a que utilicen la IA de manera respetuosa y eficiente con estas cinco ideas propuesta por Grané (2024):

- Establecer reglas conjuntas con el alumnado. Estas deben entenderse no como prohibiciones, sino como una política de uso que nos protege de la búsqueda de soluciones rápidas que no mejoran nuestras habilidades y nuestra manera de pensar. Es interesante que estas reglas sean debatidas y negociadas para aumentar su adherencia y compromiso. Es un excelente ejercicio para discutir aspectos éticos y filosóficos que nos permitan tener el control sobre esta tecnología y evitar la dependencia.
- Crear contextos para experimentar con la IA y la IAGen. Estos escenarios controlados permiten que los errores sean convertidos en aprendizajes con un bajo riesgo. Los problemas que surjan durante las actividades creativas y compartidas deben ser aprovechados siguiendo los principios de la evaluación formativa.
- Cuestionar la IA y la IAGen. Generar contenido o ser más productivos sin mejorar nuestro pensamiento y habilidades es uno de los grandes riesgos del uso de estas tecnologías, especialmente en los periodos formativos más críticos. Estas herramientas nos permiten llegar más lejos y más rápido, pero si dejamos de esforzarnos en pensar y ser críticos ante la información que genera, nuestras habilidades quedan muy comprometidas, dando lugar

al sedentarismo cognitivo (Sigman y Bilinkis, 2023). Los estudiantes deben aprender a diferenciar las tareas que son críticas para el correcto desarrollo de habilidades y competencias de las que no lo son.

- Ofrecer estrategias de verificación de fuentes e información. La búsqueda de la verdad requiere triangular información desde distintas fuentes. Percibir la IAGen como un oráculo impide profundizar en contenidos desde diversas fuentes. Los estudiantes deben desarrollar su competencia informacional para acceder a un conocimiento fiable con el que hacerse preguntas y fundamentar nuevos argumentos.
- Crear oportunidades para pensar críticamente sobre la información obtenida. Para ser creativo o crítico se requiere tener un conocimiento basado en información validada. Los estudiantes no podrán cuestionar el contenido generado por la IAGen si no lo dominan antes. Deben aprender primero a conducir su gestión de la información y el conocimiento y apoyarse después en esta tecnología para detectar patrones y relaciones que solos son - probablemente- incapaces (Serrano, 2024). Finalmente, tomarán la última decisión contrastando la información generada.

A los cinco puntos anteriores, añadimos una reflexión sobre la posible dependencia que podemos desarrollar hacia el uso de IA. No es un asunto que tenga relación directa con los estudios seleccionados en la sección especial de este número. Sin embargo, consideramos que debe formar parte en los temas a tener en cuenta durante la planificación de la integración educativa de la IA para personalizar el aprendizaje y en cómo ayudamos a los estudiantes a utilizarla.

Debemos analizar cuidadosamente qué habilidades cognitivas hay que proteger del uso de la IA y en las circunstancias que esta puede perjudicarnos más. Se trata de un asunto complejo que comienza a generar interés desde la investigación. En un estudio reciente (Zhang et al., 2024) se concluye que los estudiantes universitarios que tienen mayores expectativas de rendimiento académico y que obtienen mejores resultados, tienden a depender más de herramientas de IA. Los autores elaboraron un listado de las diez consecuencias de esta dependencia. Las tres más destacadas son: el aumento de la pereza, la creatividad restringida y el aumento de la información correcta.

La tercera -y última- cuestión que abordamos para cerrar este artículo de presentación de la sección especial de este número se relaciona con otra de las metáforas que utilizamos en Prendes y Serrano (2016): la ley de la caducidad. Los resultados de las investigaciones en el ámbito de las tecnologías están inevitablemente afectados por su naturaleza temporal y efímera. Esto se debe a la rápida evolución y obsolescencia de las tecnologías, provocando que hallazgos que parecen relevantes en el presente puedan perder su vigencia en poco tiempo. Tanto instituciones que financian estudios, investigadores como editores de revistas académicas deben reflexionar sobre qué tipo de proyectos se están llevando a cabo y cómo se enfocan. Si un estudio se centra exclusivamente en una tecnología específica, es probable que sus resultados queden obsoletos mucho antes de lo previsto, limitando su impacto y relevancia a largo plazo. Tener en cuenta esta circunstancia con tecnologías IA es realmente crucial durante los diseños de nuevas investigaciones.

4. AGRADECIMIENTOS

Finalizamos este artículo agradeciendo a los autores que contribuyeron a esta sección especial, así como a los revisores que participaron en la evaluación de los trabajos publicados. También expresamos nuestro agradecimiento al equipo editorial de la Revista Edutec por la confianza depositada. Esperamos que los lectores encuentren en los artículos seleccionados contenido valioso que informe la toma de decisiones en su práctica educativa, que inspire nuevas investigaciones y que contribuya a reducir -un poco- la brecha entre los términos "tecnología" y "educativa". Brecha que la IA podría haber ya aumentado de manera considerable.

5. REFERENCIAS

- Bartolomé, A., y Lindín, C. (2019). Posibilidades del Blockchain en Educación. *Education in the Knowledge Society*, 19, 81-93 (4). <https://doi.org/10.14201/eks20181948193>
- Bates, T., Cobo, C., Mariño, O., y Wheeler, S. (2020). Can artificial intelligence transform higher education? *International Journal of Educational Technology in Higher Education*, 17(1). <https://doi.org/10.1186/s41239-020-00218-x>
- Bilbao, A. (2024). Entonces, ¿qué necesito saber como docente sobre la IA. En Arroyo, A. (Coord.), *Inteligencia Artificial y educación: construyendo puentes* (pp.29-40). Graó.
- Bulger, M. (2016). Personalized Learning: The Conversations We're Not Having. *Data & Society*. <https://datasociety.net/library/personalized-learning-the-conversations-were-not-having/>
- Cordón, O. (2023). Inteligencia Artificial en Educación Superior: Oportunidades y Riesgos. *RiiTE Revista interuniversitaria de investigación en tecnología educativa*, 15, 16-27. <https://doi.org/10.6018/riite.591581>
- Cordón, O. (2023). Inteligencia Artificial en Educación Superior: Oportunidades y Riesgos. *RiiTE Revista interuniversitaria de investigación en tecnología educativa*, 15, 16-27. <https://doi.org/10.6018/riite.591581>
- Dúo, P., Moreno, A.J., López, J., y Marín, J.A. (2023). Inteligencia Artificial y Machine Learning como recurso educativo desde la perspectiva de docentes en distintas etapas educativas no universitarias. *RiiTE Revista interuniversitaria de investigación en tecnología educativa*, 15, 58-78. <https://doi.org/10.6018/riite.579611>
- Dúo, P., Moreno, A.J., López, J., y Marín, J.A. (2023). Inteligencia Artificial y Machine Learning como recurso educativo desde la perspectiva de docentes en distintas etapas educativas no universitarias. *RiiTE Revista interuniversitaria de investigación en tecnología educativa*, 15, 58-78. <https://doi.org/10.6018/riite.579611>
- González-Calatayud, V., Prendes-Espinosa P. y Roig-Vila, R. (2021). Artificial Intelligence for Student Assessment: A Systematic Review. *Applied Sciences* 11 (12). <https://doi.org/10.3390/app11125467>
- Grané, M. (2024). Empecemos por el centro, ¿qué opina el alumnado sobre la inteligencia artificial? En Arroyo, A. (Coord.), *Inteligencia Artificial y educación: construyendo puentes* (pp.15-28). Graó.

- Guan, R., Raković, M., Chen, G., y Gašević, D (2024). How educational chatbots support self-regulated learning? A systematic review of the literature. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-024-12881-y>
- Guettala, M., Bourekache, S., Kazar, O., y Harous, S. (2024). Generative artificial intelligence in education: Advancing adaptive and personalized learning. *Acta Informatica Pragensia*, 13(3), 460-489. <https://doi.org/10.18267/j.aip.235>
- Labadze, L., Grigolia, M. i Machaidze, L. (2023). Role of AI chatbots in education: systematic literature review. *International Journal of Educational Technology in Higher Education* 20 (56). <https://doi.org/10.1186/s41239-023-00426-1>
- Li, K., y Wong, B. (2023). *Artificial intelligence in personalised learning: a bibliometric analysis*. *Interactive Technology and Smart Education*, 20(3), 422-445. <https://doi.org/10.1108/ITSE-01-2023-0007>
- Li, K.C. y Wong, B.T.-M. (2023), Artificial intelligence in personalised learning: a bibliometric analysis, *Interactive Technology and Smart Education*, 20(3), pp. 422-445. <https://doi.org/10.1108/ITSE-01-2023-0007>
- Lim, W., Gunasekara, A., Leight, J., Pallant, J.I. y Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The International Journal of Management Education*, 21 (2). <https://doi.org/10.1016/j.ijme.2023.100790>
- Martínez, F. (2016). Sentado en el andén. *RiiTE Revista interuniversitaria de investigación en tecnología educativa*, 0. 17-22. <https://doi.org/10.6018/riite/2016/258131>
- Ouyang, F., Zheng, L. y Jiao, P. (2022). Artificial intelligence in online higher education: A systematic review of empirical research from 2011 to 2020. *Education and Information Technologies*, 27. <https://doi.org/10.1007/s10639-022-10925-9>
- Panadero, E. y Alonso-Tapia, J. (2014). ¿Cómo autorregulan nuestros alumnos? Revisión del modelo cíclico de Zimmerman sobre autorregulación del aprendizaje. *Anales de Psicología*, 30(2). <https://doi.org/10.6018/analesps.30.2.167221>
- Prendes, M.P. y Serrano, J.L. (2016). En busca de la Tecnología Educativa: la disrupción desde los márgenes. *RiiTE Revista Interuniversitaria de Investigación en Tecnología Educativa*, 0, 6-16. <http://dx.doi.org/10.6018/riite/2016/263771>
- Salinas, J. y De Benito, B. (2020). Construction of personalized learning pathways through mixed methods. *Comunicar*, 65. <https://doi.org/10.3916/C65-2020-03>
- Serrano, J.L. (2023). Cómo los gestores digitales pueden convertir nuestras notas en conocimiento. *The Conversation*. <https://tinyurl.com/2d3xp6ar>
- Serrano, J.L., y Sánchez, M.M. (2024). ¿A qué promesas y desafíos me enfrento como docente con la IA? En Arroyo, A. (Coord.), *Inteligencia Artificial y educación: construyendo puentes* (pp.57-70). Graó.
- Sigman, M., y Bilinkis, S. (2023). Artificial: La nueva inteligencia y el contorno de lo humano. Debate.

- Thorat, V. A., Rao, P., Joshi, N., Talreja, P., y Shetty, A. (2024). The Role of Chatbot GPT Technology in Undergraduate Dental Education. *Cureus* 16(2). <https://doi.org/10.7759/cureus.54193>
- UNESCO (2019). Beijing Consensus on Artificial Intelligence and Education. <https://unesdoc.unesco.org/ark:/48223/pf0000368303>
- Winne, P. H., y Hadwin, A. F. (1998). Studying as self-regulated engagement in learning. in metacognition in educational theory and practice. *Metacognition in Educational Theory and Practice*, pp. 277–304.
- Woolf, B.P. (2007). Building Intelligent Interactive Tutors. Student-centered strategies for revolutionizing e-learning. Elsevier. <https://doi.org/10.1016/B978-0-12-373594-2.X0001-9>
- Yusuf, A., Pervin, N., Román-González, M., y Noor, N. M. (2024). Generative AI in education and research: A systematic mapping review. *Review of Education*, 12(2). <https://doi.org/10.1002/rev3.3489>
- Yusuf, A., Pervin, N., Román-González, M., y Noor, N. M. (2024). Generative AI in education and research: A systematic mapping review. *Review of Education*, 12(2). <https://doi.org/10.1002/rev3.3489>
- Zawacki-Richter, O., Marín, V. I., Bond, M., y Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education– where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1). <https://doi.org/10.1186/s41239-019-0171-0>
- Zhang, S., Zhao, X., Zhou, T., y Kim, J. H. (2024). Do you have AI dependency? The roles of academic self-efficacy, academic stress, and performance expectations on problematic AI usage behavior. *International Journal of Educational Technology in Higher Education*, 21(1), 34. <https://doi.org/10.1186/s41239-024-00467-0>
- Zheng, J., Li, S., Wang, T., y Lajoie, S.P. (2024). Unveiling emotion dynamics in problem-solving: A comprehensive analysis with an intelligent tutoring system using facial expressions and electrodermal activities. *International Journal of Educational Technology in Higher Education*, 21(33). <https://doi.org/10.1186/s41239-024-00462-5>

Para citar este artículo:


Serrano, J. L., y Moreno-García, J. (2024). Inteligencia artificial y personalización del aprendizaje: ¿innovación educativa o promesas recicladas?. *EduTec, Revista Electrónica de Tecnología Educativa*, (89), 1-17. <https://doi.org/10.21556/edutec.2024.89.3577>



Diseño y simulación de un modelo de predicción para la evaluación de la competencia digital docente usando técnicas de Machine Learning

Design and simulation of a predictive model for the evaluation of teachers' digital competence using Machine Learning techniques

 Wiston Forero-Corba; wiston.forero@uib.es

 Francisca Negre Bennasar; xisca.negre@uib.es

Universitat de les Illes Balears (España)

Resumen

Machine Learning (ML) es un campo de la inteligencia artificial que, a través de técnicas, elabora predicciones de datos masivos. La competencia digital docente (CDD) refiere comúnmente a las habilidades y destrezas de los docentes en sistemas digitales y su aplicación en los procesos de enseñanza-aprendizaje. La investigación sobre CDD es importante para las instituciones, ya que de su evaluación dependen el aprendizaje, trayectoria, dirección y comportamiento de los alumnos. La CDD en Colombia se basa en 5 elementos: Comunicativa, de gestión, investigativa, pedagógica y tecnológica, y cada uno de ellos se mide en tres niveles: Explorador, integrador e innovador. Las preguntas de investigación fueron: (1) ¿Qué tipo de resultados podemos esperar de la predicción de la CDD con técnicas de ML? (2) ¿Qué técnicas de ML son efectivas para predecir la CDD? (3) ¿Qué ventajas trae predecir la CDD con técnicas de ML? La metodología pretende diseñar un modelo de predicción de la CDD en Colombia aplicando 9 técnicas de ML usando el software *Orange Data Mining*. Los resultados muestran la alta efectividad que tienen las técnicas inteligentes para predecir la CDD. El modelo muestra que es retroalimentable, escalable y permite proponer itinerarios personalizados de aprendizaje.

Palabras clave: competencia digital docente, machine learning, inteligencia artificial, aprendizaje adaptativo, tecnología emergente

Abstract

Machine Learning (ML) is a field of artificial intelligence that uses techniques to make predictions from massive data. Teachers' Digital Competence (TDC) commonly refers to teachers' skills and abilities in digital systems and their application in teaching and learning processes. TDC research is important for institutions, since student learning, trajectory, direction, and behavior depend on its evaluation. TDC in Colombia is based on 5 elements: Communicative, management, investigative, pedagogy and technology, and each of them is measured at three levels: exploratory, integrative, and innovative. The research questions are: (1) What kind of results can we expect from CDD prediction with ML techniques? (2) What ML techniques are effective in predicting TDC? (3) What are the advantages of predicting TDC with ML techniques? The methodology aims to design a prediction model of TDC in Colombia through the application of 9 ML techniques using *Orange Data Mining* software. The results show the high effectiveness of intelligent techniques to predict TDC. The model demonstrates that it is feedbackable, scalable and allows proposing personalized learning itineraries.

Keywords: teachers' digital competence, machine learning, artificial intelligence, adaptive learning, emerging technology



1. INTRODUCCIÓN

Machine Learning (ML) es un campo de la inteligencia artificial (IA) que elabora predicciones de grandes datos a partir de diferentes técnicas o algoritmos (Tarik et al., 2021) que ayudan en la toma de decisiones. El sector educativo se ve cada vez más influenciado por software con IA que potencia la productividad, la escritura, la gramática, la creación de vídeos, diseños y marketing. Herramientas como *ChatGPT 4.0*, *DALL-E*, *Midjourney*, *Instanttext*, *Bing*, *Adobe Enhance*, *Capcut*, *iContact*, *Robin* y muchas otras están revolucionando las prácticas computacionales. Incluso *Canva*, el editor online, incorpora herramientas de IA para crear diseños automáticamente.

La IA y el ML presentan un gran potencial para transformar la educación, impulsando metodologías activas e innovadoras como los proyectos STEAM en todos los niveles educativos. Estos proyectos fomentan el pensamiento crítico y la conciencia ciudadana. El ML, uno de los campos más exitosos de la IA, optimiza la productividad al automatizar tareas repetitivas, permitiendo a los docentes enfocarse en tareas más creativas y de mayor impacto (Dúo Terrón et al., 2023).

El ML y el Big Data se posicionan como pilares fundamentales en la evolución ascendente de la producción científica. Los altos indicadores bibliométricos, convierte estas áreas en focos de atención para organizaciones y países que buscan impulsar el conocimiento y la productividad (Belmonte et al., 2020).

La competencia digital docente (CDD) es fundamental para el aprendizaje de los alumnos. Esta competencia se refiere a las habilidades, estrategias y conocimientos propios de los docentes en sistemas digitales y su aplicación en los procesos de enseñanza-aprendizaje (Prendes Espinosa et al., 2018). Permite a los docentes solucionar los problemas y retos educativos que plantea la sociedad del conocimiento (Cabero-Almenara et al., 2020).

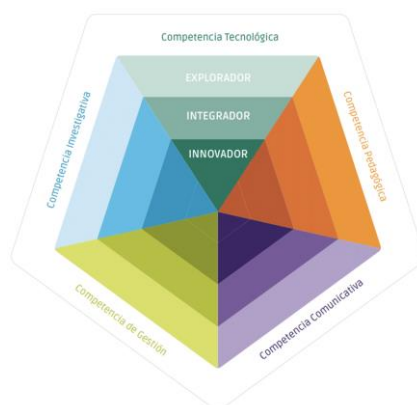
Para garantizar la eficacia de la formación, identificar los factores que la mejoran y reducir los factores que la perjudican (García-Ruiz et al., 2023) es necesario investigar mecanismos que ayuden a predecir y evaluar la CDD.

La evaluación del profesorado tiene como propósito valorar el desempeño de los docentes en las distintas facetas de su labor pedagógica (Flores et al., 2021). La CDD se adapta a cada contexto geográfico y responde a las necesidades educativas específicas de cada región. La evaluación del profesorado, estrechamente ligada a la CDD, se ha visto enriquecida por la existencia de diversos marcos de referencia de competencias, como por ejemplo el Marco Europeo de Competencia Digital del Profesorado DigCompEdu (Redecker, 2020), Estándares ISTE para Educadores (International Society for Technology in Education, 2019), Marco UNESCO de competencia TIC para Docentes (UNESCO, 2021), Marco Común español de Competencia Digital Docente (INTEF, 2017), Competencias TIC para el desarrollo profesional docente colombiano (Ministerio de Educación Nacional, 2013), Competencias y Estándares TIC para la TIC para la profesión docente en Chile (Elliot, Jaime; Gorichon, Solange; Irigoín, María; Maurizi, 2011), entre otros.

La pandemia evidenció la falta de preparación de los docentes colombianos para la transición a la virtualidad. Por ello en Colombia, fue necesario fortalecer la CDD a través de la formación docente. La CDD comprende 5 dimensiones comunicativa, de gestión, investigativa, pedagógica y tecnológica (Figura 1): Estos elementos se miden en tres niveles: explorador, integrador e innovador (Ministerio de Educación Nacional, 2013).

Figura 1

Pentágono de competencias TIC



Nota. Tomado de Ministerio de Educación Nacional (2013)

La formación y capacitación previa al empleo en CDD ayuda a los futuros docentes a desarrollar las habilidades y conocimientos necesarios para utilizar las tecnologías digitales (TD) en el aula. Esto les permite reflexionar de forma crítica sobre el valor añadido de las TD en la enseñanza (Lucas et al., 2021). Es así, que la relación de los futuros profesionales de la educación en su día a día con dispositivos electrónicos facilita la practica educativa, brindándoles confianza para optar positivamente con su continua formación digital docente (Salmerón Majadas, 2018).

La mejora de las CDD ayuda a los estudiantes a desarrollar las habilidades y conocimientos necesarios para utilizar estas tecnologías de manera efectiva facilitando la creación de ambientes de aprendizaje innovadores y atractivos. Esto les permite participar plenamente en la sociedad actual, cada vez más digital (Silva et al., 2019). Por tanto, la CDD puede ayudar al docente en la mejora de su práctica educativa y al aprendizaje por parte de los estudiantes aumentando la motivación y compromiso con su proceso académico (Forero-Corba & Negre-Bennasar, 2024).

Investigaciones previas han identificado variables predictoras de los niveles de la CDD. Cabero-Almenara et al. (2022) midieron, evaluaron y analizaron las competencias digitales de los docentes universitarios centrados en el modelo *DigCompEdu*. En este estudio, utilizaron el modelo de ecuaciones estructurales para predecir los niveles de la CDD. Gómez et al. (2022) propusieron la medición de la CDD centrada en simulaciones virtuales.

Moreno-Padilla (2019) sostiene que la implementación de la IA en la educación puede ser lenta debido a las políticas y procesos administrativos de cada país. Salas-Pilco et al. (2022) afirman que la tasa de adopción de la IA y el análisis de aprendizaje en la educación es lenta, en comparación con otros campos como la medicina, la industria y las finanzas. Sin embargo, en 2023, ya es una realidad en varios países, lo que su progreso y aceleración exponencial plantean importantes desafíos para el sector educativo (Dai et al., 2022).

La enseñanza de la IA en las escuelas primarias, junto con métodos multidisciplinares innovadores para desarrollar habilidades y competencias, ayuda a mejorar los sistemas educativos de forma integral (Denys & Klimczuk, 2022). Es por ello, que la alfabetización digital permite comprender más a profundidad sobre la verdadera utilidad y practicidad de la IA (Moreno Padilla, 2019). En este contexto, es necesario incorporar elementos necesarios para fortalecer los marcos de la CDD con técnicas inteligentes. Esto permitirá no solo estar a la vanguardia digital, sino también mejorar aspectos de los entornos educativos, como la formación docente, los procesos de alfabetización digital, la actualización del currículo y la personalización de la enseñanza (Bartolomé et al., 2018).

El objetivo de esta investigación es diseñar un modelo de predicción para evaluar la CDD de los docentes colombianos con técnicas de ML. El modelo se basó en el Marco de competencias TIC para el desarrollo profesional docente colombiano (Figura 1).

Las preguntas de investigación (PI) planteadas son:

- PI1: ¿Qué tipo de resultados podemos esperar de la predicción de la CDD con técnicas de ML?
- PI2: ¿Qué técnicas de ML son efectivas para predecir la CDD?
- PI3: ¿Qué ventajas trae predecir y evaluar la CDD con técnicas de ML?

2. MÉTODO

2.1. Diseño de la investigación

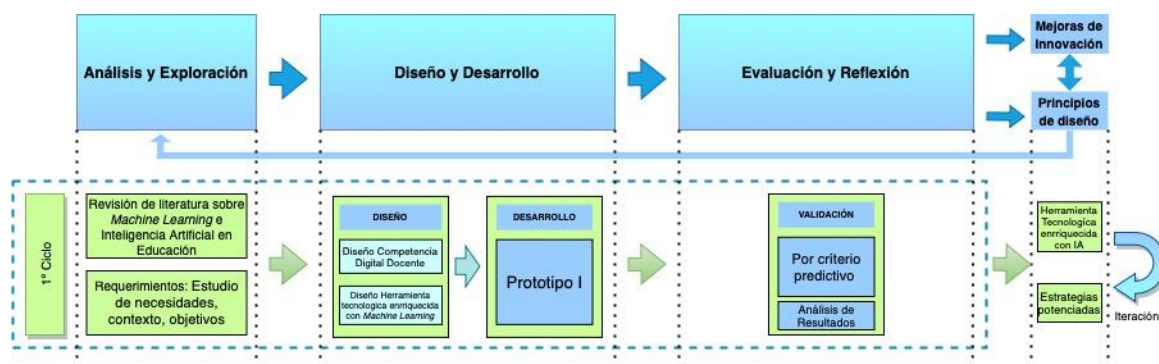
El diseño de la investigación está fundamentado en la Investigación Basada en Diseño (IBD) (De Benito & Salinas, 2016) ya que permite que el diseño y desarrollo del entorno enriquecido con ML pueda implementarse y evaluarse correctamente en futuras investigaciones para explorar procesos educativos innovadores y generar soluciones reales y de gran impacto en la práctica educativa haciendo uso de la IA.

Se implementó el modelo IBD basado en el Modelo ADDIE (Esquivel Gámez, 2014). La IBD se caracteriza por la iteración de sus fases. Esta iteración permitió un proceso continuo de revisión y reformulación del proyecto.

El modelo final para la sección del diseño se presenta en la Figura 2. Las fases de esta investigación se centraron únicamente en el diseño y desarrollo, teniendo en cuenta que las fases previas de análisis y exploración ya se efectuó en una anterior investigación. Sin embargo, el diseño de la herramienta tecnológica en la parte de evaluación se validó para confirmar las estrategias potenciadas por los principios de diseño y mejoras de innovación.

Figura 2

Modelo IBD tipo ADDIE para la investigación



2.2. Diseño Marco Competencia Digital Docente

El marco de la CDD se basa en el marco de Competencias TIC para el desarrollo del docente colombiano. Para ello, se seleccionó una rúbrica de evaluación validada (Mantilla Contreras, 2022) de 25 ítems. La rúbrica se encuentra como anexo en la sección 7.2. Esta rúbrica sirve para medir y evaluar las competencias digitales de docentes universitarios y no universitarios.

2.3. Competencias

Las siguientes definiciones de las competencias fueron tomadas del marco de *Competencias TIC para el Desarrollo Profesional Docente* (Ministerio de Educación Nacional, 2013).

Competencia comunicativa: Capacidad para expresarse, establecer contacto y relacionarse en espacios virtuales y audiovisuales a través de diversos medios y con el manejo de múltiples lenguajes, de manera sincrónica y asincrónica.

Competencia de gestión: Capacidad para utilizar las TIC en la planeación, organización, administración y evaluación de manera efectiva de los procesos educativos; tanto a nivel de prácticas pedagógicas como de desarrollo institucional.

Competencia investigativa: Capacidad de utilizar las TIC para la transformación del saber y la generación de nuevos conocimientos.

Competencia pedagógica: Capacidad de utilizar las TIC para fortalecer los procesos de enseñanza y aprendizaje, reconociendo alcances y limitaciones de la incorporación de estas tecnologías en la formación integral de los estudiantes y en su propio desarrollo profesional.

Competencia tecnológica: Capacidad para seleccionar y utilizar de forma pertinente, responsable y eficiente una variedad de herramientas tecnológicas entendiendo los principios que las rigen, la forma de combinarlas y las licencias que las amparan.

2.4. Niveles

El desarrollo profesional docente en el ámbito de las TIC se evalúa a través de tres niveles progresivos, que permiten identificar el avance y la consolidación de las competencias en cada etapa. Estos niveles son:

- Explorador
- Integrador
- Innovador

La definición particular de cada nivel, en función de la competencia requerida, se especifica en el [anexo 7.1.](#)

2.5. Instrumentos de medida

Se utilizó *Orange Data Mining* 3.35.0 para procesar y analizar la información. Los datos de entrenamiento y prueba se agruparon en una hoja de cálculo .csv. Se seleccionó este software por ser una herramienta gráfica, intuitiva y fácil de usar que implementa técnicas de ML. Esto permite que investigadores con un conocimiento mínimo en ciencia de datos y pensamiento computacional puedan implementarlo sin dificultad (Caglayan, 2019).

2.6. Datos

El marco de CDD define los datos utilizados en este estudio. La rúbrica (ver Anexo) sirve para medir y evaluar la CDD. Las variables entrenadas o variables predictoras, que en este caso son las preguntas de la rúbrica, se denominan "*Datos de entrenamiento*". Estos datos permiten enseñar a la máquina a clasificar correctamente los datos de entrada. Al enfocarse en el diseño del modelo, este estudio utiliza como datos de entrada información simulada que refleja las respuestas de los docentes a la rúbrica validada, denominándola "*Datos de prueba*".

2.6.1. Datos de Entrenamiento

En la Figura 3 se observa el widget *File Train* donde importamos los datos de entrenamiento, lo cuales nos permiten enseñarle al modelo como clasificar correctamente. Para enseñarle o entrenar nuestro modelo de predicción de las 5 competencias del CDD, en *File Train* hemos incorporado un conjunto de 21.248 variables. Estas variables nos permiten identificar patrones y relaciones clave en los datos. Las variables descriptor 1-7 (rango 1-4) y el promedio (rango 1-4) se usan para entrenar la variable objetivo "nivel de competencia" (explorador, integrador o innovador).

2.6.2. Datos de Prueba

La Figura 3 presenta el widget *File Test*, el cual importa los resultados obtenidos al aplicar la rúbrica a un conjunto de 25 docentes simulados. Estos datos sirven como insumo para nuestro modelo predictivo. En este widget se visualizan las variables utilizadas en el modelo, permitiendo evaluar su desempeño. Los docentes están identificados en una nueva variable llamada ID. Para identificarlos fácilmente etiquetamos a cada docente con las letras del [A, B, C, D, E, F,..., Z]. Para cada variable descriptor, se establecieron respuestas aleatorias. Como

existen 25 docentes, cada uno tiene definido 5 competencias, por tanto, se definen 125 instancias.

2.7. Técnicas de Machine Learning

La Tabla 1 define las nueve técnicas de *ML* utilizadas en el modelo predictivo de la CDD.

Tabla 1

Técnicas de Machine Learning utilizadas en el modelo

N	Abreviatura	Técnica	Definición
1	kNN	<i>k-Nearest Neighbors</i>	<i>kNN</i> busca los <i>k</i> ejemplos de entrenamiento más cercanos en el espacio de características y utiliza su media como predicción.
2	LR	<i>Logistic Regression</i>	<i>Logistic Regression</i> es un modelo estadístico que predice la probabilidad de que ocurra un evento. Se basa en la función logística, que transforma valores continuos en valores probabilísticos.
3	DS	<i>Decision Tree</i>	<i>Tree</i> es un algoritmo que divide los datos en nodos en función de la de clase. Para variables categóricas, utiliza la ganancia de información, y para variables numéricas, utiliza el error cuadrático medio. <i>Tree</i> es precursor de <i>Random Forest</i> . <i>Tree</i> puede manejar conjuntos de datos tanto categóricos como numéricos.
4	RF	<i>Random Forest</i>	<i>Random Forest</i> es un algoritmo que construye un conjunto de árboles de decisión. Cada árbol se desarrolla a partir de una muestra aleatoria de los datos de entrenamiento. Para reducir la varianza, cada árbol se construye utilizando un subconjunto aleatorio de atributos. El modelo final se basa en el voto mayoritario.
5	NB	<i>Naive Bayes</i>	El algoritmo <i>Naive Bayes</i> se utiliza para clasificar datos y se basa en el teorema de Bayes. Calcula la probabilidad de cada clase, dado el conjunto de características. La clase con la probabilidad más alta es la clase predicha. Es un algoritmo simple y eficiente para conjunto de datos con gran número de características.
6	GB	<i>Gradient Boosting</i>	<i>Gradient Boosting</i> es un algoritmo para problemas de regresión y clasificación, que produce un modelo de predicción fuerte a partir de conjunto de modelos de predicción débiles. El nuevo modelo se entrena para reducir el error del modelo anterior, y luego se agrega al modelo final.
7	SVM	<i>Support Vector Machine</i>	<i>SVM</i> separa datos de dos clases con un hiperplano, maximizando el margen entre clases. Es eficaz en problemas de clasificación lineal donde los datos se pueden separar con un hiperplano lineal.
8	SGD	<i>Stochastic Gradient Descent</i>	<i>SGD</i> usa la función del gradiente descendente estocástico para minimizar una función de pérdida con una función lineal. Se aproxima al gradiente real con una muestra cada vez y actualiza el modelo simultáneamente. Es útil para conjuntos de datos a gran escala y dispersos.
9	NN	<i>Neural Network</i>	<i>NN</i> es un algoritmo inspirado en el cerebro humano. Se compone de neuronas artificiales interconectadas, similar a las neuronas biológicas. Su objetivo es aprender a realizar una tarea a partir de datos de entrenamiento. La red aprende a identificar patrones y usarlos para realizar la tarea.





2.8. Validación del modelo

La validación del diseño del prototipo está basada por el método *validez del criterio predictivo* (Ramírez Martínez, 2019), donde refiere la capacidad de un instrumento de medición para predecir un futuro resultado en base a la probabilidad de éxito en un programa de entrenamiento. La validez de criterio asocia el rendimiento del instrumento y la toma de decisiones, utilizando como medida de validación el área bajo la curva AUC (*Area Under the ROC Curve*) y la precisión (López et al., 2019).

En la Tabla 2 se evidencia los 4 *widgets* usados en *Orange Data Mining* para validar el modelo y ver resultados.

Tabla 2

Widgets usados en Orange Data Mining para validar el modelo

Widgets	Icono	Definición
<i>Test and Score</i>	 Test and Score	Prueba algoritmos de aprendizaje. Muestra una tabla con diferentes medidas de rendimiento del clasificador, como la precisión de la clasificación y el AUC.
<i>Confusion Matrix</i>	 Confusion Matrix	Tabla que muestra el número o proporción de instancias correcta y erróneamente clasificadas por un modelo.
<i>Feature Statistics</i>	 Feature Statistics	Tabla que muestra el número o proporción de instancias correcta y erróneamente clasificadas por un modelo.
<i>Scatter Plot</i>	 Scatter Plot	Muestra la visualización de gráficos de dispersión con análisis exploratorio y mejoras inteligentes de visualización de datos.

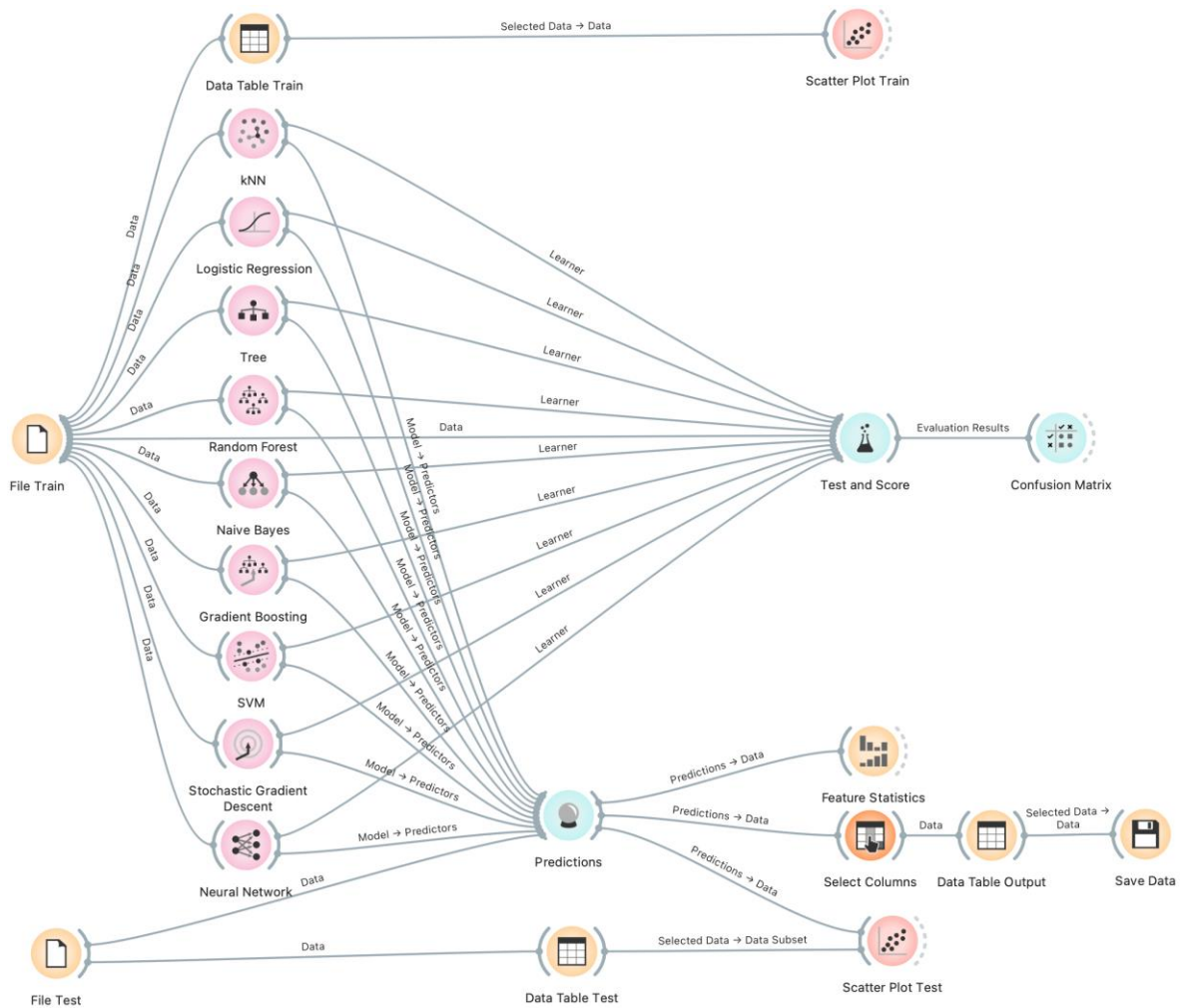
3. RESULTADOS

La Figura 3 muestra el diseño del modelo de predicción. Se incorporan 9 técnicas de ML. El widget *Predictions* permite clasificar por técnica. Los widgets *Data Table Train* y *Scatter Plot Train* visualizan los datos de entrenamiento. Los widgets *Data Table Test* y *Scatter Plot Test* visualizan los datos de prueba. La predicción de datos de prueba se visualiza combinando *File Test* y *Predictions*. *Data Table Test* está conectado con *File Test*, por lo que se puede combinar con *Predictions* para visualizar las predicciones.

Además, se introducen los widgets *Test and Score*, *Confusion Matrix* y *Feature statistics* para la validación del modelo.

Por último, se agregan tres widgets *Select Columns*, *Data Table Output* y *Save Data*, con el fin de seleccionar y exportar los resultados que necesitemos para la toma de decisiones.

Figura 3
Diseño del modelo



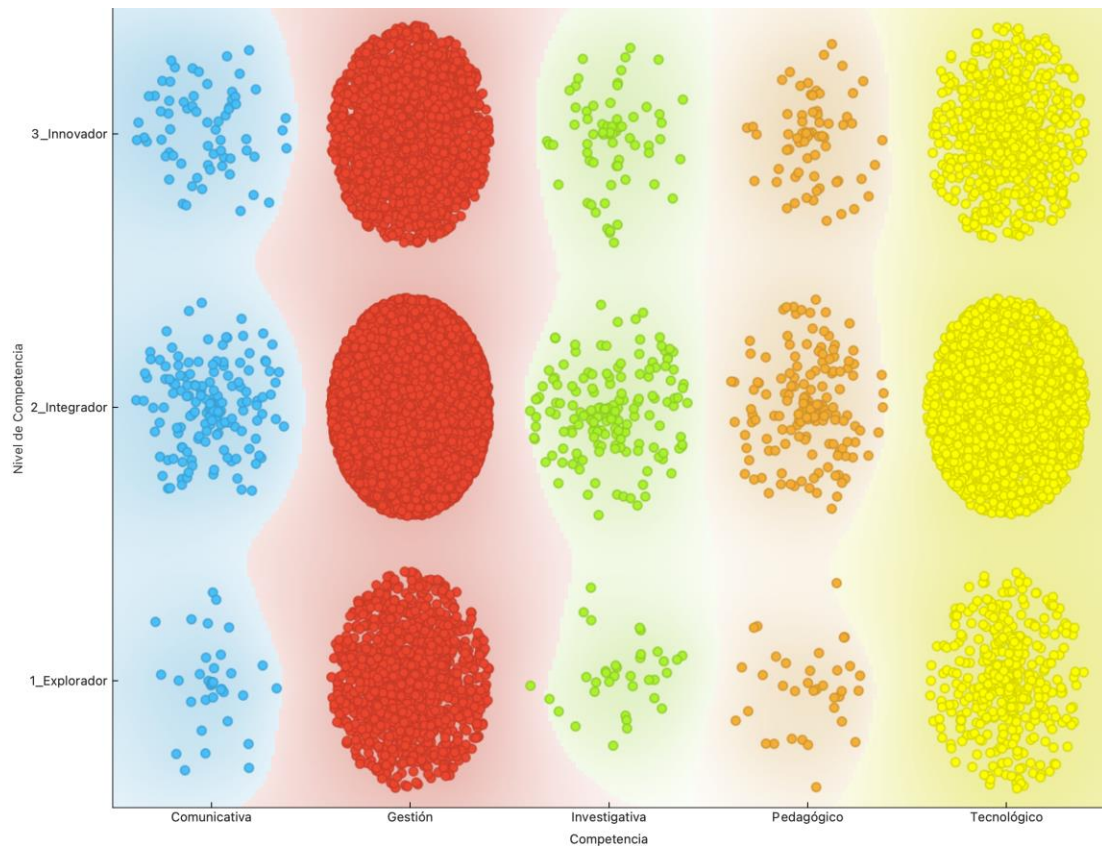
Para responder a la primera pregunta de investigación, se analizaron las Figuras 9, 10, 11 y 12, así como las Tablas 3 y 4.

PI1: ¿Qué tipo de resultados podemos esperar de la predicción de la CDD con técnicas de ML?

La Figura 4 muestra la agrupación de los 21.248 datos entrenados para la predicción de los datos de prueba frente a las variables de Competencia vs Nivel de competencia.

Figura 4

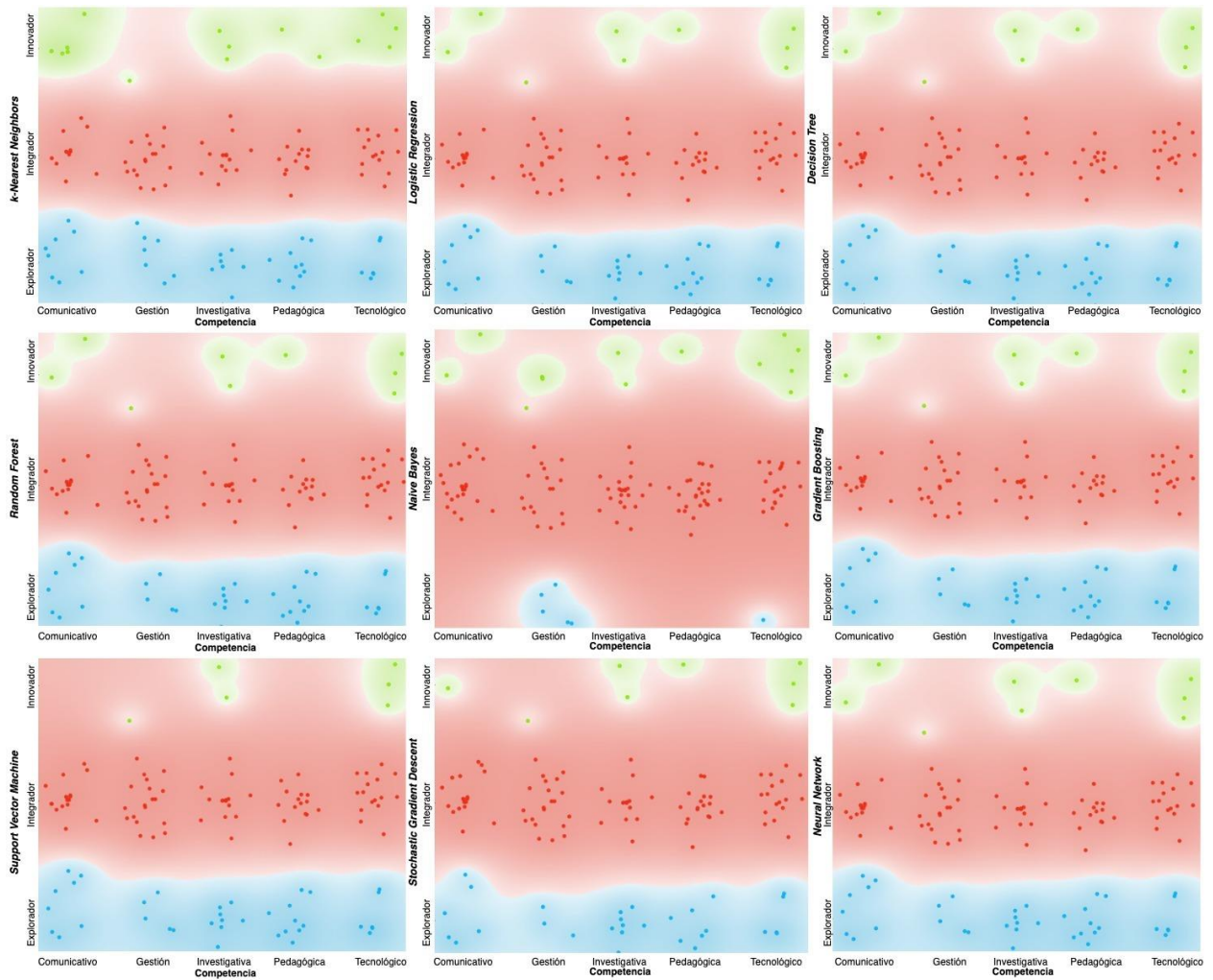
Agrupación datos de entrenamiento



La Figura 5 muestra la agrupación de las técnicas de ML para la predicción de la CDD. Cada punto representa la competencia docente en el nivel correspondiente. El contorno muestra la CDD de los 25 docentes simulados. Los datos aleatorios ubican la CDD docente en gran porcentaje en los niveles explorador e integrador para las 9 técnicas.

Figura 5

Visualización de la predicción de las CDD con las 9 técnicas de ML



La Tabla 3 muestra los resultados estadísticos centrados en los descriptores de la rúbrica. Las variables estadísticas medidas fueron media, mediana, dispersión, valor mínimo y valor máximo.

Tabla 3

Resultados estadísticos de los descriptores

Descriptores	Media	Mediana	Dispersión	Mínimo	Máximo
Descriptor 1	2.22	2	0.51	1	4
Descriptor 2	2.13	2	0.47	1	4
Descriptor 3	2.06	2	0.52	1	4
Descriptor 4	2.05	2	0.49	1	4
Descriptor 5	2.26	2	0.42	1	4
Descriptor 6	2.34	2	0.44	1	4
Descriptor 7	2.40	2	0.44	1	4

La Tabla 3 muestra el rendimiento que cada una de las técnicas usadas para la predicción de la CDD.

PI2: ¿Qué técnicas de ML son efectivas para predecir la CDD?

Para responder la segunda pregunta de investigación, se analizó la Tabla 4.

La Tabla 4 muestra que el AUC y Precisión para las técnicas *Logistic Regression*, *Decision Tree*, *Random Forest*, *Gradient Boosting* y *Neural Network* son del 100%, con ello, podemos afirmar que el entrenamiento de los 21.248 datos fue bastante efectivo para que nuestro modelo predijera la CDD y que estas técnicas son las más efectivas para poder predecir las Competencias TIC del docente colombiano.

Tabla 4

Rendimiento de las técnicas de ML mediante AUC y Precisión

Técnica	AUC	Precisión
<i>k-Nearest Neighbors</i>	0.941	0.850
<i>Logistic Regression</i>	1.000	1.000
<i>Decision Tree</i>	1.000	1.000
<i>Random Forest</i>	1.000	1.000
<i>Naive Bayes</i>	0.978	0.920
<i>Gradient Boosting</i>	1.000	1.000
<i>Support Vector Machine</i>	0.998	0.988
<i>Stochastic Gradient Descent</i>	0.902	0.952
<i>Neural Network</i>	1.000	1.000

Nota. Elaboración propia

La Figura 6 muestra la matriz de confusión únicamente para las técnicas *kNN*, *NB*, *SVM* y *SGD*, debido a que según la Tabla 4, fueron técnicas donde no llegaron al 100%, clasificando incorrectamente los datos de prueba.

Figura 6

Matriz de confusión para las técnicas kNN, NB, SVM y SGD

Predicted <i>kNN</i>					Predicted <i>Naive Bayes</i>				
	1_Explorador	2_Integrador	3_Innovador	Σ		1_Explorador	2_Integrador	3_Innovador	Σ
1_Explorador	1371	618	0	1989	1_Explorador	1550	439	0	1989
2_Integrador	538	14507	668	15713	2_Integrador	0	13879	1834	15713
3_Innovador	0	1289	2257	3546	3_Innovador	0	0	3546	3546
Σ	1909	16414	2925	21248	Σ	1550	14318	5380	21248

Predicted <i>SVM</i>					Predicted <i>SGD</i>				
	1_Explorador	2_Integrador	3_Innovador	Σ		1_Explorador	2_Integrador	3_Innovador	Σ
1_Explorador	1975	14	0	1989	1_Explorador	1143	846	0	1989
2_Integrador	13	15602	98	15713	2_Integrador	0	15698	15	15713
3_Innovador	0	133	3413	3546	3_Innovador	0	217	3329	3546
Σ	1988	15749	3511	21248	Σ	1143	16761	3344	21248

La Figura 7 muestra una matriz de comparación de las 9 técnicas de ML entre sí. En ella, podemos ver qué técnica es más efectiva contra otra técnica.

Figura 7

Matriz de comparación entre técnicas de ML

	Gradient Boosting	kNN	Logistic Regressi...	Naive Bayes	Neural Network	Random Forest	SGD	SVM	Tree
Gradient Boosting		1.000	0.500	0.999	0.500	0.500	1.000	0.955	0.500
kNN	0.000		0.000	0.000	0.000	0.000	0.989	0.000	0.000
Logistic Regression	0.500	1.000		0.999	0.500	0.500	1.000	0.955	0.500
Naive Bayes	0.001	1.000	0.001		0.001	0.001	0.999	0.001	0.001
Neural Network	0.500	1.000	0.500	0.999		0.500	1.000	0.955	0.500
Random Forest	0.500	1.000	0.500	0.999	0.500		1.000	0.955	0.500
SGD	0.000	0.011	0.000	0.001	0.000	0.000		0.000	0.000
SVM	0.045	1.000	0.045	0.999	0.045	0.045	1.000		0.045
Tree	0.500	1.000	0.500	0.999	0.500	0.500	1.000	0.955	

PI3: ¿Qué ventajas trae predecir y evaluar la CDD con técnicas de ML?

A continuación, se denotan las ventajas que puede tener el predecir la CDD con técnicas de ML.

Retroalimentación: El análisis de grandes cantidades de datos permite que las técnicas de ML se alimenten de información de forma constante para predecir de forma correcta las competencias digitales de cada docente. Esta información puede ayudar a los directivos docentes a tomar decisiones sobre cómo enfocar la alfabetización digital o el desarrollo profesional de los docentes en la institución. El modelo diseñado permite categorizar a los docentes y agregar nuevas muestras (docentes) para compararlos con muestras antiguas (docentes ya incluidos en el modelo). Con estas predicciones, se pueden tomar decisiones correctas para el crecimiento y la actualización profesional del nuevo docente en la institución.

Escalabilidad: Las técnicas de ML puede favorecer a la aplicación de modelos de predicción de la CDD en diferentes contextos. El diseño del modelo desarrollado permite adaptarse fácilmente a instrumentos o marcos de competencia digital de otros países, únicamente se debe ajustar las variables tanto de entrenamiento como de prueba, de acuerdo con la parametrización de variables que conlleve el marco.

Habilidades del siglo XXI: El uso de técnicas inteligentes en la predicción de la CDD permite desarrollar las habilidades que los docentes necesitan actualmente para desempeñarse de manera eficaz en el aula, especialmente en contextos educativos dinámicos.

Desarrollo sostenible: El uso de técnicas de ML para la predicción de la CDD permite ir a la vanguardia del desarrollo tecnológico y mejorar el cuarto objetivo de desarrollo sostenible, que se centra en la educación inclusiva, equitativa y de calidad para generar oportunidades de aprendizaje (Organización de la Naciones Unidas, 2018).

4. DISCUSIÓN Y CONCLUSIONES

El modelo diseñado y desarrollado en la investigación ayuda a predecir y evaluar la CDD usando técnicas de ML. Este diseño es escalable y adaptable a instrumentos o marcos de competencia digital de otros países. Esto se logra ajustando las variables de entrenamiento y prueba a la parametrización del marco en cuestión. Además, el modelo permite la retroalimentación a nivel de datos, ya que puede almacenar grandes cantidades de muestras. Con estas muestras, se puede predecir correctamente la competencia digital de un docente o grupo de docentes de una institución educativa.

Instituciones educativas están generando iniciativas para orientar y mejorar el desarrollo de la CDD de acuerdo con las estandarizaciones del marco, incluyendo parámetros de mejora con respecto a otros marcos de competencia de nivel internacional (Silva et al., 2019). Diseñar y desarrollar un modelo de predicción con técnicas de ML motiva a las instituciones a implementar modelos para evaluar la CDD, especialmente con herramientas enriquecidas con IA. El uso de herramientas enriquecidas con IA puede ayudar a las instituciones a mejorar la eficiencia y precisión de la evaluación de la CDD.

En el ámbito de la IA, nos encontramos con sistemas o procesos cuyo funcionamiento interno permanece oculto o inaccesible para el usuario, lo que se conoce como *black box* (Bearman & Ajjawi, 2023). A pesar de que estos sistemas inteligentes pueden generar resultados precisos, la falta de comprensión sobre el proceso que los lleva a esas conclusiones plantea interrogantes en torno a una visión política y ética (Suárez-Guerrero et al., 2020). Por ello, se debe tener en cuenta consideraciones éticas clave como la transparencia, sesgo algorítmico, privacidad y seguridad en los datos e.t.c, para garantizar que sistemas inteligentes se utilicen de manera responsable en beneficio de la sociedad (UNESCO, 2019).

La importancia de establecer niveles de desarrollo de la CDD mediante un instrumento o rúbrica permite evaluar las habilidades docentes frente a herramientas tecnológicas que son clave para autorregular los actuales procesos de enseñanza-aprendizaje (Gisbert & Lázaro, 2018). Por ello, la rúbrica ya evaluada por Mantilla Contreras (2022), es efectiva para la investigación porque mide las competencias digitales de los docentes colombianos, sean o no universitarios. La rúbrica establece las 5 competencias en los 3 niveles que son variables estandarizadas por el Ministerio de Educación Nacional.

Los datos para el análisis de la CDD se obtuvieron a partir de los descriptores establecidos en la rúbrica. La rúbrica de evaluación propuesta puede utilizarse en procesos de evaluación, ya sea de forma autoevaluativa o externa. El diseño del modelo, a partir de los resultados predictivos de los niveles de competencias, puede orientar la toma de decisiones de los docentes (en la modalidad de autoevaluación) o de las instituciones educativas y organizaciones del sector educativo (en modalidad de evaluación externa) frente a las necesidades formativas.

Investigaciones de la evaluación de la competencia digital implementan variables predictoras sociodemográficas para la predicción de la CDD como por ejemplo *edad, genero, experiencia docente, experiencia TIC*, entre otras (Cabero-Almenara et al., 2022). El implementar variables sociodemográficas junto con las variables de la rúbrica o instrumentos de evaluación mejoraría

la predicción de la CDD. Además, puede ser efectiva al momento de aplicar nuevas muestras para el proceso de retroalimentación, autoaprendizaje y escalabilidad del modelo.

La validación del modelo estuvo dada por la validez del criterio predictivo, donde el resultado predicho de las técnicas de ML se centra en la probabilidad de éxito de los datos de entrenamiento frente a los de prueba, comparando así resultados. López et al. (2019) afirma que si el área bajo la curva AUC es mayor de 0.9, el instrumento aplicado en el modelo es excelente, por lo tanto, el diseño del modelo predictivo se ajusta a los elementos de evaluación de la CDD dado que todas las predicciones de las técnicas de ML estuvieron por encima del 0.9 como se especifica en la Tabla 4.

Estudios basados en la aplicación de herramientas con técnicas inteligentes han desarrollado modelos con un valor de predicción alrededor del 99%. Un ejemplo de ello es el estudio de Hougue et al. (2022), en el que se estimó el rendimiento académico de estudiantes de una institución educativa a partir de técnicas de aprendizaje automático, obteniendo un 99% de precisión en la predicción.

El software *Orange Data Mining* fue fundamental en el diseño y desarrollo del modelo de predicción de la CDD al ser una herramienta para el análisis y minería de datos. Este software incorpora variadas técnicas de ML, lo que permite mostrar diferentes tipos de resultados, como parámetros estadísticos, matrices de evaluación, visualización predictiva numérica y/o gráfica. Estas características lo convierten en una herramienta muy poderosa para implementar modelos enriquecidos con IA. Además, Caglayan (2019) señala que esta *Orange Data Mining* es muy fácil de implementar al ser de modo gráfico, lo que le da versatilidad en ambientes tanto de desarrollo e investigación.

Las nueve técnicas de ML aplicadas se tomaron del software *Orange Data Mining*, ya que estaban predefinidas. Estas técnicas son bien conocidas en el campo de las ciencias de la computación y análisis de datos, donde su implementación en entornos de desarrollo e investigación está en constante crecimiento (Forero-Corba & Bennasar, 2024). Sin embargo, a pesar de las investigaciones sobre la implementación de la IA en educación, Moreno Padilla (2019) expone que aún quedan muchos elementos por trabajar y organizar antes de que sea posible una verdadera implementación de la IA en la educación. En este sentido, implementar técnicas de ML en procesos de alfabetización digital y desarrollo profesional docente podría ser uno de los pasos que se han dado para una implementación efectiva de la IA en el sector educativo.

La Figura 5 presenta 9 gráficos que ilustran el nivel de competencia de cada docente en cada una de las técnicas implementadas en el diseño del modelo. Estos gráficos agrupan los datos por niveles de competencia siguiendo la resolución de la rúbrica definida. Se observa una concentración de predicción de los resultados en los niveles 2 (Integrador) y 1 (Explorador). El modelo predice una cantidad mínima de datos de prueba en el nivel 3 (Innovador). Los resultados sugieren que las directivas docentes deben implementar planes y estrategias de formación y actualización pedagógica y digital para los docentes. Esto permitirá mejorar sus habilidades y aumentar su nivel de competencia, especialmente en el nivel 3 (Innovador).

La predicción y evaluación de la CDD con técnicas inteligentes permite cerrar la brecha digital mediante la mejora de los procesos de enseñanza-aprendizaje, la actualización del currículo, el

aumento del rendimiento académico y la disminución de la deserción escolar. Pérez-Escoda et al. (2021) señalan que las principales competencias digitales que debe tener todo docente para disminuir la brecha digital son: Conocimiento de las herramientas digitales, autopercepción de las propias competencias digitales, dominio de las competencias digitales para el desempeño profesional, grado de competencia digital de los alumnos e inclusión curricular de las competencias digitales.

5. LIMITACIONES Y FUTURAS INVESTIGACIONES

El modelo se entrena y prueba con las variables de la rúbrica. No obstante, para aumentar la precisión en la predicción de la CDD, se aconseja incorporar más variables al modelo, como las sociodemográficas (edad, sexo, nivel educativo) y las relacionadas con la trayectoria profesional (experiencia docente, formación en TIC). Estas variables adicionales permitirían obtener una mejor evaluación de las características de los docentes como su familiaridad con las nuevas tecnologías, su capacidad de aprendizaje y su disposición a experimentar con nuevas metodologías.

Para robustecer la evaluación de la CDD, se recomienda la aplicación de otras rúbricas e instrumentos ya validados en futuros estudios. Esto permitirá contrastar los resultados del modelo actual y fortalecer la confiabilidad de las predicciones. Diversos estudios, como los de Gisbert & Lázaro (2018), Gómez et al. (2022), García Tamarit et al. (2021), Cabero-Almenara & Palacios-Rodríguez (2020) y Touron et al. (2018), proponen rúbricas e instrumentos validados que pueden ser utilizados para este fin. La escalabilidad del modelo actual facilita la incorporación de estas herramientas adicionales, lo que permitirá obtener una medición y evaluación más completa y precisa de la CDD.

Para futuras investigaciones, se recomienda implementar este diseño de predicción de la CDD en un contexto real con usuarios y datos no simulados. Esta evaluación permitirá estimar y comparar con mayor precisión las habilidades reales de los docentes, brindando información valiosa para la toma de decisiones sobre las necesidades formativas.

En otro orden de cosas, el modelo propuesto podría arrojar información para facilitar la propuesta de itinerarios personalizados de aprendizaje (De Benito et al., 2020) para que además de predecir los niveles de competencia digital, pueda generar recomendaciones personalizadas de aprendizaje en base a las predicciones de la CDD, lo que permitiría, entre otras cosas, una mayor flexibilización del proceso de enseñanza-aprendizaje y su adaptación a las características personales del estudiante así como a la propuesta de itinerarios centrados en mejorar la autonomía del estudiante y aumentar su autonomía. Estos itinerarios, con la ayuda del modelo de ML generado, se podría ajustar a los distintos estilos de aprendizaje, a las necesidades específicas de cada estudiante, así como a otros factores relacionados con el aprendizaje como la motivación, la autonomía, la personalización, entre otros). En este caso, estos itinerarios se caracterizarían por su carácter flexible (Salinas & Agudelo, 2016; Salinas & de Benito Crosetti, 2020). En cualquier caso, es importante señalar que la CDD es un factor determinante para el diseño e implementación de estrategias didácticas centradas en la flexibilización, en especial si se trata de formación en entornos virtuales, en los que el dominio de la competencia tecnológica se convierte en un elemento clave.

6. REFERENCIAS

- Bartolomé, A., Castañeda, L., & Adell, J. (2018). Personalisation in educational technology: the absence of underlying pedagogies. *International Journal of Educational Technology in Higher Education*, 15(1). <https://doi.org/10.1186/s41239-018-0095-0>
- Bearman, M., & Ajjawi, R. (2023). Learning to work with the black box: Pedagogy for a world with artificial intelligence. *British Journal of Educational Technology*, 54(5), 1160–1173. <https://doi.org/10.1111/bjet.13337>
- Belmonte, J. L., Segura-Robles, A., Moreno-Guerrero, A. J., & Parra-González, M. E. (2020). Machine learning and big data in the impact literature. A bibliometric review with scientific mapping in web of science. *Symmetry*, 12(4). <https://doi.org/10.3390/SYM12040495>
- Cabero-Almenara, J., Barroso-Osuna, J., Llorente-Cejudo, C., & Palacios-Rodríguez, A. (2022). Validación Del Marco Europeo De Competencia Digital Docente Mediante Ecuaciones Estructurales. *Revista Mexicana de Investigacion Educativa*, 27(92), 185–208. https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1405-66662022000100185
- Cabero-Almenara, J., & Palacios-Rodríguez, A. (2020). Marco Europeo de Competencia Digital Docente «DigCompEdu». Traducción y adaptación del cuestionario «DigCompEdu Check-In». *Edmetic*, 9(1), 213–234. <https://doi.org/10.21071/edmetic.v9i1.12462>
- Cabero-Almenara, J., Romero-Tena, R., Barroso-Osuna, J., & Palacios-Rodríguez, A. (2020). Marcos de Competencias Digitales Docentes y su adecuación al profesorado universitario y no universitario. *RECIE. Revista Caribeña de Investigación Educativa*, 4(2), 137–158. <https://doi.org/10.32541/recie.2020.v4i2.pp137-158>
- Caglayan, C. (2019). Comparison of the Code-based or Tool-based Teaching of the Machine Learning Algorithm for the First-Time Learners. *1st International Informatics and Software Engineering Conference: Innovative Technologies for Digital Transformation, IISEC 2019 - Proceedings, November 2019*. <https://doi.org/10.1109/UBMYK48245.2019.8965519>
- Dai, Y., Liu, A., Qin, J., Guo, Y., Jong, M. S. Y., Chai, C. S., & Lin, Z. (2022). Collaborative construction of artificial intelligence curriculum in primary schools. *Journal of Engineering Education*, October 2022, 23–42. <https://doi.org/10.1002/jee.20503>
- De Benito, B., Moreno García, J., & Villatoro Moral, S. (2020). Entornos tecnológicos en el codiseño de itinerarios personalizados de aprendizaje en la enseñanza superior. *Eduotec. Revista Electrónica de Tecnología Educativa*, 74, 73–93. <https://doi.org/10.21556/edutec.2020.74.1843>
- De Benito, B., & Salinas, J. M. (2016). La Investigación Basada en Diseño en Tecnología Educativa Design-Based Research in Educational Technology. *Revista Interuniversitaria de Investigación en Tecnología Educativa*, 0(1), 44–59. <https://doi.org/10.6018/riite2016/260631>
- Denys, B., & Klimczuk, B. (2022). *International Cooperation Towards Digital Transformation and Digital Ecosystems in Education*. 589–593.

- Dúo Terrón, P., Moreno Guerrero, A. J., López Belmonte, J., & Marín Marín, J. A. (2023). Inteligencia Artificial y Machine Learning como recurso educativo desde la perspectiva de docentes en distintas etapas educativas no universitarias. *Revista Interuniversitaria de Investigación En Tecnología Educativa*, 58–78. <https://doi.org/10.6018/riite.579611>
- Elliot, Jaime; Gorichon, Solange; Irigoín, María; Maurizi, M. R. (2011). *Competencias y Estándares TIC para la Profesión Docente*. 98.
- Esquivel Gámez, I. (2014). Los Modelos Tecno-Educativos, revolucionando el aprendizaje del siglo XXI. In *Nucl. Phys.* (Vol. 13, Issue 1). https://www.researchgate.net/publication/280301257_Los_Modelos_Tecno-Educativos_revolucionando_el_aprendizaje_del_siglo_XXI
- Flores, N., Castelán, V., & Zamora, M. (2021). Evaluación del perfil del profesorado a partir de los atributos del desempeño docente. *Revista Innova Educación*, 3(3), 53–72. <https://doi.org/10.35622/j.rie.2021.03.003>
- Forero-Corba, W., & Bennasar, F. N. (2024). Técnicas y aplicaciones del Machine Learning e Inteligencia Artificial en educación: una revisión sistemática. *RIED-Revista Iberoamericana de Educacion a Distancia*, 27(1), 209–253. <https://doi.org/10.5944/ried.27.1.37491>
- Forero-Corba, W., & Negre Bennasar, F. (2024). Techniques and applications of Machine Learning and Artificial Intelligence in education: a systematic review. *RIED-Revista Iberoamericana de Educacion a Distancia*, 27(1), 209–253. <https://doi.org/10.5944/ried.27.1.37491>
- García Tamarit, C. O. N. S. U. E. L. O., Perochena González, P. A. O. L. A., & Orcos Palma, L. A. R. A. (2021). The design and validation of an instrument for assessing undergraduate dissertations. *Bordon. Revista de Pedagogia*, 73(2), 79–96. <https://doi.org/10.13042/Bordon.2021.89015>
- García-Ruiz, R., Buenestado-Fernández, M., & Ramírez-Montoya, M. S. (2023). Assessment of Digital Teaching Competence: Instruments, results and proposals. Systematic literature review. *Educacion XX1*, 26(1), 273–301. <https://doi.org/10.5944/educxx1.33520>
- Gisbert, M., & Lázaro, J. (2018). Una Rubrica Para Evaluar La Competencia Digital Del Profesor Universitario En El Contexto Latinoamericano a Rubric To Evaluate the Digital Competence of the University. *EDUTEC Revista Electrónica de Tecnología Educativa.*, 0(63), 1–14. <https://doi.org/10.21556/edutec.2018.63.1091>
- Gómez, R., Palacios, A., Moreno-Mediavilla, D., & Barreras, Á. (2022). Teacher competences in the use of STEM virtual simulations: design and validation of a measurement instrument (CDUSV). *Bordon. Revista de Pedagogia*, 74(4), 85–102. <https://doi.org/10.13042/Bordon.2022.94154>
- Houngue, P., Hountondji, M., & Dagba, T. (2022). An Effective Decision-Making Support for Student Academic Path Selection using Machine Learning. *International Journal of Advanced Computer Science and Applications*, 13(11), 727–734. <https://doi.org/10.14569/IJACSA.2022.0131184>
- INTEF. (2017). *MARCO COMÚN DE COMPETENCIA DIGITAL DOCENTE*.

- International Society for Technology in Education. (2019). *ISTE Standards*.
- López, R., Avello, R., Palmero, D. E., Sánchez, S., & Quintana, M. (2019). Validation of instruments as a guarantee of credibility in scientific research. *Revista Cubana de Medicina Militar*, 48(2), 441–450.
- Lucas, M., Bem-Haja, P., Siddiq, F., Moreira, A., & Redecker, C. (2021). The relation between in-service teachers' digital competence and personal and contextual factors: What matters most? *Computers and Education*, 160(October 2020). <https://doi.org/10.1016/j.compedu.2020.104052>
- Mantilla Contreras, M. A. (2022). *MODELO DE FORMACIÓN PARA EL DESARROLLO DE COMPETENCIAS DIGITALES EN DOCENTES DE UNA UNIVERSIDAD DEL NORORIENTE COLOMBIANO*. Universidad de les Illes Balears.
- Ministerio de Educación Nacional. (2013). Competencias TIC Para el Desarrollo Profesional Docente. In *Imprenta Nacional* (Oficina de, Vol. 82).
- Moreno Padilla, R. D. (2019). La llegada de la inteligencia artificial a la educación. *Revista de Investigación En Tecnologías de La Información*, 7(14), 260–270. <https://doi.org/10.36825/riti.07.14.022>
- Organización de la Naciones Unidas. (2018). La Agenda 2030 y los Objetivos de Desarrollo Sostenible Una oportunidad para América Latina y el Caribe Gracias por su interés en esta publicación de la CEPAL. In *Publicación de las Naciones Unidas*.
- Pérez-Escoda, A., Iglesias-Rodríguez, A., Meléndez-Rodríguez, Lady, & Berrocal-Carvajal, V. (2021). Competencia digital docente para la reducción de la brecha digital: Estudio comparativo de España y Costa Rica. *Tripodos*, 46, 77–96. <https://doi.org/10.51698/tripodos.2020.46p77-96>
- Prendes Espinosa, M. ^a P., Solano Fernández, I. M., Serrano Sánchez, J. L., González Calatayud, V., & Román García, M. ^a del M. (2018). Entornos Personales de Aprendizaje para la comprensión y desarrollo de la Competencia Digital: análisis de los estudiantes universitarios en España. *Educatio Siglo XXI*, 36(2 Julio), 115. <https://doi.org/10.6018/j/333081>
- Ramírez Martínez, J. L. (2019). El proceso de elaboración y validación de un instrumento de medición documental. *Acción y Reflexión Educativa*, 44, 50–63.
- Redecker, C. (2020). Marco Europeo para la Competencia Digital de los Educadores: DigCompEdu. In *Secretaría General Técnica del Ministerio de Educación y Formación Profesional de España (Original publicado en 2017)*.
- Salas-Pilco, S. Z., Xiao, K., & Hu, X. (2022). Artificial Intelligence and Learning Analytics in Teacher Education: A Systematic Review. *Education Sciences*, 12(8). <https://doi.org/10.3390/educsci12080569>
- Salinas Ibáñez, J. M., & Agudelo Velásquez, O. L. (2016). Itinerarios flexibles de aprendizaje y mapas conceptuales: un abánico de posibilidades para todos los niveles educativos. *Proc. of the Seventh Int. Conference on Concept Mapping, Tabla 1*.

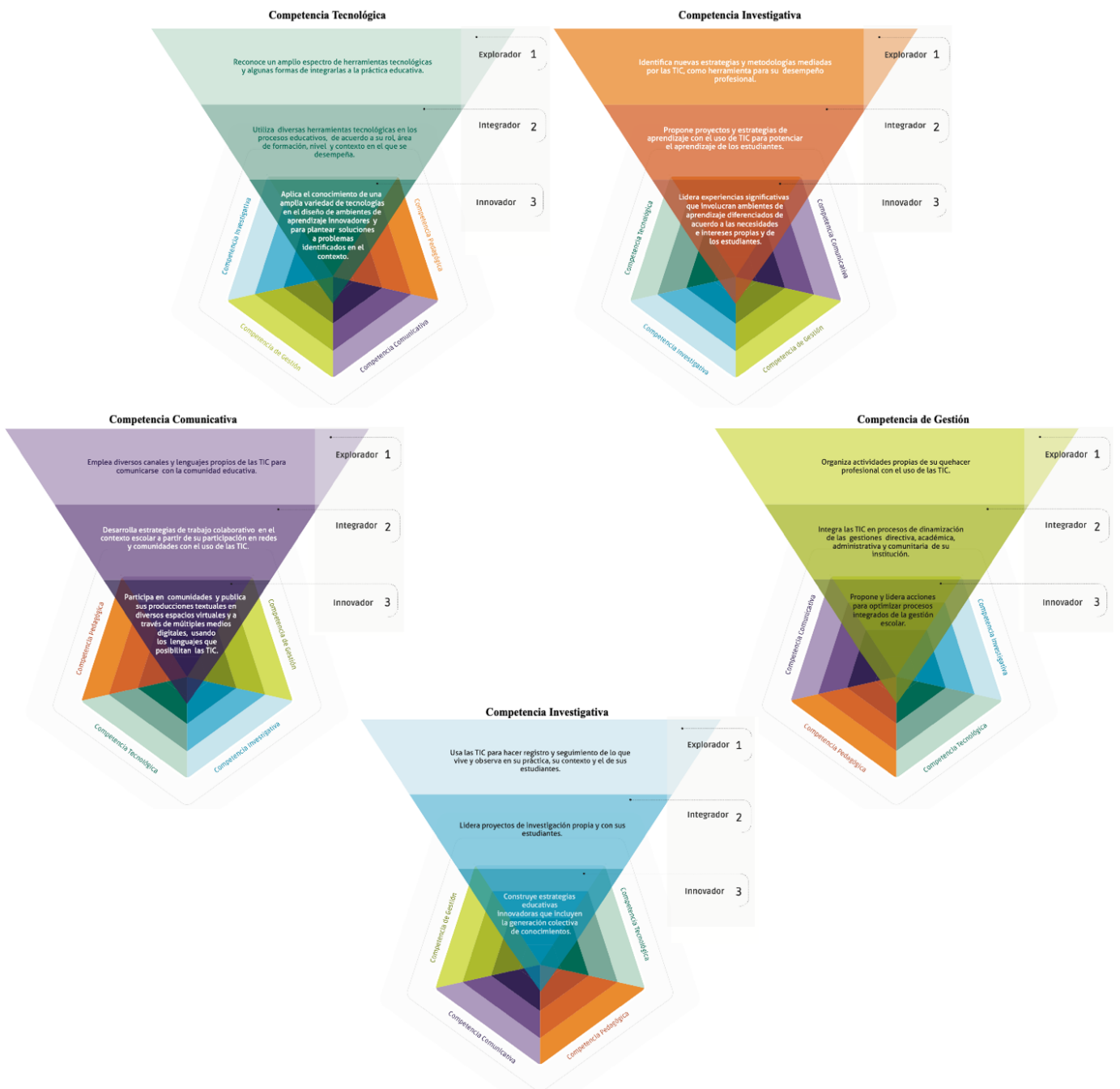
- Salinas Ibáñez, J. M., & de Benito Crosetti, B. L. (2020). Construcción de itinerarios personalizados de aprendizaje mediante métodos mixtos. *Comunicar*, 65, 31–42. <https://doi.org/https://doi.org/10.3916/C65-2020-03>
- Salmerón Majadas, S. (2018). *A Methodological Approach based on Machine Learning to Generate a Multimodal User's Affective State Model in Adaptive Educational Systems*.
- Silva, J., Usart, M., & Lázaro-Cantabrana, J. L. (2019). Teacher's digital competence among final year Pedagogy students in Chile and Uruguay. *Comunicar*, 27(61), 31–40. <https://doi.org/10.3916/C61-2019-03>
- Suárez-Guerrero, C., Rivera-Vargas, P., & Rebour, M. (2020). Preguntas educativas para la tecnología digital como respuesta. *EduTec. Revista Electrónica de Tecnología Educativa*, 73, 7–22. <https://doi.org/10.21556/edutec.2020.73.1733>
- Tarik, A., Aissa, H., & Yousef, F. (2021). Artificial intelligence and machine learning to predict student performance during the COVID-19. *Procedia Computer Science*, 184, 835–840. <https://doi.org/10.1016/j.procs.2021.03.104>
- Touron, J., Martin, D., Navarro Asencio, E., Pradas, S., & Inigo, V. (2018). Validation de constructo de un instrumento para medir la competencia digital docente de los profesores (CDD). *Revista Espanola de Pedagogia*, 75(269), 25–54. <https://doi.org/10.22550/REP76-1-2018-02>
- UNESCO. (2019). *CONSENSO DE BEIJING sobre la inteligencia artificial y la educación*. 26–39. <https://unesdoc.unesco.org/ark:/48223/pf0000368303?posInSet=3&queryId=N-EXPLORE-e2652983-c3e5-4aba-9a2e-13341d1392aa>
- UNESCO. (2021). Marco de competencias docentes en materia de TIC. In *UNESCO Publishing*.

Para citar este artículo:

Forero-Corba, W., y Negre Bennasar, F. Diseño y simulación de un modelo de predicción para la evaluación de la competencia digital docente usando técnicas de Machine Learning. *EduTec, Revista Electrónica de Tecnología Educativa*, (89) 18-43. <https://doi.org/10.21556/edutec.2024.89.3201>

7. ANEXOS

7.1. Anexo I. Niveles por competencias TIC



Nota. Tomado y adaptado de Ministerio de Educación Nacional (2013)

7.2. Anexo II. Rúbrica Competencia TIC docente

La siguiente rúbrica evalúa el nivel de la competencia digital docente en Colombia (Mantilla Contreras, 2022).

En la rúbrica aparecen descriptores organizados en cuatro niveles de competencia digital: Nivel Principiante (N1), Nivel Medio (N2), Nivel Experto (N3) y Nivel Transformador (N4). Debe marcar solo uno de los niveles en los que se encuentra.

Competencia Pedagógica	
Descriptor 1.1. Planificación docente y Competencia Digital	
N1: Planifica actividades de (EA) donde contempla el uso de las tecnologías digitales.	
N2: Incorpora en la planificación didáctica: la búsqueda, tratamiento, almacenamiento y difusión de la información digital en diferentes formatos.	
N3: Incorpora en la planificación didáctica el buen uso de las tecnologías digitales orientadas hacia al manejo pertinente de información, la publicación de productos y el trabajo colaborativo.	
N4: Diseña actividades que favorecen el desarrollo de competencias (transversales, orientadas al trabajo autónomo y colaborativo) que impliquen habilidades complejas: resolver problemas, inferir, comunicar..., donde se haga un uso especializado de las tecnologías digitales.	
Descriptor 1.2. Las tecnologías digitales como facilitadoras del aprendizaje	
N1: Utiliza herramientas digitales de apoyo en el aula para la realización de actividades de (EA).	
N2: Desarrolla actividades con los estudiantes que explican la resolución de problemas de forma colaborativa mediante el uso de recursos tecnológicos digitales.	
N3: Realiza actividades con los estudiantes que implican: el análisis de un problema en grupo, la propuesta de soluciones alternativas, la negociación de los resultados y su publicación haciendo uso complejo de los recursos tecnológicos digitales.	
N4: Estimula el aprendizaje autónomo y el trabajo colaborativo mediante procesos de investigación que conducen a la transformación y creación de conocimientos haciendo uso de recursos tecnológicos digitales.	
Descriptor 1.3. Estrategias metodológicas para la integración de las TIC	
N1: Reconoce las orientaciones de la unidad académica para la incorporación de las tecnologías digitales en el aula y las tiene en cuenta en sus programaciones didácticas.	
N2: Desarrolla actividades donde incorpora los recursos digitales en función de las orientaciones metodológicas y de los recursos de los que dispone la institución.	
N3: Incorpora los recursos digitales en actividades significativas (funcionales, transversales y que favorecen la autonomía) que implican el uso de las tecnologías digitales para construir y compartir el conocimiento.	
N4: Propone nuevas estrategias metodológicas innovadoras y sirve de modelo docente para la integración de las TIC.	
Descriptor 1.4. Práctica pedagógica	
N1: Utiliza las evaluaciones de los estudiantes para reflexionar acerca de las estrategias pedagógicas mediadas por TIC llevadas al aula.	
N2: Implementa en las prácticas pedagógicas estrategias mediadas por TIC teniendo en cuenta disposiciones de la unidad académica que se han obtenido a partir de la evaluación entre pares.	
N3: Diseña ambientes virtuales de aprendizaje para la institución que surgen de la reflexión de su práctica de aula.	
N4: Transforma sus prácticas pedagógicas a partir de investigaciones de la reflexión de su práctica pedagógica.	

Competencia Tecnológica	
Descriptor 2.1. Ambientes de aprendizaje y espacios con tecnologías digitales	
N1: Utiliza las tecnologías digitales de aula: Pizarra digital interactiva (PDI), dispositivos fijos y móviles, en función de cada situación de (EA).	
N2: Adecua las actividades de (EA) a partir de la incorporación de las tecnologías digitales en los espacios académicos disponibles en la unidad académica.	
N3: Modifica los espacios de (EA) con tecnologías digitales para mejorar y optimizar la infraestructura disponible a partir de unos criterios institucionales.	
N4: Utiliza herramientas digitales especializadas para diseñar ambientes virtuales innovadores que favorecen espacios de (EA).	
Descriptor 2.2. Infraestructuras tecnológicas digitales.	
N1: Hace un uso responsable de los espacios y tecnologías digitales, utilizando un protocolo para resolver incidencias.	
N2: Incorpora a su práctica innovaciones haciendo un uso responsable de los espacios y tecnologías digitales, evaluando los riesgos de su implementación.	
N3: Resuelve incidencias que pueden surgir de la implementación de innovaciones o de la utilización de las tecnologías digitales y hace propuestas de mejora para promover su utilización.	
N4: Gestiona el uso de las tecnologías digitales y promueve el diseño de recursos para brindar las orientaciones necesarias para acceder a la infraestructura tecnológica y resolver los problemas.	
Descriptor 2.3. Acceso libre a la información, creación y difusión de material didáctico con licencias abiertas.	
N1: Comparte materiales didácticos elaborados y distribuidos en abierto en la red.	
N2: Elabora materiales didácticos abiertos y los comparte en la red siguiendo un estándar que facilite la búsqueda y accesibilidad.	
N3: Organiza los Recursos Educativos Abiertos (REA) por tipo y áreas en función de las necesidades de la unidad académica.	
N4: Potencia el uso de (REA) mediante la creación y/o difusión de repositorios abiertos de materiales didácticos.	
Descriptor 2.4. Entorno personal de aprendizaje (PLE)	
N1: Utiliza diferentes aplicaciones de escritorio y web para gestionar los contenidos del aula y acceder a la información.	
N2: Configura su (PLE) utilizando herramientas digitales para el aprendizaje, fuentes de información y red personal de aprendizaje.	
N3: Colabora con los docentes de la unidad académica en la creación de sus (PLE).	
N4: Asesora en el uso de los (PLE) en la unidad académica.	
Descriptor 2.5. Gestión de tecnologías digitales y aplicaciones.	
N1: Identifica los espacios con tecnologías digitales, selecciona los recursos y las herramientas existentes para el trabajo en el aula.	
N2: Utiliza de manera responsable y segura, los espacios, recursos y herramientas adecuadas para las diferentes situaciones de (EA).	
N3: Combina el uso de diferentes tecnologías digitales e incorpora innovaciones en función de su potencialidad analizando de forma reflexiva el rendimiento de los estudiantes a partir de su utilización.	
N4: Investiga situaciones de (EA) basadas en la utilización de los espacios y tecnologías digitales e innova a partir de los resultados obtenidos.	
Descriptor 2.6. Ética y seguridad	
N1: Respeta los derechos de autor en sus materiales docentes y utiliza las tecnologías digitales personales de forma responsable y segura.	
N2: Fomenta el uso ético de las tecnologías digitales siendo un referente durante las actividades con los estudiantes.	
N3: Sirve de referente para otros profesionales sobre el uso responsable y seguro de las tecnologías digitales.	
N4: Promueve estudios acerca del uso responsable, ético y seguro de las tecnologías digitales, que den respuesta a las necesidades y fomenten soluciones novedosas.	

Competencia Comunicación	
Descriptor 3.1. Comunicación, difusión y transferencia del conocimiento	
N1: Utiliza herramientas digitales para comunicarse y compartir sus conocimientos con otros docentes.	
N2: Gestiona recursos abiertos en red para publicar sus experiencias y compartirlas.	
N3: Capacita docentes, mediante actividades reconocidas por la institución, en el uso de tecnologías digitales para compartir y crear conocimientos.	
N4: Sirve de referente en el uso de recursos tecnológicos para difundir y compartir su conocimiento, transformando la institución educativa en una unidad académica de innovación al servicio de la comunidad.	
Descriptor 3.2. Gestión de contenidos digitales	
N1: Accede y comenta los contenidos distribuidos en diferentes espacios digitales.	
N2: Utiliza los espacios digitales de la unidad académica como editor de alguno de ellos con el objetivo de compartir conocimientos y experiencias.	
N3: Utiliza un espacio digital propio como medio para publicar y difundir su conocimiento profesional y hacer participar a la comunidad educativa.	
N4: Gestiona espacios virtuales para difundir el conocimiento colectivo y para fortalecer la comunicación y la interacción entre los miembros de la comunidad educativa.	
Descriptor 3.3. Colaboración mediante herramientas digitales	
N1: Orienta a nivel de aula actividades de (EA) enfocadas a la colaboración mediante la creación de documentos compartidos que fortalecen el trabajo en equipo.	
N2: Utiliza las tecnologías digitales con los estudiantes en situaciones de (EA) para elaborar productos colaborativos a través de canales digitales diversos.	
N3: Capacita a los docentes de la unidad académica o institución en el uso de medios digitales de colaboración.	
N4: Utiliza las herramientas digitales de colaboración para investigar acerca de la usabilidad para favorecer procesos de (EA), compartiendo los resultados a través de redes profesionales.	
Descriptor 3.4. Comunidades de aprendizaje virtuales: formales, no formales e informales	
N1: Utiliza materiales docentes compartidos en red para la actividad académica en el aula.	
N2: Utiliza el aprendizaje en red como medio de formación permanente y promueve la participación de estudiantes en comunidades virtuales.	
N3: Promueve la participación de estudiantes en comunidades virtuales como estrategia de aprendizaje colaborativo y fomenta el aprendizaje en red entre los miembros de la unidad académica.	
N4: Gestiona un ecosistema de aprendizaje digital entre los miembros de la unidad académica y otras instituciones en la consolidación de comunidades y/o redes de aprendizaje interinstitucionales.	

Competencia Gestión	
Descriptor 4.1. Evaluación, tutoría y seguimiento de los estudiantes	
N1: Utiliza recursos digitales para la tutoría y seguimiento de los estudiantes (registrar asistencia y valoraciones, enviar mensajes, ...)	
N2: Utiliza escenarios educativos digitales para evaluar, hacer seguimiento al aprendizaje de los estudiantes y mejorar los procesos de realimentación.	
N3: Utiliza recursos digitales para compartir la evaluación y el seguimiento de los estudiantes con otros profesionales de la unidad académica (facultad, departamento, escuela, instituto, corporación, carrera, posgrado, ...), para mejorar los resultados teniendo en cuenta el despliegue de estrategias didácticas digitales.	
N4: Gestiona y utiliza recursos digitales (entornos, portafolios digitales, ...) para hacer el seguimiento escolar, la evaluación de los estudiantes a nivel de la unidad académica y utilizar los resultados para investigar la incorporación de las TIC en los procesos de (EA).	
Descriptor 4.2. Evaluación y recuperación de contenidos digitales	
N1: Reconoce las diferentes fuentes de información y contenidos digitales para la docencia y almacena los recursos disponibles.	
N2: Orienta en su práctica de aula, la valoración de contenidos digitales y los organiza	

Competencia Gestión	
considerando diferentes tipos de almacenamiento.	
N3: Evalúa la calidad de los recursos y diferencia las licencias que permiten reutilizarlos para incorporarlos a nivel de aula y la unidad académica.	
N4: Capacita a los docentes de la institución para evaluar de manera crítica las fuentes de información, perfiles académicos y comunidades a las cuales pertenece.	
Descriptor 4.3. Identidad digital de la institución	
N1: Conoce la identidad digital de la unidad académica, la necesidad de respetar los modelos documentales y protocolos relativos a la identificación visual.	
N2: Incorpora a sus documentos y espacios virtuales la identificación visual de la unidad académica e institucional.	
N3: Participa en el mejoramiento de la imagen institucional en los espacios virtuales de la unidad académica.	
N4: Transforma los espacios virtuales de la unidad académica velando por transmitir una identidad digital de la propia institución.	
Descriptor 4.4. Gestión de la Institución	
N1: Identifica elementos de la gestión de la unidad académica, que pueden ser mejorados a partir de la incorporación de las tecnologías digitales, en las diferentes actividades realizadas.	
N2: Adopta políticas de la institución en beneficio de la unidad académica, para el uso de tecnologías digitales que contribuyen a la dinamización de procesos de gestión.	
N3: Propone acciones para optimizar la gestión institucional a partir del uso de las tecnologías digitales que permiten integrar los procesos y hacerlos más efectivos.	
N4: Lidera acciones para optimizar los procesos integrados a la gestión institucional desde el uso de las tecnologías digitales.	
Descriptor 4.5. Formación Permanente	
N1: Participa en actividades de formación profesional que son orientadas por la unidad académica, reconocidas por la administración educativa y están relacionadas con las tecnologías digitales.	
N2: Gestiona su plan de formación de manera permanente, de acuerdo a las necesidades de desarrollo profesional, mediante actividades de educativas que están relacionadas con las tecnologías digitales y la innovación educativa, y son reconocidas por la administración educativa.	
N3: Transforma su práctica docente, mediante la incorporación de las tecnologías digitales a la misma, incorporando los conocimientos obtenidos en las actividades de formación: "transferencia de la formación".	
N4: Participa como formador en actividades de capacitación permanente del profesorado relacionadas con las tecnologías digitales.	
Descriptor 4.6. Participación ciudadana	
N1: Participa en los diferentes escenarios digitales promoviendo acciones socialmente responsables.	
N2: Realiza actividades pedagógicas que favorecen el empoderamiento de las tecnologías y promoción de una ciudadanía digital en los estudiantes.	
N3: Promueve actividades de capacitación a en la unidad académica con el propósito de promover la participación responsable en el mundo digital.	
N4: Desarrolla proyectos con otros profesionales de la comunidad educativa hacia la participación y promoción de la ciudadanía digital.	
Descriptor 4.7. Inclusión digital	
N1: Potencia el acceso y uso de las tecnologías digitales por parte de todos los estudiantes con la intención de compensar las desigualdades.	
N2: Participa en la organización de la atención a la diversidad de la unidad académica ejerciendo acciones para compensar las desigualdades a nivel de acceso y uso de las tecnologías digitales.	
N3: Promueve la utilización de los espacios y recursos tecnológicos digitales de la unidad académica mediante la participación en acciones orientadas a la compensación de las desigualdades.	
N4: Capacita a los integrantes de la unidad académica con acciones orientadas a la generalización del uso, la gestión y la difusión de buenas prácticas en el uso de las tecnologías digitales.	

Competencia Investigativa	
Descriptor 5.1. Tratamiento de la información y creación de conocimiento	
N1: Enseña a realizar búsquedas de información accediendo a diferentes fuentes de diversatipología, teniendo en cuenta procedimientos definidos.	
N2: Enseña a utilizar fuentes de información de diferente tipología atendiendo a criterios decalidad, veracidad y pertinencia.	
N3: Enseña a clasificar, ordenar y seleccionar la información desde diferentes fuentes aplicandocriterios de calidad, veracidad y pertinencia, promoviendo una valoración crítica de los recursosdigitales.	
N4: Enseña a crear y transformar la información en conocimiento, que previamente se ha almacenado y recuperado, siguiendo un sistema que le permita hacer un uso compartido.	
Descriptor 5.2. Proyectos de incorporación de las tecnologías digitales	
N1: Sigue las directrices acordadas a nivel de unidad académica sobre el uso de las tecnologíasdigitales en la docencia.	
N2: Forma parte activa de los equipos de la unidad académica y aporta su experiencia y conocimientos sobre las tecnologías digitales.	
N3: Lidera un equipo de trabajo de la facultad o unidad académica haciéndose cargo de gestionarla utilización de las tecnologías digitales en las prácticas pedagógicas.	
N4: Coordina proyectos interdisciplinarios o interinstitucionales en torno a la incorporación de lastecnologías digitales en la docencia.	
Descriptor 5.3. Liderazgo en el uso de las tecnologías digitales	
N1: Utiliza las tecnologías digitales con los estudiantes haciendo de referente en cuanto a su uso.	
N2: Integra las tecnologías digitales a la docencia y comparte experiencias con los colegas, siendoun referente de incorporación TIC.	
N3: Coordina el uso de las tecnologías siendo un modelo de curación de contenidos y recursosdigitales a nivel de la unidad académica.	
N4: Asesora a la unidad académica en la utilización y gestión de las tecnologías digitales ydifunde la experiencia y las buenas prácticas.	
Descriptor 5.4. Identidad y presencia digital	
N1: Usa la identificación digital profesional en las comunicaciones y mantiene su perfilactualizado en los espacios virtuales de la unidad académica.	
N2: Tiene un perfil digital y un currículum profesional actualizado online.	
N3: Utiliza las redes sociales y profesionales como medio de comunicación e interacción profesional, que le facilita conectarse con otros investigadores.	
N4: Utiliza las redes para el desarrollo profesional en la divulgación de resultados de las investigaciones y promueve el uso e importancia a los miembros de la unidad académica.	



Educational innovation: Exploring the potential of Generative Artificial Intelligence in cognitive schema building

Innovación educativa: Explorando el potencial de la Inteligencia Artificial Generativa en la construcción de esquemas cognitivos

 Bernarda Salgado Granda; bsalgadog@ups.edu.ec; Universidad Politécnica Salesiana (Ecuador)

 Yana Inzhivotkina; yana.inzhivotkina@ug.edu.ec; Universidad de Guayaquil (Ecuador)

 María Fernanda Ibáñez Apolo; maria.ibanezapolo0916@upse.edu.ec; Universidad Estatal Península de Santa Elena (Ecuador)

 Jorge Gustavo Ugarte Fajardo; jugartef@dmgscotec.edu.ec; Universidad Ecotec (Ecuador)

Abstract

This study explores the use of generative artificial intelligence to enhance teaching and learning experience, focusing on strengthening and consolidating cognitive schemas. Research reveals that schemas can profoundly influence the improvement of the learning experience and promote the assimilation of new types of information and retention in students' memory. To improve the teaching and learning experience, the advantages, obstacles, and potential future trajectories of utilizing these technologies were examined by conducting a thorough literature review and analyzing relevant studies. Findings indicate that generative artificial intelligence has the potential to personalize learning, diversify educational content, and improve teaching efficiency and scalability. However, it also poses challenges related to content quality, data privacy, and equity in access to personalized learning. Future research should focus on the effectiveness of educational tools based on generative AI that promote equity and inclusion, ethical approaches, and interdisciplinary collaboration. Overall, this study provides a solid foundation for understanding and harnessing the potential of generative artificial intelligence in enhancing cognitive schemas, thereby promoting more effective, inclusive, and personalized education.

Keywords: Generative Artificial Intelligence, Cognitive Schemas, Technology-Assisted Learning, Personalized Learning, Education.

Resumen

Este estudio explora el uso de la inteligencia artificial generativa para mejorar la experiencia de enseñanza y aprendizaje, centrándose en el fortalecimiento y consolidación de los esquemas cognitivos. Las investigaciones revelan que los esquemas cognitivos mejoran la experiencia de aprendizaje y favorecen la asimilación de nuevos tipos de información y su retención en memoria de los estudiantes. Mediante una exhaustiva revisión de literatura y análisis de estudios relevantes, se examinaron las ventajas, obstáculos y posibles trayectorias futuras de la utilización de estas tecnologías para mejorar la enseñanza. Los hallazgos muestran que la IA generativa puede personalizar el aprendizaje, diversificar el contenido educativo, mejorar la eficiencia y escalabilidad de la enseñanza, pero también plantea desafíos relacionados con la calidad del contenido, la privacidad de los datos y la equidad en el acceso al aprendizaje personalizado. Futuras investigaciones deberían centrarse en la efectividad del aprendizaje personalizado, el desarrollo de herramientas de IA generativa para la educación, la promoción de la equidad y la inclusión, la exploración de enfoques éticos y la colaboración interdisciplinaria. Este estudio proporciona una base sólida para comprender y aprovechar la IA generativa en la mejora de los esquemas de aprendizaje, promoviendo una educación más efectiva, inclusiva y personalizada.

Palabras clave: Inteligencia Artificial Generativa, Esquemas Cognitivos, aprendizaje asistido por tecnología, Aprendizaje personalizado, Educación.



1. INTRODUCTION

In the educational context, the schema concept emerges as a central element in the knowledge construction process and offers valuable insights to improve the teaching-learning process. Schemas, defined as mental structures that organize and process information, are pivotal in facilitating deeper understanding, solid knowledge retention, and their application in various contexts. Understanding cognitive processes and robust schema formation facilitates knowledge assimilation and stimulates the human brain's ability to creatively address problems, foster critical thinking, and adapt to complex contexts, resulting in more efficient and effective learning (Parrales et al., 2020).

On the other hand, the exponential growth of artificial intelligence technologies has revolutionized students' learning experiences, offering new tools with enhanced capabilities to transform the teaching-learning process (Tyagi et al., 2022). Schools and universities are changing their educational models to adapt to new technologies and offer programs that incorporate artificial intelligence (Sanabria-Navarro et al., 2023). This interest primarily stems from their potential to create personalized educational content, generate immersive learning experiences, and promote academic proficiency (Pendy, 2021). New Generative AI models focus on creating new and realistic content using existing digital content from a specific domain, such as video, images/graphics, text, and audio, which have immense potential to enhance educational processes (Yu & Guo, 2023). This cohesion between learning schemas and artificial intelligence technologies fosters more dynamic, accessible, and effective learning environments (Ahmad et al., 2023).

In the current context, education faces challenges due to rapid technological progress, socioeconomic changes affecting equity, and the increasing number of students with different learning styles and needs (Ubah et al., 2022). This complexity leads to continuous changes in the roles of educators and educational systems to adapt and meet the changing workforce demands of society and prepare students with the necessary competencies in an ever-evolving world (Maj-Waśniowska et al., 2022).

One of the most pressing challenges is improving students' cognitive schemas. Cognitive education focuses on enhancing the quality of education by addressing systematic and logical thinking, metacognitive processes, and the development of learning tools such as abstraction, inference, relation, comparison, and judgment, which are essential for effective learning (Tzuriel, 2021). However, in a dynamic and diverse educational environment, traditional teaching methods may not be sufficient to develop robust and adaptable schemas.

Therefore, there is a need for innovative methods and educational tools that can effectively enhance the building of schemas in an ever-changing learning environment (Darejeh et al., 2022). It is crucial to explore new strategies and technologies that can boost the formation of more efficient, flexible, and personalized learning schemas by adapting to individual students' needs and promoting a deep and lasting understanding of knowledge (Voskoglou, 2020). Generative AI technologies emerge as promising tools with significant potential to transform the teaching-learning process by strengthening the formation and consolidation of learning schemas.

The main objective of this study is to examine the use of generative AI tools to personalize and enhance the teaching-learning process, as well as to strengthen students' cognitive schemas, facilitating the assimilation and retention of knowledge. We examine the role of generative artificial intelligence in the face of the growing need to explore new tools to create more effective and adaptable learning systems that strengthen cognitive schemas in an educational environment characterized by student diversity, rapid technological evolution, and changes in the labor market demands.

This work contributes to knowledge advancement in the educational field by providing insights and recommendations that can guide future research on the application of generative artificial intelligence in the learning process.

1.1. Cognitive Schemas

In the field of cognitive psychology, schemas are defined as organized mental structures or patterns that organize and interpret information, facilitating knowledge construction through the deep understanding, interpretation, and long-term retention of information (Torney-Purta, 1991; Soriano, 2017). These schemas are developed from experience and prior knowledge and play a crucial role in assimilation, interpretation, and information retention (Derry, 1996).

As individuals interact with their environment and access new information, their schemas adjust to accommodate new experiences and situations. This continuous adaptation allows for a deeper and more comprehensive understanding of the world around them, facilitating the learning process and cognitive development (Brewer, 2000).

The relevance of cognitive schemas lies in their ability to organize process, understand, and remember information. By developing schemas in a specific knowledge domain, individuals can process information more efficiently, identify patterns, establish connections, and reach conclusions swiftly (Moreira et al., 2002). Furthermore, schemas also influence how individuals assimilate and retain knowledge, enhancing their ability to use and apply it in different contexts and resolve problems creatively (McVee et al., 2005).

In his study, Hejny (2018) proposed schema-oriented education to improve mathematics teaching. This approach was based on constructing schemata to create a dynamic network of mathematical knowledge and skills for the student.

According to Akram (2019), the human brain can store various types of knowledge in the form of schemas that are retained in long-term memory. These schemas are crucial in improving the learning experience and consolidating new knowledge in students' memory. To improve schema formation in the classroom, teachers should incorporate activities that stimulate the connection between new information and existing knowledge and provide students with the learning environment to apply new knowledge in real-world situations.

Furthermore, Dang et al. (2019) argued that individuals employ mental structures to orient themselves in time and space, as well as to categorize and make sense of their experiences in a relevant manner. The consolidation of these schemas can be enhanced through the repetition of experiences and reinforcement of certain beliefs.

In another study, Sweller (2020), based on the theory of cognitive load, proposed recommendations regarding the design of instructional procedures that facilitate the construction of high-quality knowledge schemas in long-term memory. Most instructional procedures can be implemented more effectively through technology-based learning environments that support constructive activity, metaconceptual awareness, and cognitive flexibility. Sweller suggests that the principal learning mechanisms are schema acquisition and automation, when considering intellectual activities. Thus, he deems it necessary to reduce extraneous cognitive load in instructional design to mitigate its effects on schema formation during the learning process. An overview of the primary effects of cognitive load in technology-assisted instructional design is shown in Table 1.

Table 1

Relevant effects in technology-assisted instructional design (Sweller, 2020)

Instructional Effect	Description
Worked	Example Studying worked examples is superior to solving the equivalent problems
Split-attention	If multiple sources of information need to be considered simultaneously, physically integrating them is superior to requiring learners to split their attention between them
Modality	If a diagram and text need to be considered simultaneously and the text is simple and short, presenting the text in spoken rather than written form is superior
Transient	High element interactivity information should be presented in permanent rather than transient form or presented in smaller chunks
Redundancy	Eliminating unnecessary information results in superior learning
Expertise reversal and element Interactivity	With increases in expertise and decreases in element interactivity, information that is essential for novices becomes redundant for more expert learners, decreasing learning
Working memory depletion	Working memory use depletes working memory resources that recover after rest

Jung et al. (2022) presented a schema-based instructional design model for self-paced learning environments. In their study, they highlighted the importance of the schema-based approach, which involves an initial analysis of students' needs and the learning environment, followed by the creation of a knowledge map to classify and visualize knowledge structure. Subsequently, in the design and development phase, instructional strategies were proposed to promote schema activation, construction, and automation and to update existing schemas.

The schema concept provides an understanding of how students organize and assimilate information, enabling the design of more effective educational activities. Brain functions are closely related to cognitive schemas, facilitating information processing and interpretation.

1.2. Generative Artificial Intelligence

Generative artificial intelligence is a subtype of artificial intelligence that can generate innovative content across diverse formats, including images, text, or music (Lv, 2023). This technology enhances productivity and creativity, automates content creation, and aids in

problem solving. These models are trained on large amounts of information, allowing them to capture complex relationships and underlying structures in the data and use this information to generate meaningful and valuable outcomes (Feuerriegel et al., 2023).

Since 2013, generative AI has experienced a remarkable impulse thanks to the introduction of innovative neural network designs like Variational AutoEncoder (VAE), Generative Adversarial Network (GAN), and diffusion models. These models highlighted a high capacity to create data in various formats, like text, images, videos, and voice, and to generate synthetic data mimicking datasets. In 2017, Vaswani A. proposed a novel neural network architecture—attention-layer-based transformer and positional encoding—that incorporates a positional token containing information about the word order in the input (Vaswani et al., 2017). In 2018, two transformer-based models made their debut: Google introduced BERT, and OpenAI unveiled the first pre-trained generative transformer model (GPT-1). Other notable LLMs introduced afterward include GPT-4 (by OpenAI), PaLM (by Google), GPT-Neo (by EleutherAI), and LaMDA (by Google), among others, garnering significant and widespread attention. Large Language Models (LLMs) have emerged as game-changing innovations, largely because of their ability to grasp intricate connections between words and phrases in the input text.

As research in generative AI advances, Large Multimodal Models (LMMs) have gained significant attention. These models stand out from their predecessors by accepting multiple data types, such as text, images, video, code, and audio (Gozalo-Brizuela & Garrido-Merchán, 2023). This helps them provide more accurate insights and interpretations of the complex and varied data we encounter in the digital realm. Multimodal systems can receive information from multiple sources, giving them a deeper understanding of context and making them more reliable and high-quality (Suzuki & Matsuo, 2022). In 2021, OpenAI introduced DALL-E, an LMM capable of generating images from text. In 2022, several other models with multimodal capabilities were unveiled, such as PaLM-E (by Google), GPT-4 Vision (by OpenAI), Stable Diffusion (by Runway ML), Midjourney V6 (by Midjourney), Gemini (by Google), Gopher (by DeepMind), Claude (Anthropic), Codex (by OpenAI), and Vall-E (by Microsoft). The most recent language model released by OpenAI in May 2024, GPT-4o, represents a significant evolutionary advancement in the development of generative artificial intelligence models. This model constitutes an improved version of GPT-4, offering native multimodal capabilities and reduced latency compared to its predecessor. One of the most prominent aspects of GPT-4o is its enhanced reasoning capacity, which enables deeper understanding and the generation of more coherent responses. Furthermore, the proposed model offers improvements in terms of quality control and security, which increases the reliability of content generation. Additionally, the GPT-4o's ability to perform real-time analysis of images and videos captured by the camera expands the range of applications and enables more natural and automated interactions (OpenAI, 2024).

The creativity produced by this technology has had a significant impact across various industries. Large Language Models (LLMs) have demonstrated proficiency in executing tasks related to natural language with precision and coherence, including generating high-quality creative text, advanced automatic translation, contextually relevant and coherent dialog responses, sentiment analysis, text classification, and summaries (Epstein et al., 2023). In addition, generative AI-based image generators, commonly referred to as multimodal models (LMMs), have demonstrated exceptional capabilities in producing images and videos from textual descriptions. These models are used to create innovative designs, create stunning visual

effects, and render realistic scenes. Additionally, they are used to compose original music and generate images with remarkable precision (Fareed et al., 2024).

In the educational field, generative AI possesses significant potential for creating and delivering educational content. In addition, these technologies can generate a wide variety of resources, including images, videos, and interactive applications, which can be adapted to students' learning needs and their particular learning styles. The disruption caused by Generative AI in all spheres of human life makes it imperative to instruct educators and students on how to handle and formatively use these tools (Mills et al., 2023). This academic training could include generating exercises and problems adapted to each student's skill level and comprehension or crafting explanations and examples to address specific areas of confusion or difficulty (Menekse, 2023). Generative AI tools can also facilitate the creation of interactive learning environments and immersive learning experiences with personalized feedback that encourages active participation and student engagement. Such technologies can generate realistic simulations, games, and practice activities that allow students to explore concepts practically and experience the learning process interactively, adaptively, and informatively (Chheang et al., 2024).

However, the misuse of Generative AI can compromise the quality of content and raise ethical and legal concerns regarding the outcomes obtained. Accuracy and relevance evaluation of AI content generated is crucial to overcome these challenges and address potential biases. In addition, guidelines and policies should be established to regulate the use of AI-generated content, focusing on the quality and originality of ideas. Collaboration among educators, researchers, and professionals is essential for developing best practices and guidelines for AI model usage that promote the ethical and responsible use of Generative AI in the educational field (Chen et al., 2023).

2. METHOD

In this study, a systematic review was conducted according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement to provide transparent and high-quality results presented in a clear and reproducible manner. The selected studies, which focused on improving cognitive processes, provided exhaustive and up-to-date details on the application of Generative AI and empirical results supported by substantiated conclusions. The most relevant findings obtained from the data analysis were synthesized and summarized.

In the literature search, rigorous selection criteria were implemented to ensure the comprehensiveness and quality of the research on the application of Generative AI in the learning process. The search for relevant studies in the Scopus database was conducted using Boolean combinations of the following keywords: SCHEMA, SCHEME-ORIENTED, GENERATIVE ARTIFICIAL INTELLIGENCE, GENERATIVE AI, GENERATIVE MODEL, LEARNING, COGNITIVE, located in the title, abstract, and keywords of the articles, covering original journal articles and reviews. In total, 255 articles were identified using the search strategy.

Temporal filters were established to limit the search to studies published within the last five years, thus ensuring the relevance and timeliness of the collected information. Furthermore, inclusion criteria were applied to select relevant studies for the analysis.

- Document types: scientific articles and reviews.
- Time frame: the last five years to ensure relevance and currency.
- Publication Stage: Article in press and final.
- Languages: English and Spanish.
- Data sources: Scopus database.

Based on the retrieved metadata, articles were excluded according to the following criteria: articles without authors, articles lacking DOIs, duplicate documents, and retracted articles. Subsequently, the abstracts and full texts were reviewed to confirm their relevance to the research topic. Because of this rigorous selection process, 41 valid records were obtained and subsequently used for analysis.

A term co-occurrence network was generated using the corpus of documents with VOSviewer software to examine the relationships between terms and identify the organization of subthemes within the research topic. VOSviewer was configured with the normalization method "Association Strength," attraction set to 2, and resolution set to 1.1 to visualize the co-occurrence network.

Drawing from the insights obtained through co-occurrence analysis, we examined the documents to identify use cases for generative AI. The most significant use cases of Generative AI for educational applications were derived from the selected information.

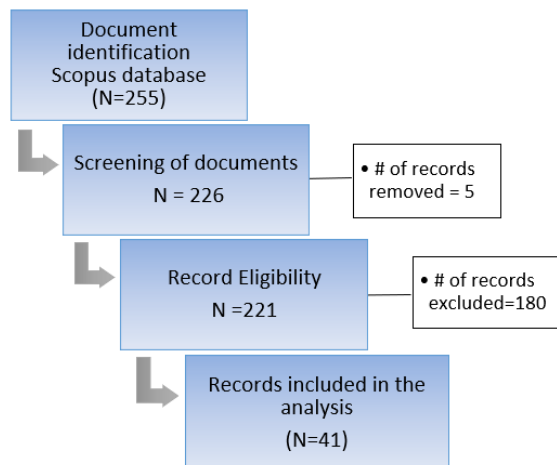
We identified use cases of Generative AI that can enhance the development of cognitive schemas. These cases focused primarily on improving conceptual understanding, tailoring content and activity difficulty to individual learning needs and styles, stimulating creativity and innovation, and fostering collaborative learning. These use cases can be integrated into the design of particular instructional strategies to promote schema activation, construction, and automation and update existing schemas.

Furthermore, we proposed several Generative AI tools suitable for implementing these use cases. These tools were identified through online research, considering criteria that align with the defined services in the use case classification.

One potential limitation in this systematic review is the rapid evolution of generative AI, which presents a significant challenge due to the frequent changes in tools and applications. This accelerated development can lead to the swift obsolescence of the IA studies reviewed, thereby limiting the longevity of several of our results.

Figure 1

PRISMA flow chart of research strategy



3. RESULTS

Developing the search strategy (Figure 1) resulted in 255 collected documents within the categories of original articles and reviews. After a thorough cleaning of the metadata, ineligible articles were excluded through an individual assessment of titles, keywords, abstracts, and text. This process yielded a refined dataset comprising 41 records, which were subsequently subjected to analysis.

The comprehensive literature review and relevant studies provided an integrated and detailed insight into the application of Generative AI in learning processes, aimed at strengthening cognitive schemas. The most relevant articles identified in the systematic literature review are detailed in Table 2, located in the Annexes section.

3.1. Analysis of Thematic Structure

The co-word analysis enabled us to explore the conceptual and thematic framework of the investigated scientific domain. Figure 2 presents the co-word network, highlighting the words with the highest frequency of occurrence in the analyzed documents and their interrelations.

Four clusters are identified, represented by the colors yellow, red, blue, and green.

In Cluster 1 (yellow), the most prominent terms are artificial intelligence, LLM, ChatGPT, generative AI, chatbot, prompt engineering, GPT model, game-based learning, and critical thinking. This clustering indicates that the primary focus is the integration and impact of advanced AI technologies in educational settings, with an emphasis on the use of Large Language Models (LLMs) and generative AI tools like ChatGPT for interactive learning experiences, as well as employing game-based learning strategies to enhance critical thinking skills.

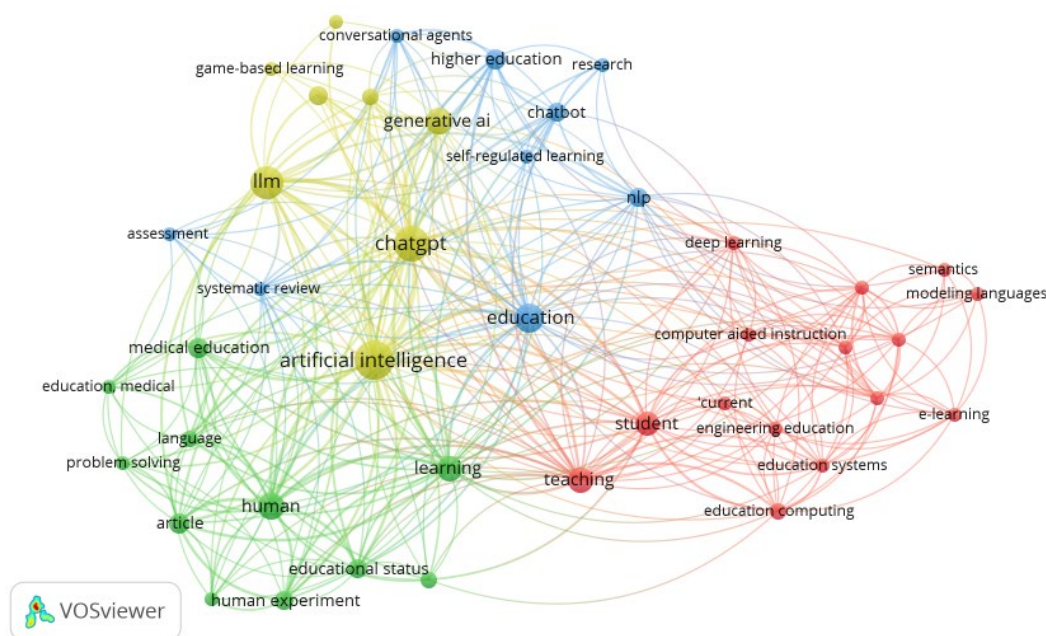
In Cluster 2 (red), the most relevant terms include teaching, student, engineering education, education computing, education systems, e-learning, deep learning, computer-aided instruction, language processing, modeling language, and semantics. These elements reflect the role of digital and computational methods in modern education through using deep learning and language processing in e-learning and computer-aided instruction, especially within engineering education.

Cluster 3 (blue) encompasses various terms, such as education, higher education, conversational agent, self-regulated learning, chatbot, NLP, and assessment. This cluster focuses on human-centric educational tools, emphasizing using conversational agents and chatbots powered by natural language processing (NLP) to facilitate self-regulated learning and assessment in higher education.

Cluster 4 (green) encompasses terms such as medical education, human, learning, problem-solving, human, human experiment, education status, simulation, and curriculum. The focus of this group is on medical and experiential education, highlighting the importance of human-computing relationships in learning. This includes using simulations and problem-solving.

Figure 2

Co-occurrence network of words



Based on the insights obtained from the co-occurrence analysis, we analyzed the documents to derive use cases for generative AI. Generative AI models are being implemented across a wide range of tools and applications in various industries. Generative AI models can create content in various categories, such as natural language processing, image processing, music, and video creation, that can be used to enhance educational processes by implementing tools for personalized experiences, realistic content generation, intelligent conversational dialogue, data analysis, and virtual instruction.

3.2. Guidance for Schema-Based Instructional Design

Given that teaching is intrinsically linked to the cognitive processes of creating and maintaining schemas in learning, establishing a schema-based instructional design process that considers the importance of schemas and cognitive load in learning new forms of knowledge and data within classroom teaching and learning contexts is crucial.

The stages of the process are described in detail below.

1. Context analysis

Context analysis serves as a foundational step in instructional design, ensuring that educational strategies are tailored to the specific needs, capabilities, and circumstances of learners and the learning environment, thereby enhancing the overall effectiveness of the educational process.

Needs Analysis: Assess the existing conditions and available resources necessary for optimal learning outcomes, and identify the gaps between the instructional goals and the current state.

Learner Analysis: Assess and identify students' prior knowledge and learning strategies to align them with learners' preferred learning styles. Additionally, determine learners' motivation, confidence, and proficiency in using technology, their capacity for self-directed learning, and their confidence in engaging in online communications.

Learning Environment (Context) Analysis: Determine the physical or virtual setting and logistics of learning. Establish whether the instruction will be online, face-to-face, or hybrid, and determine the group size for the instructional sessions.

Identify Potential Solutions and Feasibility Analysis: Develop strategies to address identified problems and achieve learning outcomes. Evaluate the financial and technical resources and constraints and determine the operational feasibility of implementing the instructional design.

2. Schema Analysis

Schema analysis serves as the pivot of the design process, encompassing the development of a schema model that facilitates the categorization and visualization of organizational knowledge.

Cognitive Task Analysis: Involve identifying the mental models and cognitive processes used by experts in the field. Behavioral task analysis captures observable behaviors, whereas procedural task analysis delves into the documentation of covert actions and the underlying conceptual frameworks.

Cognitive Task Analysis: Involve identifying the mental models and cognitive processes used by experts in the field. Behavioral task analysis captures observable behaviors, whereas procedural task analysis delves into the documentation of covert actions and the underlying conceptual frameworks.

Knowledge mapping: create a schema knowledge map that offers an efficient and effective method for categorizing and visualizing the organizational knowledge inherent in schemas.

3. Design and Development

Using compiled learning experiences and materials, instructional techniques and strategies should be framed to foster schema activation, construction, automation, modification, and elaboration. Table 3 provides an overview of the processes associated with the design and development phase and the applicable instructional strategies.

4. Evaluation

To measure students' mental efforts, cognitive load can be used as a proxy to assess the depth and quality of learners' understanding and determine their engagement with a given learning task. Assessments can be conducted using the following methods: self-explanation of acquired knowledge, a checklist approach, and concept mapping.

Table 3

Description of Schema-Based Instructional Design and Development Process

Stage	Description	Instructional Strategies
Schema activation	Schema Activation involves leveraging a learner's prior knowledge and experiences to expedite the learning process.	pre-reading activities, pre-organizers, advance organizers, previews, and thematic organizers.
Schema construction and Schema automation	Critical considerations: <ul style="list-style-type: none"> Utilizing goal-free problems to minimize extraneous cognitive load associated with means-end analysis. Employing worked examples to streamline means-end searches and enhance working memory capacity by reducing extraneous cognitive load. 	Schema construction strategies: Cases, interactive video-recordings, hierarchical concept maps and vee diagrams
	<ul style="list-style-type: none"> Offering completion problems that provide partial solutions, enabling learners to deduce remaining solutions. Minimizing split attention whenever feasible in instructional settings. Integrating visuals such as diagrams and presenting verbal information audibly to enhance learning. Eliminating redundant information and reducing working memory load. Implementing the variability effect to expose learners to different contexts of a concept. 	Schema automation strategies: Goal-free problems, worked examples, completion problems, split-attention effect, modality effect, redundancy effect and variability effect.
Schema modification	The schema modification process, known as accommodation, focuses on re-evaluating a learner's existing schema to enhance the efficient utilization of mental efforts.	Unlearning/relearning and Reflection
Schema elaboration	Schema elaboration involves the process of maturing, saturating, or refining acquired schemas through instruction in a different domain, which fosters connections among related schemas.	Pattern recognition and Parallel schema interaction

3.3. Specific use case of Generative AI in the learning process

The use of generative AI in education is a topic of growing interest because of its ability to develop functions as part of a human-AI cooperative and social learning system. The implementation of generative AI in education has the potential to redefine instructional methods and enhance student learning outcomes. This technology enables the personalization of educational experiences, interactive content creation, and the provision of adaptive assessments, stimulating greater student engagement and improved knowledge retention.

Recent studies have investigated the use of chatbots and advanced language models in the educational domain, focusing on student performance monitoring and evaluation, and interactive and engaging educational resource creation. In addition, generative AI has demonstrated proficiency in enhancing educational strategies, such as personalized learning experiences and interactive content creation. Using generative AI in open collaboration environments allows for emerging approaches that can help educators navigate and thrive in a rapidly evolving AI era.

Below are the educational use cases of generative AI that enhance learning schemas, primarily aimed at improving conceptual understanding, adapting content and activity difficulty according to individual learning needs and styles, stimulating creativity and innovation, and fostering collaborative learning. Table 4 presents the generative AI-Based Instructional Strategies classified according to the stages of the schemes.

3.3.1. Learning personalization

Generative AI can create personalized educational content tailored to individual student needs. This approach involves adjusting learning materials, creating customized learning paths, and developing exercises and problems that align with each student's skill level and comprehension.

IA tools: Rewordify and DreamBox Learning.

3.3.2. Engaging educational content creation

In education, Generative AI can produce relevant and up-to-date educational content that promotes student engagement, including text, videos, images, and interactive applications. These resources can be automatically generated from large datasets.

IA tools: ChatGPT, Bard, Copilot, Dall-E 2, Midjourney, Stable Diffusion, Pictory AI, Synthesia, and Gamma.

3.3.3. Active-learning facilitation

Generative AI tools can drive active-learning by generating interactive and participatory learning experiences. These technology-enriched environments encompass simulations, games, and practical activities that allow students to explore concepts actively.

IA tools: Course Hero, Vortex Studio, Labster, Nearpod, PhET, and Kahoot.

3.3.4. Support for teaching and tutoring

Generative AI tools can generate complementary educational resources, provide feedback and automated assessment of student progress, assist in setting up adaptive learning environments, and analyze student performance data to identify areas needing support (Amani et al., 2023). This enables educators to focus on personalized activities for students and dedicate more time to analyzing individual academic progress. Generative AI tools can help educators become more efficient and productive.

IA tools: M-Powering Teachers, Rewordify, Gradescope, and Turnitin.

Despite the promising progress of generative AI, integrating this emergent technology into education involves addressing concerns related to data privacy, copyright, and algorithmic bias. A major challenge is its impact on the academic integrity of scholarly work, because of its ability to generate responses similar to those of humans in all domains. Furthermore, it requires responses to concerns regarding the opacity and accuracy of the generated data. Many of these issues will be mitigated by the advanced capabilities of new multimodal models, such as the recently launched GPT-4o.

Table 4

Generative AI based Instructional Strategies

Use cases of Generative AI	Schema Activation	Schema Construction	Schema Automation	Schema Modification	Schema Elaboration
Learning Personalization	Pre-reading activities generated by AI, tailored to each student's prior knowledge.	Personalized Goal-free problems to individual student levels to minimize cognitive load.	AI-generated completion problems tailored to each learner's progress.	AI-driven reflection prompts that encourage the re-evaluation and modification of existing schemas.	Personalized thematic organizers created by AI to enhance and refine existing schemas.
	Custom pre-organizers and advanced organizers created by AI to prime learners for new content.	Worked examples customized by AI to align with each student's skill level.	Interactive exercises and problems generated by AI to automate schema construction through practice.		
Engaging Educational Content Creation	AI-generated previews and thematic organizers to introduce new topics and activate prior knowledge.	Interactive video recordings and hierarchical concept maps created by AI to build foundational schemas.	AI-generated simulations and practical activities that promote schema automation through repetition and application.		AI-developed cases and interactive applications that encourage deeper exploration and connection of concepts across different domains.
Active-Learning Facilitation	AI-generated simulations and games designed to activate relevant prior knowledge	Practical activities and interactive learning experiences generated by AI to	AI-driven participatory exercises that promote repeated application and automation of		Parallel schema interaction activities created by AI to facilitate pattern recognition and

Use cases of Generative AI	Schema Activation	Schema Construction	Schema Automation	Schema Modification	Schema Elaboration
	before diving into new content.	construct new schemas.	newly constructed schemas		schema refinement through diverse, yet related, contexts.
Support for Teaching and Tutoring	AI-generated advanced organizers and previews that prepare students for new content.	AI-created hierarchical concept maps and Vee diagrams to facilitate the construction of new knowledge structures.	Adaptive learning exercises and AI feedback to support the ongoing practice and automation of schemas. AI-generated automated assessments that help students identify and solidify automated knowledge	Reflection prompts and unlearning/relearning activities generated by AI to assist in modifying existing schemas based on feedback.	AI-facilitated pattern recognition tasks and parallel schema interaction activities that deepen and elaborate existing knowledge structures.

4. DISCUSSION AND CONCLUSIONS

In the structure of the educational process, cognitive schemas are fundamental pillars for enhancing the formative experience and fostering the assimilation of new knowledge. These schemas act as catalysts by allowing students to intertwine newly acquired information with their prior knowledge, resulting in deeper understanding and improved learning retention. Generative artificial intelligence tools have emerged as cutting-edge resources for enriching educational dynamics. These tools enhance attention, promote adaptability in the teaching process, increase creativity, and encourage collaborative work, stimulating the brain and facilitating the assimilation and retention of knowledge. Moreover, these AI tools can facilitate the development of higher-order cognitive skills, such as critical analysis, reflective evaluation, and innovative capacity.

In the educational context, Generative AI tools are becoming increasingly important resources for improving schemas. These tools promote intelligent tutoring, facilitate the generation of personalized educational content, and boost the creation of simulations, immersive learning environments, and the automation of feedback and evaluation processes.

In a study conducted by Jaboob et al. (2024), students experienced notable satisfaction when using AI applications, which positively influenced their behavior and cognitive performance. This finding aligns with the conclusions of Howard-Jones (2010), who emphasizes the importance of emotions in knowledge acquisition, given their impact on students' motivation, attention, memory, and decision-making. Generative AI integration in education promotes active participation, personalized adaptation, effectiveness, and idea generation. Furthermore, the use of technology in classrooms, especially visual learning and recent innovations, is increasingly gaining ground among the student community.

These examples illustrate the great potential of generative AI in educational applications and demonstrate that it can be innovatively employed to enrich learning schemas. However, it is crucial to persist in the research and development of these technologies to optimize their capacity and overcome the ethical challenges and dilemmas inherent in their implementation in the educational sphere.

Some key advantages of integrating generative AI in educational environments include the following:

- Significant improvement in student motivation and engagement.
- Acceleration of the learning process and facilitation of information retention.
- Promotion of active participation and the acquisition of meaningful knowledge.
- Increase in accessibility to educational content adapted to various learning modalities.
- Provision of personalized and timely feedback to students regarding their performance.
- Fostering student autonomy and responsibility in the learning process.

Generative AI implementation in education brings many advantages, but it also presents challenges and issues that need attention for the best results. These challenges involve the possibility of creating content that's not up to standard or not relevant, worries about protecting intellectual property, safeguarding student data privacy and security, and the potential for reinforcing biases present in current educational materials, among other important considerations.

It is important to emphasize that generative AI tools do not take over the teacher's role in education. Teachers still play a vital role in choosing the right tools, evaluating the content they generate, and integrating it effectively into the teaching-learning process.

For generative AI to be effectively integrated into educational settings, teachers should gain more skills and knowledge about how to use this technology professionally. Moreover, it is crucial to create a strong framework that guarantees the responsible development of AI models focused on transparency and explainability and supported by ethical guidelines to ensure the responsible use of these tools in education.

This study explored how generative AI contributes to building cognitive schemas in education, uncovering evidence that these tools can tailor educational content to individual student needs, enhancing both the effectiveness and relevance of learning. Additionally, we observed that generative AI can produce diverse educational resources, enriching the learning journey and addressing various cognitive styles. By leveraging generative AI, educators can develop more impactful and student-centered teaching methods, fostering deeper and more personalized learning experiences.

Generative AI holds promise for education but presents challenges. These include the risk of low-quality or irrelevant content, privacy and intellectual property concerns, and the potential reinforcement of biases. The teacher's participation is essential in selecting and evaluating tools and integrating them effectively. Educators need training to use generative AI responsibly, aligned with a framework with clear ethical guidelines for proper implementation.

Future research should focus on assessing the effectiveness of generative AI tools compared with traditional teaching methods, which will provide valuable insights into how AI-generated

educational content impacts both academic achievements and student engagement. Considering the advancements in the development of novel generative AI tools, it's advisable to implement a continuous evaluation process for these tools. This would streamline the creation and dissemination of educational content tailored to individual student needs. Furthermore, exploring how generative AI can enhance equity and inclusion, especially for marginalized students or those with special needs, is imperative.

5. REFERENCES

- Akram, N. (2019). *Role and Importance of Schemas in Pedagogy and Learning: A Cognitive Approach* (SSRN Scholarly Paper 3829364). <https://papers.ssrn.com/abstract=3829364>
- Brewer, W. F. (2000). Bartlett, Functionalism, and Modern Schema Theories. *The Journal of Mind and Behavior*, 21(1/2), 37–44.
- Chen, C., Fu, J., & Lyu, L. (2023). *A Pathway Towards Responsible AI Generated Content* (arXiv:2303.01325). arXiv. <https://doi.org/10.48550/arXiv.2303.01325>
- Chheang, V., Sharmin, S., Marquez-Hernandez, R., Patel, M., Rajasekaran, D., Caulfield, G., Kiafar, B., Li, J., Kullu, P., & Barmaki, R. L. (2024). *Towards Anatomy Education with Generative AI-based Virtual Assistants in Immersive Virtual Reality Environments* (arXiv:2306.17278). arXiv. <https://doi.org/10.48550/arXiv.2306.17278>
- Dang, S., Sharma, P., & Shekhawat, L. S. (2019). Cognitive Schemas among Mental Health Professionals and Other Health Professionals. *Indian Journal of Psychological Medicine*, 41(3), 258–265. https://doi.org/10.4103/IJPSYM.IJPSYM_194_18
- Darejeh, A., Mashayekh, S., & Marcus, N. (2022). Cognitive-based methods to facilitate learning of software applications via E-learning systems. *Cogent Education*, 9(1), 2082085. <https://doi.org/10.1080/2331186X.2022.2082085>
- Derry, S. J. (1996). Cognitive schema theory in the constructivist debate. *Educational Psychologist*, 31(3–4), 163–174. <https://doi.org/10.1080/00461520.1996.9653264>
- Epstein, Z., Hertzmann, A., Herman, L., Mahari, R., Frank, M. R., Groh, M., Schroeder, H., Smith, A., Akten, M., Fjeld, J., Farid, H., Leach, N., Pentland, A., & Russakovsky, O. (2023). Art and the science of generative AI: A deeper dive. *Science*, 380(6650), 1110–1111. <https://doi.org/10.1126/science.adh4451>
- Fareed, M. W., Bou Nassif, A., & Nofal, E. (2024). Exploring the Potentials of Artificial Intelligence Image Generators for Educating the History of Architecture. *Heritage*, 7(3), Article 3. <https://doi.org/10.3390/heritage7030081>
- Feuerriegel, S., Hartmann, J., Janiesch, C., & Zschech, P. (2023). *Generative AI* (SSRN Scholarly Paper 4443189). <https://doi.org/10.2139/ssrn.4443189>
- Gozalo-Brizuela, R., & Garrido-Merchán, E. C. (2023). *A survey of Generative AI Applications* (arXiv:2306.02781). arXiv. <https://doi.org/10.48550/arXiv.2306.02781>
- Hejny, M. (2018). Exploring the Cognitive Dimension of Teaching Mathematics through Scheme-oriented Approach to Education. *ORBIS SCHOLAE*, 6(2), 41–55.

- Hsu, C. H. C., Tan, G., & Stantic, B. (2024). A fine-tuned tourism-specific generative AI concept. *Annals of Tourism Research*, 104, 103723. <https://doi.org/10.1016/j.annals.2023.103723>
- Jung, E., Lim, R., & Kim, D. (2022). A Schema-Based Instructional Design Model for Self-Paced Learning Environments. *Education Sciences*, 12(4), Article 4. <https://doi.org/10.3390/educsci12040271>
- Lv, Z. (2023). Generative artificial intelligence in the metaverse era. *Cognitive Robotics*, 3, 208–217. <https://doi.org/10.1016/j.cogr.2023.06.001>
- Maj-Waśniowska, K., Stanienda, J., & Wyrobek, J. (2022). Challenges for the education system in the era of the Fourth Industrial Revolution. En *Public Goods and the Fourth Industrial Revolution*. Routledge.
- Menekse, M. (2023). Envisioning the future of learning and teaching engineering in the artificial intelligence era: Opportunities and challenges. *Journal of Engineering Education*, 112(3), 578–582. <https://doi.org/10.1002/jee.20539>
- Mills, A., Bali, M., & Eaton, L. (2023). How do we respond to generative AI in education? Open educational practices give us a framework for an ongoing process. *Journal of Applied Learning and Teaching*, 6(1), Article 1. <https://doi.org/10.37074/jalt.2023.6.1.34>
- Moreira, M. A., Greca, I. M., & Rodríguez Palmero, M. L. (2002). *Modelos mentales y modelos conceptuales en la enseñanza & aprendizaje de las ciencias*. <https://lume.ufrgs.br/handle/10183/204845>
- OpenAI. (2024). *Introducing GPT-4o and more tools to ChatGPT free users*. <https://openai.com/index/gpt-4o-and-more-tools-to-chatgpt-free/>
- Parrales, E. B. A., Palma, J. K. T., Álava, R. A. Q., & Campuzano, M. F. P. (2020). The cognitive process and influence in learning. *International Journal of Linguistics, Literature and Culture*, 6(2), Article 2. <https://doi.org/10.21744/ijllc.v6n2.875>
- Pendy, B. (2021). Artificial Intelligence: The Future of Education. *Jurnal Indonesia Sosial Sains*, 2(11). <https://doi.org/10.59141/jiss.v2i11.801>
- Sanabria-Navarro, J.-R., Silveira-Pérez, Y., Pérez-Bravo, D.-D., & de-Jesús-Cortina-Núñez, M. (2023). Incidencias de la inteligencia artificial en la educación contemporánea. *Comunicar: Revista Científica de Comunicación y Educación*, 31(77), 97–107. <https://doi.org/10.3916/C77-2023-08>
- Soriano, Á. E. (2017). Los esquemas de aprendizaje: Kant y Piaget. Introducción filosófica-psicológica. *Revista Torreón Universitario*, 6(16), Article 16. <https://doi.org/10.5377/torreon.v6i16.6557>
- Suzuki, M., & Matsuo, Y. (2022). A survey of multimodal deep generative models. *Advanced Robotics*, 36(5–6), 261–278. <https://doi.org/10.1080/01691864.2022.2035253>
- Sweller, J. (2020). Cognitive load theory and educational technology. *Educational Technology Research and Development*, 68(1), 1–16. <https://doi.org/10.1007/s11423-019-09701-3>

- Torney-Purta, J. (1991). Schema Theory and Cognitive Psychology: Implications for Social Studies. *Theory & Research in Social Education*, 19(2), 189–210. <https://doi.org/10.1080/00933104.1991.10505636>
- Tyagi, M., Ranjan, S., Smiti, & Gupta, A. (2022). Transforming Education System through Artificial Intelligence and Machine Learning. *2022 3rd International Conference on Intelligent Engineering and Management (ICIEM)*, 44–49. <https://doi.org/10.1109/ICIEM54221.2022.9853195>
- Tzuriel, D. (2021). Cognitive Education Programs. En D. Tzuriel (Ed.), *Mediated Learning and Cognitive Modifiability* (pp. 413–459). Springer International Publishing. https://doi.org/10.1007/978-3-030-75692-5_15
- Ubah, A. E., Onakpojeruo, E. P., Ajamu, J., Mangai, T. R., Isa, A. M., Ayansina, N. B., & Al-Turjman, F. (2022). A Review of Artificial Intelligence in Education. *2022 International Conference on Artificial Intelligence of Things and Crowdsensing (AloTCs)*, 38–45. <https://doi.org/10.1109/AloTCs58181.2022.00104>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is All you Need. *Advances in Neural Information Processing Systems*, 30. https://proceedings.neurips.cc/paper_files/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html
- Voskoglou, M. (2020). *New Challenges for Education in the Forthcoming Era of the Fourth Industrial Revolution* (new-challenges-for-education-in-the-forthcoming-era-of-the-fourth-industrial-revolution) [Chapter]. <https://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/978-1-7998-4882-0.ch004>; IGI Global. <https://www.igi-global.com/gateway/chapter/www.igi-global.com/gateway/chapter/258035>
- Yu, H., & Guo, Y. (2023). Generative artificial intelligence empowers educational reform: Current status, issues, and prospects. *Frontiers in Education*, 8. <https://doi.org/10.3389/feduc.2023.1183162>

Para citar este artículo:

Salgado Granda, B., Inzhivotkina, Y., Ibáñez Apolo, M. F., & Ugarte Fajardo, J. G. Educational innovation: Exploring the potential of Generative Artificial Intelligence in cognitive schema building. *EduTec, Revista Electrónica de Tecnología Educativa*, (89), 44-63. <https://doi.org/10.21556/edutec.2024.89.3251>

6. APPENDIX

Table 2

List of relevant studies on the application of generative AI in education.






Title	Focus	Authors
Generative Artificial Intelligence as a Tool for Teaching Communication in Nutrition and Dietetics Education-A Novel Education Innovation	Describes the creation of an AI-driven platform to simulate virtual patients to improve communication skills in nutrition and dietetics students, offering cost-effective training and receiving positive initial feedback.	Barker, L. A.; Moore, J. D.; Cook, Helmy A.
Artificial intelligence and medical education: application in classroom instruction and student assessment using a pharmacology & therapeutics case study	Explores the application of three generative AI tools to create educational content with specific learning outcomes, test items, and standard-setting parameters for medical students.	Sridharan, K.; Sequeira, R. P.
Physical education teaching mode assisted by artificial intelligence assistant under the guidance of high-order complex network	Analyzes the integration of artificial intelligence (AI) teaching assistants in tennis instruction using compressed sensing=FCM (CS-FCM) to learn fuzzy cognitive maps. FCMs can be integrated with generative AI to enhance their capabilities and applications through enhanced modeling and simulation, which generate intuitive and interactive visual representations.	Song, X.
Leveraging the Potential of Large Language Models in Education Through Playful and Game-Based Learning	Examines LLMs' educational potential and proposes playful and game-based learning strategies to mitigate risks, ensuring expertise development while enhancing human-AI interaction in education.	Huber, S. E.; Kiili, K.; Nebel, S.; Ryan, R. M.; Sailer, M
Can artificial intelligence support creativity in early design processes?	Explores Generative AI to enhance design creativity and quality while reducing cognitive load, demonstrating its transformative potential in design education through improved design outcomes and efficient cognitive load management.	Chandrasekera, T.; Hosseini, Z.; Perera, U.
Digital transformation in engineering education: Exploring the potential of AI-assisted learning	Studies ChatGPT's potential in enhancing engineering education, emphasizing the need for student training and guidance, and proposes an AI-assisted learning flow for effective integration into educational practices.	Pham, T.; Nguyen, B.; Ha, S.; Ngoc, T.N.
Old dog, new tricks? Exploring the potential functionalities of ChatGPT in supporting educational methods in social psychiatry	Inspect ChatGPT's roles in social psychiatry education, highlighting its potential as a teaching tool, including information provider, debate facilitator, self-directed learning tool, and content creator for course materials	Smith, A.; Hachen, S.; Schleifer, R.; Bhugra, D.; Buadze, A.
Natural language processing for automatic evaluation of free-text answers — a feasibility study based on the European Diploma in Radiology examination	Assesses NLP's capability to automate free-text answer evaluation in medical exams to support the review process as a decision support system for reviewers and to support examination design.	Stoehr, F.; Kämpgen, B.; Müller, L.; Zufiria, L. O.

Title	Focus	Authors
Unlocking the opportunities through ChatGPT Tool towards ameliorating the education system	Examines ChatGPT’s role in education, emphasizing its applications, such as automated grading, personalized learning experiences, and content customization. The results highlight its capability to automate grading, translate text, and customize educational content based on student needs.	Javaid, M.; Haleem, A.; Singh, R. P.; Khan, S.; Khan, I.
Race with the machines: Assessing the capability of generative AI in solving authentic assessments	Presents a framework to evaluate generative AI tools in assessments, revealing strengths at lower Bloom's taxonomy levels and weaknesses at higher levels, thus urging a reevaluation of educational goals and assessments. The findings suggest a need to reimagine educational goals and assessments to emphasize higher cognitive skills and enhance educators’ capabilities to prepare students for a rapidly changing professional environment.	Thanh, B. N.; Vo, D. T.; Nhat, M. N.; Pham, T. T. T.; Trung, H. T.; Xuan, S. H.
Integrating generative AI in knowledge building	Explores high school students’ use of ChatGPT for knowledge building and the development of AI literacy through empirical accounts following a teacher’s pedagogical design. This study offers practical implications for integrating GenAI in K-12 education and urges educators to create spaces and scaffolds for students to engage with GenAI mindfully in the classroom.	Chen, B.; Zhu, X.; Díaz del Castillo H., F.
ChatGPT in physics education: A pilot study on easy-to-implement activities	Demonstrates the potential of implementing ChatGPT to enhance learning and support teachers' skills in secondary-school physics classrooms, highlighting its potential to enhance learning and critical thinking despite inherent limitations and biases.	Bitzenbauer, P.
How ChatGPT can inspire and improve serious board game design	Examines ChatGPT's role in educational board game design and underscores the need for human oversight to mitigate biases and inaccuracies in game development. The research highlights ChatGPT's utility in suggesting game themes and mechanisms, providing customization templates, and offering feedback on prototypes to enhance game components and mechanisms.	Junior, W. G.; Marasco, E.; Kim, B.; Behjat, L.; Eggermont, M.
Teaching and assessment of the future today: higher education and AI	Examines the capabilities of LLMs in performing tasks traditionally requiring human cognition, including data analysis, problem-solving, and content generation. It also discusses the rapid progression of AI models, noting their expanding functionalities, such as real-time information access and image analysis.	Lacey, M. M.; Smith, D. P.
Multimodality and English for Special Purposes: Signification and Transduction in Architecture and Civil Engineering Models	Investigates the necessity of multimodal literacy in architecture and civil engineering education, emphasizing the importance of transforming text-centric learning into multimodal communication skills required in professional practice. The study tasked students in English for Architecture and Civil Engineering courses to create digital, multimodal artifacts explaining concepts to a lay audience.	Hellwig, A. F. J.; Jones, P. T. ; Matruglio, E.; Georgiou, H.
Educational Design Principles of Using AI Chatbot That Supports Self-Regulated Learning in Education: Goal Setting, Feedback, and Personalization	Proposes integrating AI chatbots in education through prompting, reverse prompting, and learning analytics, emphasizing goal setting, self-assessment, and personalization to improve students’ self-regulated learning and ethical use of ethical concerns using SRL and JOL frameworks.	Chang, D. H.; Lin, M. P.; Hajian, S.; Wang, Q. Q.
Challenge, integration, and change: ChatGPT and future anatomical education	Discusses the application of ChatGPT in anatomy teaching in medical education. ChatGPT enhances student engagement and independent learning in anatomy, offering individualized and immediate medical knowledge through interactive simulations.	Leng, L.



La evaluación en educación matemática: aportes de chatbots y futuros profesores de matemática

Assessment in Mathematics education: contributions from chatbots and future mathematics teachers

-  Patricia Sureda; psureda@niem.exa.unicen.edu.ar; NIEM - CONICET-UNCPBA (Argentina)
-  Ana Corica; acorica@niem.exa.unicen.edu.ar; NIEM - CONICET-UNCPBA (Argentina)
-  Veronica Parra; yparra@niem.exa.unicen.edu.ar; NIEM - CONICET-UNCPBA (Argentina)
-  Daniela Godoy; daniela.godoy@isistan.unicen.edu.ar; ISISTAN - UNCPBA/CONICET (Argentina)
-  Silvia Schiaffino; silvia.schiaffino@isistan.unicen.edu.ar; ISISTAN - UNCPBA/CONICET (Argentina)

Resumen

La evaluación del aprendizaje de los estudiantes es un tema de investigación relevante de la didáctica de la matemática. Evaluar en matemática requiere mucho más que la resolución de un ejercicio. Se trata de evaluar todo el proceso. En este sentido, el diseño de evaluaciones no es trivial ni inmediato. Requiere formación, objetivos claros y propuestas relevantes. En este trabajo se analizan las evaluaciones propuestas por tres futuros profesores de matemática y por tres chatbots basados en modelos de Inteligencia Artificial (IA) generativa. Se comparan los tipos de evaluaciones propuestas sobre nociones de estadística (población y muestra) y se determina la funcionalidad de los chatbots como posibles asistentes para la generación de diferentes tipos de evaluaciones. Se concluye que los chatbots pueden resultar en asistentes valiosos a la hora de crear evaluaciones, ya que ofrecen diferentes tipos de evaluaciones, tanto tradicionales, como puede ser una prueba escrita, como no tradicionales, como un proyecto de investigación.

Palabras clave: Inteligencia artificial, chatbots, futuros profesores de matemática, población y muestra, evaluación.

Abstract

The assessment of student learning is a relevant research topic for the didactic of mathematics. Assessment in mathematics requires much more than solving an exercise. It is about evaluating the entire process. In this regard, the design of evaluations is neither trivial nor immediate. It requires training, clear objectives and relevant proposals. This work analyzes the evaluations proposed by three future mathematics teachers and by three chatbots based on generative Artificial Intelligence (AI) models. The types of evaluations proposed on notions of statistics (population and sample) are compared and the functionality of chatbots as possible assistants for the generation of different types of evaluations is examined. It is concluded that chatbots might become valuable assistants when creating evaluations, since they offer different types of evaluations, both traditional, such as a written test, and non-traditional ones, such as a research project.

Keywords: Artificial intelligence, chatbots, Mathematics teachers, sample and population, assessment.



1. INTRODUCCIÓN

La palabra evaluación tiene, según el diccionario de la Real Academia Española (RAE), dos acepciones: *acción y efecto de evaluar*, y *examen escolar*. Para la primera, la acción de evaluar, la RAE propone sinónimos como valoración, cálculo, estimación, apreciación, tasación y peritaje. Esta noción de evaluación está ligada a la idea de asignarle un *valor* a algo o a alguien (Chevallard, 2012). Sin embargo, para la segunda acepción, examen escolar, la RAE propone el ejemplo: *hoy tengo la evaluación de matemática*. Esta noción de evaluación aparece más ligada a la forma de acreditar un saber o una materia. En Didáctica de la Matemática, la noción de evaluación se aleja de estas dos definiciones. A pesar de la diversidad de marcos teóricos y la propia asunción de evaluación en cada uno de ellos, hay una generalidad aceptada por la comunidad, la de asumir a la evaluación como un proceso continuo, formativo (Scriven, 1967), relativo, que va más allá de un examen escolar. Incluso, se plantea la cuestión de si es necesaria una teoría diferenciada de la evaluación en matemática (Webb, 1992).

El acto de evaluar es una tarea que puede estar presente en cualquier parte de la sociedad, y no sólo en la escuela: se evalúa uno mismo, se evalúa en la familia, en la calle, en la empresa y en el bar. Se evalúa en todas partes, pero restringir la evaluación al ámbito escolar conduce a asociar la idea de evaluar a la de *corregir*, a la de acreditar, a la de adecuarla a la idea de que hay un trabajo que debe hacerse correctamente. De hecho, el verbo corregir proviene del latín: *corrigere*, que significa enderezar (Chevallard, 2004). La evaluación, por su parte, aun sin asociarla a la idea de corregir, no existe por sí misma sin un contexto. Por ejemplo, la densidad del agua a temperatura ambiente (20°C aproximadamente) es de en torno a 1.000 kg/m³. Esta medida toma sentido, por ejemplo, para calcular la flotación de objetos, respecto a cierto proyecto. Sin embargo, Chevallard (2012) indica que muchos profesores consideran que el examen del alumno al que le han puesto un 8 vale un 8, de forma absoluta, independientemente de cualquier proyecto. Ambos ejemplos representan la relatividad de la evaluación. Tanto la densidad del agua como el número 8 tienen valores relativos dentro de un proyecto previamente determinado. La cuestión clave a formular entonces respecto a la relatividad de la evaluación es si el objeto, medida, etc. tiene valor, *¿para qué? o para hacer qué?*

En la enseñanza, en cualquier nivel escolar, se suele confundir la evaluación con el proceso de *acreditación* de un curso. Por lo general, las actividades de evaluación (exámenes, tareas y trabajos extraescolares) son un momento aparte del curso, una interrupción de éste para examinar el conocimiento de los estudiantes. Como la evaluación se usa exclusivamente para asignar una calificación al trabajo de los estudiantes, resulta que, si no realizan sus tareas correctamente, tendrán una baja calificación que lo puede llevar a no aprobar el curso. De este modo, el interés del estudiante se centra en pasar los exámenes y hacer bien las tareas; siendo ése su único objetivo y no el aprendizaje de un conocimiento (Flores Samaniego y Gómez Reyes, 2009). En este contexto, el estudio de la evaluación (cómo se define y concibe en cada marco didáctico, los diferentes tipos de evaluación propuestos en cada uno de ellos, sus características, etc.) es parte de la formación de los futuros profesores de matemática (FPM), como el caso de los que participan de esta investigación.

En relación con lo indicado en párrafos anteriores, el acto de evaluar en el contexto educativo no es trivial, y requiere de tomar decisiones reflexivas en relación con el tipo de evaluación empleada y al proceso de estudio desarrollado. En este sentido, este trabajo tiene como foco la temática de la evaluación en estadística (población y muestra), específicamente la evaluación propuesta por tres FPM y por tres chatbots, basados en modelos de Inteligencia Artificial (IA) generativa. Nos proponemos caracterizar y comparar los medios, técnicas e instrumentos de evaluación propuestos por los FPM y los chatbots, así como también, determinar la funcionalidad de los chatbots como posibles asistentes del profesor para la generación de diferentes tipos de medios, técnicas e instrumentos de evaluación.

2. ANTECEDENTES

La evaluación en Matemática

En Educación Matemática, el momento de evaluación constituye una temática de investigación que considera diferentes aspectos: la parte institucional, la normativa de cuantificar numéricamente o categorizar el desempeño de los estudiantes, en relación a una disciplina, los instrumentos de evaluación, sus potencialidades, limitaciones, etc. A su vez, la problemática de la evaluación posiciona a la comunidad docente en el dilema sobre si el valor numérico o cualitativo asignado al desempeño de los estudiantes da cuenta del conocimiento alcanzado por él, en un espacio académico de su formación profesional y si garantiza su idoneidad, competencia profesional y laboral (Hamodi et al., 2015).

En la comunidad científica y en la comunidad educativa no existe un consenso sobre un modelo de evaluación. En este sentido, Bermúdez y Osorio (2012) indican que la evaluación constituye un elemento del currículo, que no está aislada y que debe integrarse al proceso de enseñanza-aprendizaje de la matemática. De esta forma, asumimos que la evaluación y la calificación no son sinónimos. Diversos autores destacan, por ejemplo, cómo el profesorado tiende a confundir ambos conceptos (Álvarez, 2005; Fernández, 2006; Santos Guerra, 2003). Siguiendo a Sanmartí (2007), consideraremos a la evaluación como un proceso que permite recoger información, a través de algún instrumento que puede ser de diversa naturaleza (escrito, oral, etc.). Una vez recolectada esta información, su análisis permitirá emitir un juicio sobre la misma con la consecuente toma de decisiones, para determinar cuantitativamente o cualitativamente la calificación correspondiente.

Adoptamos las ideas de Hamodi et al. (2015), quienes sostienen la necesidad de un sistema de evaluación que clasifique a los medios, las técnicas y los instrumentos, y que contemple a los estudiantes y su participación en el proceso evaluativo. En la sección correspondiente al método se describen detalladamente estos componentes.

La IA y la evaluación en el sistema educativo

La IA está redefiniendo la forma en que se evalúa a los alumnos, en los diferentes niveles de educación. Según algunos autores, esta transformación de los procesos de evaluación

promete mejorar la eficiencia, precisión, y personalización en la evaluación, adaptándose a las necesidades individuales de los estudiantes. La IA ofrece oportunidades para un aprendizaje más centrado en el estudiante y una educación más inclusiva (Méndez-Mantuano, 2024). Sin embargo, también presenta desafíos éticos, de privacidad, y de equidad, especialmente en el manejo de datos estudiantiles. En Martínez-Comesaña et al. (2023) se presenta una revisión bibliográfica que muestra las posibilidades y los usos que la IA puede aportar a la educación, concretamente en la evaluación del rendimiento del alumnado de primaria y secundaria. Según los autores, los principales aportes de la IA en la evaluación del alumnado de estos niveles educativos inferiores se centran en la predicción de su rendimiento, evaluaciones más objetivas y automatizadas mediante técnicas como redes neuronales o procesamiento del lenguaje natural, el uso de robots educativos para analizar su proceso de aprendizaje y la detección de factores específicos que hacen más atractivas las clases.

Algunas instituciones gubernamentales en diferentes países han analizado el impacto de la IA en la educación, tanto en la enseñanza, como en el aprendizaje y en la evaluación. En un documento generado por el Departamento de Educación de EEUU (U.S. Department of Education, 2023) se proporcionan algunas guías sobre cómo la IA puede usarse para la evaluación de los aprendizajes, se esbozan algunas políticas y se dan recomendaciones respecto a la gestión de la IA en estos procesos. Como ejemplo de herramientas que usen IA para evaluación, se mencionan herramientas de calificación automática de ensayos (Automated Essay Scoring, AES). De manera similar algunas universidades se han pronunciado al respecto, tales como la Universidad de Monash en Melbourne, Australia (2024). En esta institución, se ha elaborado un documento respecto a los principios que deben seguir los procesos de evaluación en la universidad, particularmente si se usa IA. Se sugiere a los docentes, por ejemplo, probar si sus evaluaciones están diseñadas de manera que la IA pueda completarlas fácilmente, y se los impulsa a generar evaluaciones que abarquen más los procesos (por ej. el razonamiento subyacente) que en los resultados. En Owan et al. (2023) se analiza el potencial de diferentes herramientas de IA en la evaluación en educación. Por ejemplo, la IA puede ayudar a automatizar el proceso de calificación, ahorrando tiempo a los profesores y proporcionar a los estudiantes retroalimentación inmediata sobre sus tareas. También puede proporcionar información sobre gramática, ortografía y sintaxis analizando ensayos, informes y otras tareas escritas. Al utilizar sistemas de calificación automatizados, los profesores pueden centrarse más en tareas esenciales como la planificación de las clases, y apoyar a los estudiantes, lo que resulta en importantes ahorros de tiempo según Adiguzel et al. (2023).

Considerando particularmente la IA Generativa (IAGen), algunos trabajos han estudiado su impacto en los procesos de evaluación. Estos trabajos se pueden agrupar en dos grandes categorías. Por un lado, aquellos que analizan las implicancias del uso de la IAGen por parte de los alumnos y cómo deben responder o planificar los docentes las evaluaciones teniendo en cuenta este factor, y por otro lado, aquellos que ven su potencial como herramienta para confeccionar las evaluaciones y realizar las correcciones de los trabajos entregados por los alumnos. Entre los trabajos de la primera categoría, podemos encontrar a Sánchez Mendiola (2023), donde los autores analizan aquellos tipos de evaluación más afectados por la irrupción de la IAGen y qué acciones se podrían adoptar. Algunas sugerencias son ampliar el abanico de instrumentos de evaluación usando exámenes orales, proyectos, observación de discusiones, y elaboración de diagramas, en lugar de las tradicionales pruebas escritas. Los

documentos generados por el Departamento de Educación de EEUU y la Universidad de Monash, mencionados anteriormente, también dedican secciones especiales a la IAGen, su uso en las aulas, y sobre cómo contemplar estos factores en la generación de evaluaciones que puedan medir de manera efectiva el aprendizaje.

Respecto a los trabajos en la segunda categoría, en Tobler (2024) los autores proponen *GenAI Smart Grading*, una aplicación basada en ChatGPT que permite corregir los trabajos de los estudiantes de una forma automatizada. En el trabajo se describen potenciales contextos de aplicación de la propuesta, tales como ayudar a corroborar las correcciones realizadas por el docente que pueden tener algún sesgo, evaluar preguntas abiertas dándole a la herramienta conocimiento para que verifique la presencia o no de ciertos elementos en la respuesta, entre otros. La aplicación se validó con estudiantes de grado y posgrado de Ciencias Naturales y de Tecnología en una universidad suiza. En Owan et al. (2023) los autores analizan potenciales aplicaciones de la IA generativa en evaluación. Por ejemplo, sugieren que un LLM (del inglés, *Large Language Model*) puede analizar una gran cantidad de datos de texto relacionados con un curso o tema específico, identificar los temas y conceptos principales en ellos, y sugerir ítems de examen apropiados que permitan medir con precisión la capacidad de los estudiantes para comprender esos conceptos. En Nasution (2023) se propone el uso de ChatGPT para crear evaluaciones de tipo *multiple-choice* en el área de Biología. Se analiza la calidad de las preguntas generadas por la IA según su legibilidad, nivel de dificultad, precisión y relevancia respecto al tema estudiado.

En Wang y Chen (2024) se explora el uso de ChatGPT-3.5 para escribir comentarios a las respuestas escritas por los estudiantes a preguntas conceptuales en el área de Física utilizando las técnicas de *prompt engineering* (el proceso de desarrollar y perfeccionar un *prompt* o instrucción) y *few-shot learning* (brindar unos pocos ejemplos para que el modelo pueda generalizarse sobre nuevos escenarios). Se evaluó la habilidad de la IA generativa para dar *feedback* a los alumnos, comparándola con el que brindaría el docente.

En cuanto a la aplicación de IA para la evaluación de alumnos en el área de Matemática, en Luzano (2024) se analizan las oportunidades y amenazas que puede presentar esta tecnología. En Adair et al. (2023) se presenta el diseño y la prueba inicial en laboratorios virtuales con soporte de IA que ayudan a los estudiantes de matemática a practicar en el desarrollo de modelos matemáticos de fenómenos de la ciencia. Los laboratorios evalúan automáticamente las competencias de modelado matemático de los estudiantes en función de las acciones que toman para construir sus modelos matemáticos dentro de los laboratorios.

A pesar de estos y otros trabajos en esta área interdisciplinar, hasta donde hemos analizado y estudiado, no hay estudios donde se utilice la IA generativa concretamente para proponer diferentes alternativas de evaluación ante un tema dado en el área de Matemática. En este contexto los objetivos de este trabajo son:

- Caracterizar los medios, técnicas e instrumentos de evaluación propuestos por tres FPM y por tres chatbots, sobre nociones de estadística (población y muestra).
- Comparar los medios, técnicas e instrumentos de evaluación propuestos por los FPM y los chatbots.

- Determinar la funcionalidad de los chatbots como posibles asistentes del profesor para la generación de diferentes tipos de medios, técnicas e instrumentos de evaluación.

3. MÉTODO

Contexto de recolección de datos: profesores y chatbots

Esta investigación analiza las evaluaciones sobre las nociones de población y muestra propuestas por tres futuros profesores de Matemática de secundaria (FPM) (P1, P2 y P3) y por tres chatbots: ChatGPT¹, Gemini² y Copilot³ (creativo, equilibrado y preciso). Los FPM diseñaron estas evaluaciones durante un curso de Didáctica de la Matemática del tercer año de su formación. El diseño de la evaluación es una de las actividades que los FPM deben realizar y entregar dentro del curso. En este trabajo, consideramos solamente las propuestas de los FPM que optaron por las nociones de población y muestra, ya que, a lo largo de todo el curso, se los entrena en el diseño de propuestas y evaluaciones de este tipo. El análisis se realiza sobre el documento entregado por los FPM al final del curso.

En el caso de los chatbots, la tarea de generación de texto recae en los modelos de lenguaje o LLMs en los que se fundamentan. ChatGPT utiliza los modelos de la serie GPT (Brown et al., 2020), como GPT 3.5 y GPT 4, empleado por Copilot. Gemini, en un modelo de lenguaje multimodal sucesor de PaLM-2, usado por Bard, y que tiene su origen en PaLM (Chowdhery et al., 2022). La generación de contenido en el caso de los chatbots responde a *prompts* o instrucciones. En este caso, se utilizaron los siguientes *prompts* dentro un mismo hilo de conversación, lo cual asegura que la segunda solicitud se dé dentro de un mismo contexto que la primera.

Prompt 1: Eres un profesor de matemática de escuela secundaria y tienes que evaluar el tema de Estadística: población y muestra, a estudiantes que van a la escuela secundaria de Argentina. ¿Cuáles serían las posibles maneras de evaluar el tema?

El objetivo de este primer *prompt* fue indagar en los posibles tipos de evaluaciones que podrían sugerir los *chatbots*, ya que en los cursos de Didáctica (y afines) y previo al diseño de evaluaciones, los FPM estudian formas de evaluar en matemática que provienen de concepciones más amplias de evaluación, que van más allá de un examen individual escrito.

Prompt 2: ¿Puedes generar un ejemplo concreto de una evaluación de alguna de esas formas?

Este segundo *prompt* surge a partir de las respuestas al primero ya que, como se verá en la sección de análisis y discusión, los *chatbots* proponen diferentes maneras de evaluar ante la misma pregunta. Los FPM recibieron una tarea a resolver con la misma intención que los

1 <https://chat.openai.com/>

2 <https://gemini.google.com/>

3 <https://copilot.microsoft.com/>

prompt, pero con una redacción en lenguaje natural, menos automatizada, y de acuerdo a las formas de comunicación habituales en las prácticas docentes.

Meta-categorías de análisis del sistema evaluativo

El análisis de los datos se realiza en función de los componentes del sistema evaluativo definidos por Hamodi et al. (2015): los medios, las técnicas y los instrumentos.

3.1.1. Los medios de evaluación

Los medios de evaluación se refieren a las producciones de los estudiantes que el profesor puede recoger, ver y/o escuchar. Pueden adoptar tres formas diferentes:

- a) escritos, por ejemplo, examen, carpeta dossier, diario de clase, proyecto, trabajo escrito, etc.
- b) orales, por ejemplo, comunicación, debate, presentación oral, etc.
- c) prácticos, por ejemplo, práctica supervisada, representación, juego de roles, etc.

3.1.2. Las técnicas de evaluación

Las *técnicas* de evaluación son las estrategias que el profesor utiliza para recoger información sobre las producciones y evidencias creadas por los estudiantes de los medios. Las técnicas a utilizar son diferentes en función de si los estudiantes participan o no en el proceso de evaluación.

- a) Si las técnicas son aplicadas sólo por el profesor, se utilizan unas u otras dependiendo del medio (escrito, oral o práctico); si el medio a evaluar es *escrito*, se utilizará la técnica del *análisis documental* y de *producciones* (o revisión de trabajos); si el medio a evaluar es *oral* o *práctico*, se utilizará la observación o el análisis de una grabación (audio o video).
- b) Si el estudiante participa en el proceso evaluativo, las técnicas de evaluación pueden ser las siguientes:
 - i) autoevaluación: refiere a la evaluación que hace el estudiante de su propia evidencia o producción, atendiendo a criterios consensuados con anterioridad.
 - ii) coevaluación: el estudiante evalúa de manera recíproca a sus compañeros del grupo-clase, aplicando criterios de evaluación que fueron consensuados previamente.
 - iii) evaluación colaborativa o compartida: se refiere al proceso dialógico que mantiene el profesor con el estudiante sobre la evaluación de los procesos de enseñanza-aprendizaje que se han dado. Estos diálogos pueden ser individuales o grupales.

3.1.3. Los instrumentos de evaluación

Los *instrumentos* de evaluación corresponden a las herramientas que los estudiantes y el profesor emplean para organizar la información recogida a través de una determinada técnica de evaluación. Esa información debe registrarse de manera sistemática y precisa para que la evaluación sea un proceso riguroso. Algunos ejemplos de estos instrumentos pueden ser: diario del profesor, rúbrica, ficha de observación, ficha de seguimiento individual o grupal, fichas de autoevaluación, fichas de evaluación entre iguales, informe de expertos, informe de autoevaluación, etc.

Clasificación de los datos

En la Tabla 1, presentamos la clasificación de las maneras de evaluar propuestas por los chatbots en respuesta al primer *prompt* y las producidas por los FPM. Esta clasificación es propia en función de las definiciones de los componentes del sistema evaluativo propuestos por Hamodi et. al (2015): medios (C_j^m); técnicas (C_k^t); e instrumentos (C_l^i).

Tabla 1

Análisis de los componentes del sistema evaluativo

FPM/Chatbots	C_j^m : medios evaluativos							C_k^t : técnicas evaluativas	C_l^i : instrumentos evaluativos	
	Escritos				Orales					
	Prueba escrita	Proyecto de investigación	Tarea práctica	Evaluación formativa	Presentación oral en clase	Examen oral	Debate	Uso de tecnologías / Simulaciones	Autoevaluación/evaluación entre pares	Infografía

4. RESULTADOS

En la sección 4.1 se presentan las categorías inductivas correspondientes a cada metacategoría que compone el sistema evaluativo. En la sección 4.2, se describen y analizan los dos únicos ejemplos de evaluación propuestos por los FPM y los chatbots: prueba escrita (C_1^m) y proyecto de investigación (C_2^m).

Medios, técnicas e instrumentos identificados

Se identificaron 8 medios de evaluación distintos (4 escritos, 3 orales y 1 práctico), 1 única técnica propuesta y 2 instrumentos de evaluación. A continuación, se detallan cada uno de ellos.

4.1.1. C_j^m : Medios evaluativos

- a) Los medios escritos que se obtuvieron de forma inductiva son los siguientes:

C_1^m Prueba escrita: se refiere al examen escrito tradicional, compuesto por ejercicios específicos a resolver de manera individual. Por ejemplo, las usuales “pruebas de matemática” del nivel secundario empleadas al finalizar un tema de estudio.

C_2^m Proyecto de investigación: consiste en llevar a cabo una investigación, una exploración o indagación, que involucra los saberes a evaluar, con la presentación de esa investigación en algún formato de entrega, por ejemplo, la redacción de un informe.

C_3^m Tarea práctica: resolución de tareas matemáticas relativas al tema, con datos que brinda el profesor. Se diferencia del proyecto de investigación en “el grado de libertad” por parte de los estudiantes. Por ejemplo, en la selección de la base de datos a analizar. Mientras que en C_2 los estudiantes son los que deciden sobre qué investigar, en C_3 , es el profesor quien determina el contexto y conjunto de datos.

C_4^m Evaluación formativa: se refiere a un medio de evaluación que puede contemplar otros medios. Se realiza durante todo el proceso de estudio, que permite monitorear el progreso de los estudiantes en particular y de la clase en general. En esta categoría no hay un momento específico, determinado en tiempo y espacio, donde se desarrolle la evaluación.

- b) Los medios orales que se obtuvieron a partir de las respuestas de los chatbots y FPM son:

C_5^m Presentación oral en clase: se refiere a la presentación del tema a evaluar, mediante algún soporte como puede ser una presentación por diapositivas.

C_6^m Examen oral: es el análogo a C_1 , en la modalidad oral. Se trata de un tipo de evaluación donde el profesor realiza preguntas a un estudiante por vez, quien debe responder de manera inmediata, desarrollándose completamente de forma verbal.

C_7^m Debate: se trata de generar un espacio dentro de la clase en el que se generen discusiones con todos los integrantes del grupo con relación a la temática a evaluar.

- c) Los prácticos que se obtuvieron a partir de las respuestas de los chatbots y FPM son como por ejemplo una práctica supervisada es:

C_8^m Software de estadística: se trata de un tipo de evaluación donde el conocimiento a evaluar se realiza utilizando algún software para analizar un conjunto de datos dados.

4.1.2. C_k^t : Técnicas evaluativas

La única técnica evaluativa propuesta fue la siguiente:

C_1^t Autoevaluación/evaluación entre pares: se trata de un tipo de evaluación donde los estudiantes evalúan sus propias producciones y/o las de otros estudiantes, con el fin de desarrollar habilidades críticas y de auto-reflexión.

Esto puede deberse a que el *prompt* 1 solicitó maneras de evaluar en términos generales, no necesariamente ajustada a los medios de evaluación definidos por Hamodi et al. (2015). Esto se debe a que este marco teórico no forma parte de la formación de los profesores de matemática, ni fue especificado a los chatbots para contextualizar su respuesta, por lo que en ambos casos respondieron con la opción que les es más conocida. La categorización en medios, técnicas e instrumentos fue una técnica metodológica de análisis de los investigadores.

4.1.3. C_j^i : Instrumentos evaluativos

Mientras que los instrumentos evaluativos fueron los siguientes:

C_1^i Infografía: se refiere a una modalidad de evaluación que demanda recuperar lo estudiado, y transformarlo para ser comunicado a través de alguna forma de difusión escrita. Por ejemplo, a través de algún tipo de cartelera. Este tipo de evaluación requiere, además del poder de síntesis del emisor, de las habilidades interpretativas del receptor para entender y otorgar sentido a la información comunicada. En esta categoría se puede acceder a toda la información a la vez, aun cuando pueda tener un orden cronológico.

C_2^i Recurso audio/video educativo: se refiere a una modalidad de evaluación que demanda recuperar lo estudiado y transformarlo para ser comunicado a través de algún medio audiovisual de difusión. Este tipo de evaluación se diferencia de C_8 porque, al ser dinámico, requiere, además del poder de síntesis del emisor, del orden cronológico y la organización del contenido en una línea temporal. Otra característica que la distingue es la disponibilidad de la información. En este caso, el receptor no dispone de todo el contenido a la vez.

La Tabla 2 muestra las maneras de evaluar propuestas por los chatbots en respuesta al primer *prompt* y las producidas por los FPM para cada meta-categoría.

Tabla 2

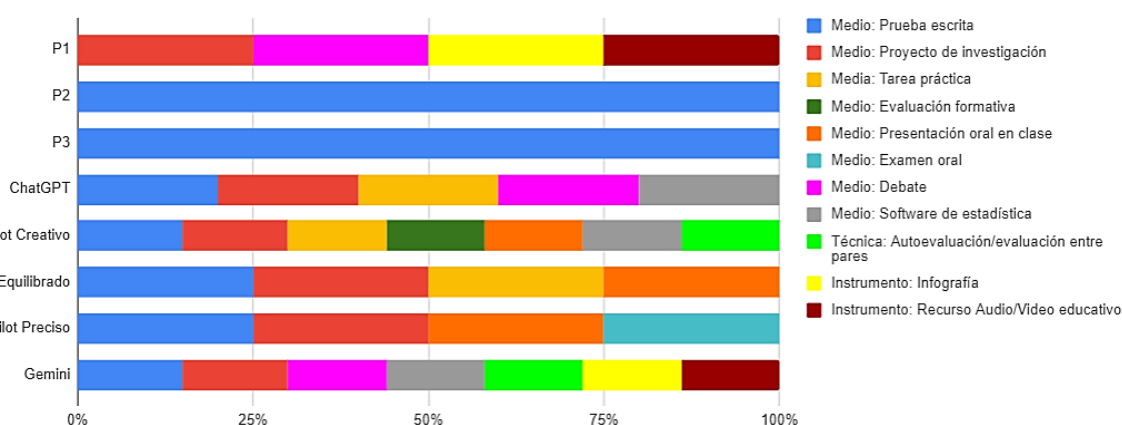
Análisis de los componentes del sistema evaluativo

FPM/ Chatbots	C_j^m : medios evaluativos								C_k^t : técnicas evaluativas	C_l^i : instrumentos evaluativos	
	Escritos				Orales			Práctico			
	Prueba escrita	Proyecto de investigación	Tarea práctica	Evaluación formativa	Presentación oral en clase	Examen oral	Debate	Uso de tecnologías/Simulaciones	Autoevaluación/evaluación entre pares	Infografía	Recurso Audio/Video educativo
P1	X	✓	X	X	X	X	✓	X	X	✓	✓
P2	✓	X	X	X	X	X	X	X	X	X	X
P3	✓	X	X	X	X	X	X	X	X	X	X
ChatGPT	✓	✓	✓	X	X	X	✓	✓	X	X	X
Copilot Creativo	✓	✓	✓	✓	✓	X	X	✓	✓	X	X
Copilot equilibrado	✓	✓	✓	X	✓	X	X	X	X	X	X
Copilot Preciso	✓	✓	X	X	✓	✓	X	X	X	X	X
Gemini	✓	✓	X	X	X	X	✓	✓	✓	✓	✓

La Figura 1 permite apreciar visualmente con mayor claridad la diversidad de medios, técnicas e instrumentos empleados por cada profesor y chatbot. Considerando las tres meta-categorías, el único que propone los tres componentes del sistema evaluativo es Gemini: 4 medios (C_1^m : prueba escrita, C_2^m : proyecto de investigación, C_7^m : debate y C_8^m : software de estadística), 1 técnica (C_1^t : autoevaluación) y 2 instrumentos (C_1^i : infografía y C_2^i : recurso audio/video educativo). Copilot creativo, si bien no cubre los tres aspectos, ofrece la mayor variedad de medios (todos excepto examen oral y debate) y una técnica (C_1^t autoevaluación). Los otros chatbots se limitan al uso de medios. Es importante resaltar que, en el extremo opuesto, P2 y P3 sólo proponen un medio: prueba escrita.

Figura 1

Formas para evaluar población y muestra según los FPM y los chatbots



El tipo de medio más frecuente fue *prueba escrita* (C_1^m). Es el único medio de evaluación propuesto por dos de los FPM (P2 y P3). Todos los chatbots lo proponen como una opción más, entre otras. De los FPM, P1 es el único que considera dos de los componentes de evaluación diferentes: medios, *proyecto de investigación* (C_2^m) y *debate* (C_7^m), e instrumentos, *infografía* (C_1^i) y *recursos audio/video educativo* (C_2^i). Todos los chatbots propusieron *prueba escrita* (C_1^m) y *proyecto de investigación* (C_2^m). Copilot equilibrado y Copilot preciso, además de estas dos opciones (C_1^m y C_2^m), incorporan el medio *presentación oral en clase* (C_5^m). El cuarto medio de evaluación para Copilot equilibrado es *tarea práctica* (C_3^m) y para Copilot preciso es *examen oral* (C_6^m). Estos chatbots son los que menos opciones propusieron (cuatro cada uno). ChatGPT, además de C_1^m y C_2^m , propone otros tres medios posibles: *tarea práctica* (C_3^m), *debate* (C_7^m) y *software de estadística* (C_8^m). Copilot creativo, incluye, además: *tarea práctica* (C_3^m), *presentación oral en clase* (C_5^m) y *evaluación formativa* (C_4^m). Gemini, por su parte, incorpora: *debate* (C_7^m), *infografía* (C_1^i) y *recurso audio/video educativo* (C_2^i).

En el apartado siguiente, se analizan las respuestas de los *chatbots* al *prompt* 2: ¿Puedes generar un ejemplo concreto de una evaluación de alguna de esas formas? y las propuestas presentadas por los FPM. De este análisis, se determinó que los únicos ejemplos de medios evaluativos propuestos fueron: *prueba escrita* (C_1^m) y *proyecto de investigación* (C_2^m).

Prueba escrita (C_1^m) y proyecto de investigación (C_2^m)

En esta sección se describen detalladamente cada uno de estos medios. Para la prueba escrita (C_1^m) se consideraron los siguientes descriptores:

- **Cantidad de ejercicios:** se detalla el número total de tareas que conforman la prueba escrita.
- **Objetivo:** el o los propósitos de evaluación de cada una de las tareas propuestas. Por ejemplo, indicar que con la tarea 1 se espera evaluar la distinción entre población y muestra.
- **Instrucciones:** recomendaciones respecto a la resolución y entrega de la prueba escrita. Por ejemplo, indicar al inicio de la misma que todos los cálculos realizados deben hacerse explícitos.
- **Ejercicios/Partes:** las tareas que conforman la prueba escrita. Algunos de los chatbots las segmentan en secciones o partes, indicando que contiene esa parte, por ejemplo, ejercicios que refieren a resolver problemas.
- **Puntuación:** asignación de puntos a la prueba escrita y al detalle de su distribución por tarea y/o por procedimiento de resolución.

Para los proyectos de investigación (C_2^m), los respectivos descriptores son:

- **Objetivo:** el o los propósitos del proyecto. Por ejemplo, aplicar la estadística para el análisis de datos reales.
- **Tema de interés:** el contexto en el que se involucra el proyecto. Por ejemplo, realizar un análisis de la educación de niños en comunidades rurales de Argentina.
- **Definición y/o selección de la población y muestra:** una vez seleccionado el tema de interés, establecer si corresponde el análisis sobre una población o una muestra y justificar la selección de las mismas. Por ejemplo, en la temática de la educación de niños en comunidades rurales de Argentina, identificar y justificar el método de muestreo que se utilizará.
- **Diseño del dispositivo:** consideraciones acerca de la propuesta de las preguntas que conforman el dispositivo. Por ejemplo, las preguntas deben ser claras, concisas y relevantes para el tema de investigación.
- **Recolección y análisis de datos:** caracterizar la forma de recolección y análisis de los datos generados con el instrumento. Por ejemplo, si van a realizar la encuesta de forma presencial, online o por teléfono y si los datos recolectados se analizarán con herramientas estadísticas o realizando cálculos manualmente.
- **Transformación de Datos:** caracterizar la forma de presentar los resultados del análisis de los datos. Por ejemplo, confección de tablas, gráficos, cálculo de medidas de tendencia central y dispersión, etc.
- **Presentación de resultados:** forma de entrega de las conclusiones sobre el análisis de los datos recolectados. Por ejemplo, presentando un informe con una estructura específica.
- **Criterio de evaluación:** aspectos a ponderar de la actividad desarrollada por los estudiantes al realizar el proyecto. Por ejemplo, claridad y precisión en la definición del tema de investigación; adecuación de la población objetivo y la muestra seleccionada; organización en la presentación de resultados, etc.

- **Puntuación:** asignación de puntos a las diferentes etapas que contempla el desarrollo del proyecto de investigación.

En la sección siguiente se presentan los resultados y discusión.

4.1.4. Ejemplo 1: prueba escrita (C_1^m)

C_1^m fue propuesta por dos FPM (P2 y P3) y por Copilot (creativo, equilibrado y preciso), siendo la de los chatbots las que contienen la mayor cantidad de ejercicios (5/6). Los dos FPM y Copilot creativo explicitan los objetivos mientras que Copilot Equilibrado y Preciso no lo hacen. Copilot creativo y copilot equilibrado son los únicos que proponen instrucciones relativas a la explicitación de los cálculos y las justificaciones al inicio de la propuesta.

Hay diferencias importantes en la naturaleza de los ejercicios propuestos por los FPM y los chatbots. Los FPM proponen ejercicios fuertemente centrados en *calcular*: 2 de los 3 ejercicios propuestos por P2, y los 2 propuestos por P3. De las tres modalidades de Copilot, el equilibrado es el que propone más ejercicios de *cálculo* (dos de tres), mientras que Copilot creativo solo uno y Copilot preciso ninguno. Se destaca que las tareas de *cálculo* que los chatbots proponen son más complejas que las de los dos FPM. Mientras los FPM solicitan calcular *mediana, moda, cuartiles y variación* de una muestra dada; los chatbots piden calcular la *media, la mediana, desviación estándar, el tamaño* de una muestra para una población dada, e *inferir* la cantidad de individuos de una población a partir de la muestra. Por otra parte, los ejercicios propuestos por los chatbots además del cálculo, requieren de habilidades como *definir, identificar* un ejemplo de población, *describir* una situación que requeriría estudiar una muestra, etc. Las tres versiones de Copilot solicitan que el estudiante *defina* una *población* y una *muestra*, y/o la diferencia entre ellas. Copilot equilibrado, además, pide una *explicación* sobre la importancia de seleccionar una *muestra representativa*. Copilot creativo y Copilot preciso, en otro ejercicio, proponen *preguntas de opción múltiple*, dando ejemplos de *población* y de *muestra*. El estudiante debe determinar cuál es una u otra (ver Fig. 2). Dado que este chatbot solicitaba justificar mediante cálculos y razonamientos, la tarea no es de resolución inmediata, sino que por el contrario requiere conocimiento, y capacidad de explicitación que le permita al estudiante justificar su elección.

Figura 2

Primer ejercicio propuesto por Copilot Creativo

Parte 1: Preguntas de Opción Múltiple
¿Qué es una población en estadística?
Un grupo de personas en un país.
El conjunto de todos los elementos que estamos estudiando.
Una parte del conjunto que estamos estudiando.
Un grupo de animales en un ecosistema.

2. ¿Qué es una muestra?
Un grupo seleccionado al azar de una población.
El conjunto completo de datos que estamos estudiando.
Una encuesta nacional.
Un gráfico estadístico.

Copilot creativo y Copilot preciso proponen además *describir una situación* en la que sería más apropiado estudiar una muestra en lugar de una población, y *describir cómo seleccionar una muestra representativa* para un estudio particular. Un aspecto positivo a favor de los FPM, y desfavorable para los chatbots, es el uso del *marco de representación gráfico*. Mientras los FPM piden como tarea realizar un gráfico para representar los datos, o la interpretación de los datos a partir de un gráfico; los chatbots no lo mencionan entre sus opciones. Esto se puede deber a que los chatbots son inteligencias artificiales conversacionales, y en consecuencia presentan problemas con la representación gráfica (Parra et al., 2024).

P3 y Copilot creativo no especifica la puntuación asignada a cada ejercicio. P2, Copilot equilibrado y Copilot preciso realizan asignación de puntos con diferencias notorias en el detalle a favor de P2 y Copilot preciso.

La Tabla 4 sintetiza estos resultados.

Tabla 4

Descriptores de C_1^m

	P2	P3	C. Creativo	C. Equilibrado	C. Preciso
Cant. de ejercicios	3	2	6	5	6
Objetivo	Explicita	Explicita	Explicita	No explicita	No explicita
Instrucciones	No explicita	No explicita	Explicitar cálculos y razonamientos.	Explicitar cálculos y justificaciones	No explicita
Ejerc./Parte 1	Tabla con datos reales del Censo Nacional Argentino. Cálculo de frecuencias absolutas, representación gráfica de los datos de la tabla y conclusiones.	Gráfico de barras con datos ficticios. Cálculo de promedio, moda, mediana, cuartiles, desvío estándar e interpretación de resultados.	Preguntas de opción múltiple sobre población y muestra (2 ejercicios).	Definición de población y muestra. Importancia de muestra representativa (2 ejercicios).	Definición de población y muestra y su diferencia (3 ejercicios).
Ejerc./Parte 2	Conjunto de 20 datos (discretos) ficticios. Cálculo de moda, media y mediana. Establecer cuál representa mejor los datos.	Tabla con datos ficticios. Cálculo de promedio, moda, mediana, cuartiles, desvío estándar. Interpretación de resultados.	Diferencia entre población y muestra y ejemplificar (2 ejercicios).	Conjunto de 10 datos ficticios. Cálculo de media y mediana (ejercicio 3). Generación de una muestra de 20 registros. Cálculo de media y desviación estándar (ejercicio 4).	Preguntas de opción múltiple sobre población y muestra (2 ejercicios).
Ejerc./Parte 3	Gráfico de torta con datos ficticios. Confección de tabla a partir del gráfico y determinación de	No corresponde.	Descripción de selección de muestra representativa (ejercicio 5). Cálculo	Identificación de población y la muestra en un conjunto de datos ficticios.	Justificación de selección de muestra. Descripción de selección de

	P2	P3	C. Creativo	C. Equilibrado	C. Preciso
	población y/o muestra. Justificar.		del tamaño de muestra necesario para un nivel de confianza del 95% y margen de error del 5%, para una población de tamaño n=500 (ejercicio 6)	Determinación de frecuencia relativa y concluir (ejercicio 5).	muestra representativa. (ejercicio 6).
Puntuación	No específica	Análisis detallado para cada ejercicio y puntuación correspondiente.	No específica	Puntos asignados por partes, no por ejercicio.	Detalle de puntos por ítem de cada ejercicio.

4.1.5. Proyecto de investigación (C_2^m)

C_2^m fue propuesto por un FPM (P1) y dos chatbots (ChatGPT y Gemini), siendo P1 quien presenta los objetivos más detallados en función de cada noción estadística involucrada en C_2^m . ChatGPT formula un objetivo general sobre población y muestra, y Gemini no explicita. P1 propone un tema de interés relevante para los jóvenes, como lo es el consumo de tabaco, mediante la lectura e interpretación de resultados de una encuesta mundial. Luego, solicita que los estudiantes diseñen una encuesta para relevar la situación en su escuela. El tema de interés de ChatGPT refiere a la educación de niños en comunidades rurales de Argentina. Gemini, si bien da ejemplo de posibles temas, afirma que es a elección de los estudiantes. Previo al diseño del dispositivo, tanto el FPM como ChatGPT y Gemini, aluden a nociones de población y muestra. P1 propicia una discusión para distinguir estas nociones tanto en la encuesta mundial como en la escolar. Gemini solicita ejemplo de cada una, además hace hincapié en la selección de una muestra representativa, y justificación del método de muestreo. ChatGPT requiere además la definición de población y muestra. P1 y Gemini proponen diseñar un dispositivo de recolección de datos: P1 propone un cuestionario y Gemini una encuesta con preguntas claras, concisas y relevantes. Además, P1 sugiere realizar una prueba piloto con los estudiantes de la clase, y luego, implementarla en toda la escuela. ChatGPT sugiere ejemplificar preguntas de una encuesta o entrevista. La recolección de datos a partir del cuestionario (P1) se propone en horario de clase, mientras que Gemini no indica en qué momento, pero sugiere varias modalidades: presencial, online o por teléfono. ChatGPT recomienda indicar cómo se recopilarían los datos sin dar precisiones como P1 y Gemini. Para el análisis de los datos, P1 alude a la identificación de variables cualitativas y cuantitativas (continuas y discretas), ChatGPT poco específica al respecto y Gemini propone analizar los datos con softwares estadísticos. Respecto a la transformación de los datos, tanto P1 como ChatGPT sugieren realizar tablas, gráficos u otros recursos visuales, mientras que Gemini, el uso de software. La presentación de resultados se realiza, según P1, mediante una infografía, spot para la radio escolar y nota para la misma revista. Ambos chatbots solicitan un informe escrito con algún tipo de estructura y formato. Los criterios de evaluación de P1 provienen claramente de un proceso de estudio a lo largo de las clases, no de un momento determinado de la actividad. Detalla siete aspectos a ponderar. ChatGPT no especifica ningún criterio mientras que Gemini da un detalle de cinco aspectos. Finalmente, la asignación de puntos para P1 es valorativa en cada etapa del proyecto, ChatGPT asigna puntos a cada fase y Gemini no explicita. La Tabla 5 sintetiza esta descripción.

Tabla 5

Descriptores de C_2^m

	P1	ChatGPT	Gemini
Objetivo	6 objetivos específicos detallados para cada noción.	Un objetivo muy general	No explicita.
Tema de interés	¿Cómo afecta el cigarrillo en la economía y salud de los estudiantes?" Etapa 1: responder preguntas sobre una encuesta real. Etapa 2: diseñar una encuesta para estudiantes de su escuela.	Educación de niños en comunidades rurales de Argentina: analizar las condiciones de estudio.	A elección de los estudiantes.
Definición y/o selección de la población y muestra	Distinguir entre la población de la etapa 1 y la de la etapa 2.	Definir población y muestra en general. Determinar la población. Seleccionar una muestra representativa. Identificar y justificar el método del muestreo. Calcular el tamaño de la muestra.	Determinar la población. Seleccionar la muestra representativa y justificar el método de muestreo.
Diseño del dispositivo	Diseñar, evaluar (prueba piloto) e implementar un cuestionario a estudiantes de la escuela.	Ejemplificar con preguntas de encuesta o entrevista.	Diseñar encuesta a los estudiantes con preguntas claras, concisas y relevantes.
Recolección y análisis de datos	Encuestar en el horario de clase. Identificar variables cualitativas y cuantitativas discretas y continuas.	Indicar cómo se recopilarían los datos, posibles preguntas y posible análisis.	Aplicar la encuesta (presencial, online o por teléfono).
Transformación de Datos	Realizar tablas, gráficos y extraer conclusiones.	Sugiere usar recursos visuales (tablas, gráficos u otros) y proporcionar recomendaciones.	Analizar datos con softwares estadísticos.
Presentación de resultados	Infografía, spot para la radio escolar y nota para la revista escolar.	Informe escrito, con una extensión entre 3 y 5 páginas (sin incluir gráficos o tablas), formato PDF. Presentación oral en clase.	Informe estructurado en introducción, metodología, resultados y conclusiones.
Criterio de Evaluación	Evaluación como proceso. Detalle de 7 aspectos.	No especifica	Detalle de 5 aspectos.
Puntuación	Valorativa a cada etapa del proyecto.	Asigna puntos a cada etapa del proyecto.	No explicita.

En síntesis, el FPM, P1, propone la implementación del proyecto (C_2^m) con objetivos específicos bien detallados, con un tema de interés muy relevante, proveniente de datos mundiales reales, con criterios de evaluación bien definidos y con una implementación escolar clara y novedosa tanto en el desarrollo como en la difusión de los resultados. ChatGPT se queda en el orden de la sugerencia, con muy pocas (casi nulas) especificaciones. Gemini podría ponderarse en un intermedio entre P1 y ChatGPT en términos de descriptores del proyecto.

5. DISCUSIÓN Y CONCLUSIONES

En lo que respecta a los objetivos 1 y 2, caracterizar y comparar los medios, técnicas e instrumentos de evaluación propuestos por tres FPM y por tres chatbots, sobre nociones de estadística (población y muestra), concluimos que Gemini es el único que propone los tres componentes del sistema evaluativo: 4 medios (C_1^m : prueba escrita, C_2^m : proyecto de investigación, C_7^m : debate y C_8^m : software de estadística), 1 técnica (C_1^t : autoevaluación) y 2 instrumentos (C_1^i : infografía y C_2^i : recurso audio/video educativo). Copilot creativo, es quien ofrece la mayor variedad de medios (todos excepto examen oral y debate) y una técnica (C_1^t : autoevaluación) pero no propone ningún tipo de instrumento. Los restantes chatbots (ChatGPT, Copilot preciso y equilibrado) se limitan sólo al uso de medios, sin proponer ni técnicas ni instrumentos. Para el caso de los futuros profesores, P1 considera dos componentes de evaluación diferentes: dos medios, *proyecto de investigación* (C_2^m) y *debate* (C_7^m), y dos instrumentos, *infografía* (C_1^i) y *recursos audio/video educativo* (C_2^i) sin aludir a las técnicas. P2 y P3 únicamente postulan un medio: prueba escrita (C_1^m).

Los chatbots propusieron mayor variedad de medios de evaluación que los FPM. Una razón posible sería la forma en que los chatbots generan las respuestas, a partir de una gran variedad de información con la que fueron entrenados. Por otra parte, la preponderancia de los FPM (dos de tres) en la elección de la prueba escrita como único medio de evaluación, puede deberse a las restricciones asociadas a la escasez de tiempo, como una de las principales variables.

Al continuar interactuando con los chatbots, con la intención de obtener ejemplos bien elaborados de sistemas evaluativos, de los 8 tipos de medios de evaluación identificados en interacciones previas, los únicos medios ofrecidos por todos los chatbots y los FPM (excepto P1) como ejemplos específicos fueron *prueba escrita* (C_1^m) y *proyecto de investigación* (C_2^m).

Respecto al análisis de la prueba escrita (C_1^m), propuestas por P2, P3 y por Copilot (en sus tres modalidades), se advierte una mayor potencialidad de las tareas propuestas por los chatbots, que por los FPM. Mientras los ejercicios de los FPM se resuelven mediante un cálculo, como, por ejemplo, parámetros de posición y de dispersión, las de los chatbots van más allá, pues su resolución requiere pensar, reflexionar, analizar, definir y justificar. En este sentido, el uso de las IAGen mejoraría el desempeño de los futuros profesores de matemática en el diseño de evaluaciones de tipo prueba escrita.

En cambio, respecto al análisis del proyecto de investigación (C_2^m), sugerido por P1, ChatGPT y Gemini, la propuesta del primero es más potente que la de los chatbots: ofrece mayores detalles, con objetivos bien delimitados y precisos, una temática relevante para los estudiantes con un método de recolección de datos y comunicación de resultados novedosa para ese medio de evaluación. Por otra parte, los chatbots, presentan para este caso, un proyecto mucho más acotado y estático.

Finalmente, respecto al tercer objetivo: determinar la funcionalidad de los chatbots como posibles asistentes del profesor para la generación de diferentes tipos de medios, técnicas e instrumentos de evaluación; concluimos que Gemini resultó ser el chatbot que sugiere los tres componentes del sistema evaluativo definidos por Hamodi et. al (2015). De esta forma,

Gemini podría ser considerado el mejor de los tres chatbots, al momento de pensar en un asistente para diseñar el proceso evaluativo.

Por otro lado, para el caso de las propuestas evaluativas específicas; las que formularon los FPM son superiores en la categoría C_2^m (proyecto de investigación), y las de los chatbots en la categoría C_1^m (prueba escrita). Esto permite pensar a los chatbots como asistentes valiosos a la hora de crear pruebas escritas, ya que pueden ofrecer ejercicios matemáticos complejos.

Una posible limitación de este trabajo es la cantidad de estudiantes para profesor con los cuales se llevó a cabo el estudio. Las conclusiones obtenidas se restringen a esos tres estudiantes y se requeriría de una muestra mucho más amplia para poder generalizar los resultados. Sin embargo, es aceptable desde el punto de vista de la investigación cualitativa ya que hemos considerado un estudio de casos, conformado por los tres FPM y los tres chatbots. Estas limitaciones pretenden ser abordadas en futuros trabajos. Nos proponemos, como objetivos a corto y mediano plazo, continuar indagando con los futuros profesores de matemática y con profesores en ejercicio. Además, se espera poder integrar los resultados obtenidos dentro de una situación de aprendizaje con estudiantes del nivel secundario de Argentina.

6. REFERENCIAS

- Adair, A., Pedro, M.S., Gobert, J., Segan, E. (2023). Real-Time AI-Driven Assessment and Scaffolding that Improves Students' Mathematical Modeling during Science Investigations. *Artificial Intelligence in Education*. AIED 2023.
- Adiguzel, T., Kaya, M. H., Cansu, F. K. (2023). Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. *Contemporary Educational Technology*, 15(3), ep429.
- Álvarez, J. (2005). *Evaluar para conocer, examinar para excluir*. Madrid: Morata.
- Brown, T., Mann, B., Ryder, N., Subbiah, J. Kaplan, M., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A. Krueger, G., Henighan, t., Child, R., Ramesh, A., Ziegler, D., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, L., Amodei, D. (2020). Language models are few-shot learners. In: *Proceedings of the 34th International Conference on Neural Information Processing Systems (NIPS'20)*, 1877–1901.
- Chevallard, Y. (2004). *Le moment de l'évaluation, ses objets, ses fonctions : déplacements, ruptures, refondation*. Journée de formation de formateurs. IUFM d'Aix-Marseille, 1-6 http://yves.chevallard.free.fr/spip/spip/article.php3?id_article=44
- Chevallard, Y. (2012). ¿Cuál puede ser el valor de evaluar? Notas para desprenderse de la evaluación 'como capricho y miniatura'. En G. Fioriti, C. Cuesta [comps.]. *La evaluación como problema. Aproximaciones desde las didácticas específicas* (pp. 9-21). Buenos Aires: Miño y Dávila-UNSAM Edita.

- Chowdhery, A., Narang, S., Devlin, J., Bosma, M., Mishra, G., Roberts, A..., and Fiedel, N. (2022). PaLM: Scaling language modeling with pathways, *Journal of Machine Learning Research* 24(240), 1–113.
- Fernández, J. (2006). ¿Evaluación? No gracias, calificación. *Cuadernos de Pedagogía*, 243, 92-97.
- Flores Samaniego, A., Gómez Reyes, A. (2009). Aprender Matemática, haciendo Matemática: la evaluación en el aula. *Educación Matemática*, 21(2), 117-142.
- Hamodi, C., López Pastor, V., López Pastor, A. (2015). Medios, técnicas e instrumentos de evaluación formativa y compartida del aprendizaje en educación superior. *Perfiles educativos*, 37(147), 146-161. Isaza, G. (2002). *Análisis, Interpretación y Construcción Teórica en la Investigación Cualitativa*. Centro de educación a distancia. Universidad de Manizales.
- Luzano, J. (2024). Assessment in Mathematics Education in the Sphere of Artificial Intelligence: A Systematic Review on Its Threats and Opportunities. *International Journal of Academic Multidisciplinary Research (IJAMR)*. 8(2),100-104.
- Martínez-Comesaña, M., Rigueira-Díaz, X., Larrañaga-Janeiro, A., Martínez-Torres, J., Ocaranza-Prado, I., Kreibel, D. (2023). Impacto de la inteligencia artificial en los métodos de evaluación en la educación primaria y secundaria: revisión sistemática de la literatura. *Revista de Psicodidáctica*. 18(2), 93-103.
- Méndez-Mantuano, M. O., Morán, M. Y. O., Mayorga, I. I. C., Valdez, A. Y. L., Rosado, Ángel R. H., & Robles, D. V. A. (2024). La evaluación académica en la era de la inteligencia artificial (IA). *South Florida Journal of Development*, 5(1), 119–148.
- Monash University (4 de abril de 2024). AI and assessment. <https://www.monash.edu/learning-teaching/teachhq/Teaching-practices/artificial-intelligence/ai-and-assessment>
- Nasution, N. E. A. (2023). Using artificial intelligence to create biology multiple choice questions for higher education. *Agricultural and Environmental Education*, 2(1), em002.
- Owan, V. J., Abang, K.B., Idika, D.O., Etta, E.O., Bassey, B.A. (2023). Exploring the potential of artificial intelligence tools in educational measurement and assessment. *EURASIA Journal of Mathematics, Science and Technology Education*, 19(8), em2307.
- Parra, V., Sureda, P., Corica, A., Schiaffino, S., Godoy, D. (2024). Can generative AI solve Geometry problems? Strengths and weaknesses of LLMs for geometric reasoning in Spanish. *International Journal of Interactive Multimedia and Artificial Intelligence*. 8(5), 65-74.
- Sánchez Mendiola, M. (2023). La inteligencia artificial generativa y la evaluación: ¿Qué pasará con los exámenes? *Investigación en Educación Médica*, 12(48), 5-8.

- Sanmartí, N. (2007). *10 ideas clave: evaluar para aprender*. Madrid: Graó
- Santos Guerra, M. (2003). *Una flecha en la diana: la evaluación como aprendizaje*. Madrid: Narcea.
- Scriven, M. (1967). The methodology of evaluation. En Stake, R: Perspectives of curriculum evaluation, *American Educational Research Association. Monograph series on curriculum*. Rand McNally, 38-39.
- Tobler, S. (2024). Smart grading: A generative AI-based tool for knowledge-grounded answer evaluation in educational assessments, *MethodsX*, 12, 102531.
- U.S. Department of Education (2023). *Office of Educational Technology, Artificial Intelligence and Future of Teaching and Learning: Insights and Recommendations*, Washington, DC, USA.
- Wan, T. & Chen, Z. (2024). Exploring Generative AI assisted feedback writing for students' written responses to a physics conceptual question with prompt engineering and few-shot learning. *arXiv:2311.06180*.
- Webb, N. (1992). Assessment of Students Knowledge of Mathematics. Steps Toward a Theory. En D. Grouws (Ed.) *Handbook of Research on Mathematics Teaching and Learning*. New York: Macmillan.

Para citar este artículo:

Sureda, P., Corica, A., Parra, V., Godoy, D., y Schiaffino, S. La evaluación en educación matemática: aportes de chatbots y futuros profesores de matemática. *EduTec, Revista Electrónica de Tecnología Educativa*, (89) 64-83.
<https://doi.org/10.21556/edutec.2024.89.3243>



Un análisis experimental de la relación entre las evaluaciones proporcionadas por la inteligencia artificial y las proporcionadas por los docentes en formación

An experimental analysis of the relationship between the evaluations of artificial intelligence and pre-service teachers

 Héctor Galindo-Domínguez¹; hector.galindo@ehu.eus;

 Nahia Delgado¹; nahia.delgado@ehu.eus;

 Martín Sainz de la Maza¹; martin.sainzdelamaza@ehu.eus;

 Ernesto Expósito²; ernesto.exposito@univ-pau.fr;

Resumen

Uno de los beneficios potenciales de la inteligencia artificial (IA) es que puede permitir la optimización de las tareas de los docentes. Este estudio tuvo como objetivo analizar las posibles diferencias entre las evaluaciones realizadas por docentes en formación y las realizadas por diferentes IA. Participaron un total de 507 docentes en formación, a quienes se les proporcionó una rúbrica para evaluar 12 textos de distintos tipos y calidades. Los resultados mostraron cómo el desempeño de las IA en la evaluación de tareas escritas replicó con bastante precisión el funcionamiento de los docentes en formación, siendo ChatGPT la IA que mejor replicó el comportamiento de los docentes en formación, con una precisión cercana al 70% de la evaluación proporcionada por humanos. Del mismo modo, hubo diferencias mínimas en las evaluaciones realizadas por los docentes en formación según su género y año académico. Asimismo, las evaluaciones realizadas por los docentes en formación con mejor desempeño estuvieron más alineadas con las proporcionadas por la IA en comparación con los estudiantes con menor desempeño. Estos resultados son útiles, al destacar cómo la IA podría ser una herramienta de apoyo que guíe el conocimiento pedagógico de los docentes en formación en tareas de evaluación.

Palabras clave: Evaluación, Inteligencia Artificial, ChatGPT, Formación Docente

Abstract

One of the potential benefits of artificial intelligence (AI) is its ability to optimize teachers' tasks. The aim of this study was to analyze the possible differences between assessments carried out by pre-service teachers and those performed by various AI systems. A total of 507 pre-service teachers participated, and they were provided with a rubric to evaluate 12 texts of different types and quality. The results showed that AI performance in evaluating written tasks closely replicated the functioning of pre-service teachers, with ChatGPT being the AI that most accurately mirrored the teachers' evaluations, achieving approximately 70% precision compared to human assessments. Similarly, there were minimal differences in the assessments made by pre-service teachers based on gender and academic year. Moreover, evaluations conducted by higher-performing pre-service teachers were more aligned with those provided by AI, compared to those from lower-performing students. These findings are valuable, highlighting how AI could serve as a supportive tool to guide the pedagogical knowledge of pre-service teachers in assessment tasks.

Keywords: Assessment, Artificial Intelligence, ChatGPT, Teacher Training

¹ Universidad del País Vasco/Euskal Herriko Unibertsitatea (España)

² Université de Pau et des Pays de l'Adour (Francia)



1. INTRODUCCIÓN

1.1. La evaluación en la formación docente: responsabilidades, desafíos y factores influyentes

En la Ley Orgánica 3/2020, la principal ley educativa en España, el artículo 91 detalla las funciones del profesorado, incluyendo la evaluación de los procesos de aprendizaje de los estudiantes. Trabajos previos, como el de Stiggins (2014), han demostrado que los docentes invierten entre un tercio y hasta la mitad de su tiempo profesional en tareas relacionadas con la evaluación y calificación, lo que muestra que es una tarea que consume mucho tiempo, especialmente con el aumento de la ratio de estudiantes por profesor (Ramesh y Kumar, 2022). Sin embargo, a pesar de ser una labor que requiere esfuerzo y dedicación, se ha observado que un aumento en la alfabetización evaluativa de los docentes tiene un impacto directo en los resultados de aprendizaje de los estudiantes (por ejemplo, Mellati y Khademi, 2018; Xu y Brown, 2016).

Considerando esta función que cualquier docente a nivel mundial realiza con sus estudiantes, enseñar a los futuros docentes nuevos conocimientos y habilidades para evaluar a sus futuros alumnos podría ser esencial para hacer la tarea de enseñanza lo más eficiente posible (Atjonen et al., 2022). En esta misma línea, estudios previos muestran importantes limitaciones en la manera en que se enseña a los docentes en formación a evaluar, ya sea porque es excesivamente teórico o porque está desconectado de las tareas diarias de un docente (Atjonen, 2017; DeLuca et al., 2019; Salama y Subahi, 2020). Como resultado, en muchos casos, los futuros docentes aplican estrategias de evaluación que se utilizaron con ellos cuando eran estudiantes (Hill et al., 2017).

Asimismo, la cantidad de conocimientos y habilidades de un docente en formación al evaluar trabajos podría estar condicionada por una serie de variables personales y académicas. Aunque la evidencia existente hasta la fecha es escasa, algunos estudios, como el de Salama y Subahi (2020), observaron que la alfabetización evaluativa de los docentes en formación era relativamente baja y similar, independientemente de su género, rendimiento académico o años de experiencia. Además, Lovorn y Reza (2011) observaron cómo la formación recibida puede influir en la forma en que se lleva a cabo la evaluación mediante rúbricas, y Deneen y Brown (2016) notaron cómo el rendimiento académico de los docentes en formación desempeña un papel determinante en la profundidad de una evaluación. No obstante, debemos considerar que el grado en que se detalla la evaluación de una tarea puede estar influenciado por el rendimiento académico, que a su vez podría deberse al impacto de otros factores socioemocionales, como la motivación de los estudiantes hacia la tarea o su estado emocional actual (Eklöf, 2010). Por esta razón, el grado de detalle en la evaluación de tareas podría ser un fenómeno multidimensional.

1.2. La integración de la inteligencia artificial en la evaluación digital

Con los avances tecnológicos de los últimos años, algunos de los métodos de evaluación que se están utilizando para abordar problemas como la alta ratio de estudiantes por docente, la instrucción personalizada y la reducción del tiempo excesivo, implican el uso de sistemas basados en inteligencia artificial (por ejemplo, Vij et al., 2019).

La inteligencia artificial (IA de aquí en adelante) es la capacidad de una máquina para replicar el comportamiento humano inteligente (Wang, 2019). Hoy en día, existe una abundancia de herramientas basadas en IA. Dentro del campo de la IA, un área en expansión es la IA generativa, que se entiende como una IA enfocada en crear contenido basado en las entradas del usuario. Algunas de las IA generativas más importantes en la actualidad son ChatGPT, de OpenAI, Copilot de Bing, de Microsoft, o Gemini, de Google, por nombrar algunas.

La IA generativa tiene aplicaciones potenciales en la evaluación de tareas estudiantiles, ya que diversos estudios han demostrado su precisión al proporcionar retroalimentación. Estos sistemas de IA, apoyados por el procesamiento de lenguaje natural, ofrecen respuestas personalizadas para complementar el esfuerzo de los docentes (Ocaña-Fernández et al., 2019; González-Calatayud et al., 2021). Jani et al. (2020) destacaron el papel de la IA en la evaluación formativa, utilizando el aprendizaje automático para monitorear el progreso de los estudiantes y mejorar las prácticas clínicas. De manera similar, la IA se ha aplicado en la formación médica (Mirchi et al., 2020), la educación en ingeniería (Samarakou et al., 2016; Liu et al., 2017) y las evaluaciones de programación (Grivokostopoulou et al., 2017), proporcionando retroalimentación automatizada basada en el rendimiento. Otros estudios (Rhienmora et al., 2011; Oguengay et al., 2015; Ulum, 2020; Choi y McClenen, 2020) demuestran además la capacidad de la IA para calificar y evaluar habilidades en diversos campos.

Asimismo, ciertos estudios han comparado el efecto de utilizar sistemas basados en IA frente a no utilizarlos. Por ejemplo, Grivokostopoulou et al. (2017) realizaron una comparación entre los resultados obtenidos por una IA y las evaluaciones manuales realizadas por docentes para verificar la precisión de esta tecnología. Los resultados mostraron una correlación entre ambos, con solo ligeras diferencias observadas en trabajos excelentes donde los docentes tendían a sobrevalorar la tarea en comparación con las puntuaciones otorgadas por la IA. Estos resultados son coherentes con los obtenidos por Houtao et al. (2022), quienes observaron cómo la retroalimentación de la IA generativa podría ser tan útil como la de los docentes, aunque con algunas diferencias. En particular, en la corrección de textos, mientras que la retroalimentación del docente tendía a centrarse en la estructura y el contenido de la tarea, la retroalimentación de la IA era más detallada en vocabulario y gramática. Estos hallazgos subrayan el valor potencial de integrar ambas formas de retroalimentación para garantizar una evaluación más exhaustiva.

Como comentó Dillenbourg (2016), la transición de la educación tradicional a la digital no significa la obsolescencia de los docentes en el futuro. En lugar de debatir si la IA reemplazará a los docentes, Hrastinski et al. (2019) proponen reconocer los beneficios potenciales de la IA y cómo estas ventajas podrían redefinir su papel en el aula. Por lo tanto, en la evaluación educativa, los docentes continúan desempeñando un papel crucial en garantizar la utilización adecuada de la IA para los objetivos de medición y evaluación. Algunas de estas responsabilidades incluyen la creación de evaluaciones y el establecimiento de objetivos de aprendizaje, la contextualización de las preguntas de evaluación para hacerlas más relevantes y significativas para los estudiantes, la interpretación de los resultados para proporcionar retroalimentación personalizada adaptada a las fortalezas y debilidades de los estudiantes, y el monitoreo del progreso estudiantil, entre otras (Owan et al., 2023).

1.3. Propósito del estudio

Como se ha observado, la gran mayoría de los estudios mencionados anteriormente utilizan sistemas basados en IA (generativa) para proporcionar retroalimentación a los estudiantes, pero no comparan experimentalmente dicha retroalimentación con la que podrían proporcionar los docentes. Asimismo, los estudios que realizan análisis comparativos entre evaluaciones proporcionadas por IA y las realizadas por docentes en activo son mínimos (por ejemplo, Grivokostopoulou et al., 2017; Houtao et al., 2022), pero, según el conocimiento de los autores, no existen estudios que evalúen la relación entre las evaluaciones proporcionadas por IA generativa y las de los docentes en formación. Además, revisiones sistemáticas recientes han demostrado el creciente interés de la comunidad científica en el uso de sistemas de inteligencia artificial y aprendizaje automático para automatizar la puntuación de ensayos, con el fin de abordar el aumento de la proporción de estudiantes y las tareas que consumen mucho tiempo, como la retroalimentación y la calificación (Ramesh y Kumar, 2022). Por esta misma razón, y dado que la literatura sobre el tema es escasa, es imperativo determinar si dichas herramientas están o no preparadas para servir como una ayuda adicional en el conocimiento pedagógico del docente durante sus procesos de evaluación digital.

Basándose en estas necesidades, los objetivos de este estudio son:

- O1: Analizar si existen diferencias estadísticamente significativas entre las evaluaciones realizadas por la IA generativa y las evaluaciones realizadas por los docentes en formación sobre diferentes textos escritos.
- O2: Analizar si las diferencias en las evaluaciones entre la IA generativa y los docentes en formación dependen de su género.
- O3: Analizar si las diferencias en las evaluaciones entre la IA generativa y los docentes en formación dependen de su nivel de formación.
- O4: Analizar si las diferencias en las evaluaciones entre la IA generativa y los docentes en formación dependen de su rendimiento académico.

2. MÉTODO

2.1. Participantes

Un total de 507 estudiantes universitarios participaron en el presente estudio, con una edad promedio de 20.56 años ($DT = 5.42$). De ellos, 155 eran hombres, 348 mujeres y 4 personas que no se identificaron con ninguna categoría de género específica. En cuanto a la titulación académica, 130 estudiantes provenían del Grado en Educación Infantil, 327 del Grado en Educación Primaria y 50 de áreas relacionadas como Pedagogía o Educación Social. En términos de progreso académico, 172 estudiantes estaban en su primer año, 137 en su segundo, 168 en su tercero, 25 en su cuarto y 5 en su quinto año de estudios. Una parte de la muestra fue seleccionada por proximidad, compuesta por estudiantes de los investigadores involucrados en este estudio. Otra parte de la muestra fue seleccionada mediante la difusión de un mensaje institucional que invitaba a la participación de los estudiantes matriculados en los programas de grado ofrecidos por las tres Facultades de Educación de la Universidad del País Vasco. A pesar

de que la muestra no es probabilística, existe evidencia previa que indica que las muestras por conveniencia pueden arrojar resultados similares a los obtenidos de muestras aleatorizadas (por ejemplo, Coppock et al., 2018).

2.2. Instrumentos

Para recopilar todos los datos, se utilizaron dos instrumentos diferentes. En primer lugar, se solicitó a los participantes información sobre variables personales como el género, la edad, el año académico, el grado universitario que estaban cursando y su rendimiento académico. Para el objetivo O3, el nivel de formación se codificó en función del año académico del estudiante. Específicamente, los estudiantes de primer año se ubicaron en formación inicial, los de segundo año en formación intermedia y los de tercer y cuarto año en formación avanzada de grado universitario. Los estudiantes de cuarto año se agruparon con los de tercer año debido a que tienen menos asignaturas. Asimismo, para el objetivo O4, el rendimiento académico se midió a través de la media aritmética de las calificaciones obtenidas por los estudiantes en el año anterior al que estaban matriculados. Este valor se conoce a partir del expediente académico que se les envía al final de cada año. En base a sus puntuaciones, se crearon grupos de bajo (percentil < 33), medio (percentil 34 a 66) y alto (percentil > 67) rendimiento académico.

En segundo lugar, se generaron una serie de textos *ad-hoc* escritos en español utilizando ChatGPT 3.5, y posteriormente fueron supervisados por expertos en educación. Estos textos fueron evaluados por docentes en ejercicio (n = 3). El *prompt* utilizado para generar los textos fue el siguiente:

Escribe un texto [tipo de texto] de 5 a 10 líneas escrito por un estudiante de 10 años sobre un tema de libre elección, con contenido, organización, vocabulario, coherencia y cohesión de [tipo de calidad].

En total, se desarrollaron 12 textos de diferentes tipos (3 textos descriptivos, 3 textos narrativos, 3 textos argumentativos y 3 textos instructivos) y calidades (4 excelentes, 4 normales y 4 significativamente mejorables). Los textos generados se encuentran en el Anexo I. Además, se creó una rúbrica *ad-hoc* con ChatGPT 3.5, también supervisada por docentes en ejercicio (n = 3).

En cuanto a la evaluación de ensayos mediante IA, varios estudios han analizado los principales criterios que los docentes deben utilizar. En este sentido, algunos criterios están relacionados con características estadísticas, como la longitud del ensayo en relación con el número de palabras, la longitud de las oraciones o la longitud promedio de las palabras (Contreras et al., 2018; Kumar et al., 2019; Mathias y Bhattacharyya, 2018). Respecto a los criterios de estilo o basados en la sintaxis, los más relevantes incluyen la estructura de la oración, la puntuación, la gramática, los operadores lógicos y el vocabulario (Cummins et al., 2016; Darwish y Mohamed, 2020; Ke et al., 2019). Asimismo, en cuanto a los criterios basados en el contenido, los más relevantes son la cohesión entre oraciones en un documento, la relevancia de la información, la corrección y la consistencia (Dong et al., 2017).

Basándose en estos criterios, se construyó la herramienta *ad-hoc* con 4 criterios diferentes (Contenido, Organización, Vocabulario, Coherencia y Cohesión) y 4 niveles de logro (excelente, bueno, regular y pobre). Por lo tanto, cada texto podría obtener una puntuación máxima de 16 puntos (4 criterios x 4 niveles de logro), aunque, para fines prácticos y alinearse con el sistema

educativo español, estos 16 puntos se ponderaron en una escala de 10, siendo 10 la máxima calificación para cada texto (equivalente a 16 puntos no ponderados). Los niveles de logro fueron establecidos por ChatGPT basados en criterios derivados de trabajos empíricos previos ya mencionados. Estos niveles fueron revisados por los investigadores, mostrando un alto acuerdo entre los proporcionados por la IA y los encontrados en otras rúbricas de evaluación de trabajos escritos (por ejemplo, Gobierno de Terranova y Labrador, 2014).

Esta rúbrica, recogida en el Anexo II, fue creada para ayudar a los docentes en formación a evaluar los textos generados. Todos los estudiantes participantes habían recibido formación específica sobre cómo usar rúbricas de evaluación, ya que a lo largo de su grado cursan asignaturas de didáctica general y didáctica específica en las que se les enseña a diseñar, interpretar y aplicar rúbricas de evaluación. Este aspecto es importante, ya que estudios previos han demostrado que un docente en formación sin capacitación en el uso de rúbricas emite una evaluación tan subjetiva como un docente en formación sin ninguna herramienta de evaluación (Lovorn y Reza, 2011). La fiabilidad de los diferentes textos usando la rúbrica de evaluación se muestra en la Tabla 1. Como se puede observar, dado que todos los valores están por encima del punto de corte de $\alpha > .70$ (Tavakol y Dennick, 2011), se puede asumir que existe una buena consistencia interna en los criterios utilizados para la evaluación de cada texto.

Tabla 1

Índices de fiabilidad para cada texto

Orden	Texto	Caldiad	Fiabilidad
1	Texto descriptivo	Significativamente mejorable	.792
2	Texto argumentativo	Normal	.801
3	Texto instructivo	Significativamente mejorable	.741
4	Texto narrativo	Excelente	.904
5	Texto descriptivo	Excelente	.904
6	Texto argumentativo	Significativamente mejorable	.781
7	Texto instructivo	Excelente	.922
8	Texto narrativo	Normal	.816
9	Texto descriptivo	Normal	.857
10	Texto argumentativo	Excelente	.875
11	Texto instructivo	Normal	.869
12	Texto narrativo	Significativamente mejorable	.814

Finalmente, se solicitó a cada una de las IAs analizadas que evaluara los diferentes textos utilizando el siguiente *prompt*:

El siguiente texto ha sido producido por un estudiante de 10 años. Considerando los criterios de contenido, organización, vocabulario, y coherencia y cohesión, proporciona a cada texto una calificación del 0 al 10: [Texto elaborado por ChatGPT].

2.3. Procedimiento

El proceso comenzó con la creación de los diferentes textos. Estos textos fueron elaborados utilizando ChatGPT, siguiendo la estructura: Escribe un [tipo de texto] de 5 a 10 líneas escrito por un estudiante de 10 años sobre un tema de libre elección, con contenido, organización, vocabulario, coherencia y cohesión de [tipo de calidad]. Estos textos fueron supervisados por expertos en educación para detectar errores de coherencia, léxicos y gramaticales. Posteriormente, los textos fueron convertidos digitalmente a formularios de Google (*Google Forms*), y la rúbrica para evaluar cada texto también se añadió para permitir a los estudiantes evaluar cada criterio (contenido, organización, vocabulario, y coherencia y cohesión).

Inicialmente, se seleccionó un subconjunto de la muestra de los estudiantes del grupo de los investigadores. Asimismo, el equipo de investigación solicitó permiso a la decanatura para invitar a participar a todos los demás estudiantes que no estaban bajo su instrucción. En este proceso, se siguieron estrictamente todas las medidas éticas aceptadas en la Declaración de Helsinki. En ambos casos, se requería que los participantes revisaran y aceptaran los procedimientos y términos de participación antes de aportar cualquier dato, garantizando el cumplimiento de los estándares éticos. Este acuerdo incluía informar a los participantes sobre los objetivos del estudio, el tiempo de respuesta anticipado, el manejo confidencial y anónimo de los datos, y la naturaleza voluntaria de la participación, que incluía la opción de retirar sus respuestas durante la encuesta. Los datos del grupo de estudiantes de los investigadores se recopilaron durante el horario laboral regular, mientras que las respuestas de los estudiantes externos se aceptaron en cualquier momento. Finalmente, tras completar el análisis del estudio, se envió un informe que resumía los hallazgos clave a aquellos que expresaron su deseo de recibir los resultados.

2.4. Análisis de datos

El proceso de análisis de datos se llevó a cabo completamente utilizando el software estadístico *SPSS Statistics 27*. Inicialmente, considerando las respuestas de todos los participantes, se calcularon las puntuaciones para cada texto. Como se mencionó anteriormente, aunque cada texto podía recibir un máximo de 16 puntos (4 criterios x 4 niveles de logro), para facilitar la interpretación, estos 16 puntos se ponderaron a 10 (una puntuación de 10 equivale a 16 puntos en la rúbrica), dado que, en el sistema educativo español, un 10 representa la máxima calificación.

Luego, para abordar el primer objetivo, se calcularon las medias aritméticas y las desviaciones estándar para cada grupo (IA vs. docentes en formación). Posteriormente, para identificar posibles diferencias significativas, se realizó una prueba t de Student complementada con la d de Cohen para conocer el tamaño del efecto. Además, se calculó el porcentaje de precisión, que se define como el número de textos que se desviaron en menos de 1 punto de las medias aritméticas de los docentes en formación, dividido por el número total de textos. Si la precisión fuera del 100%, significaría que las evaluaciones proporcionadas por la IA coincidirían estrechamente (en todos los casos, con una desviación de menos de 1 punto) con las evaluaciones dadas por los docentes en formación.

Para abordar el segundo objetivo, se recalcularon las medias aritméticas y las desviaciones estándar para cada grupo, seguidas de un ANOVA de un factor para examinar posibles

diferencias entre los grupos. Este análisis se acompañó de una prueba post-hoc de Tukey para identificar entre qué grupos se hallaban diferencias significativas. Para el tercer objetivo, se siguió el mismo procedimiento que para el segundo.

Finalmente, para el cuarto objetivo, dado el rendimiento académico de los docentes en formación en una escala, se formaron tres grupos diferentes: bajo rendimiento (percentil 1 a 33), rendimiento medio (percentil 34 a 66) y alto rendimiento (percentil 67 a 99). Una vez clasificados, se realizó otro ANOVA de un solo factor para identificar posibles diferencias significativas. Este análisis se complementó con un análisis post-hoc utilizando la prueba de Tukey para determinar entre qué grupos se encontraron diferencias significativas.

3. RESULTADOS

Primero, en relación con el objetivo 1, para determinar si había diferencias significativas entre las puntuaciones asignadas a los diferentes textos utilizando la rúbrica de evaluación y las puntuaciones asignadas por las IAs, se realizaron diversas pruebas T.

Como se observa en la Tabla 2, de los 12 textos, se encontraron diferencias estadísticamente significativas en solo 4 de ellos entre las calificaciones proporcionadas por los docentes en formación y las IAs. Sin embargo, vale la pena señalar que, aunque no había diferencias significativas en la mayoría de los casos, en aquellos donde sí se encontraron diferencias, el tamaño del efecto fue grande [Argumentativo Excelente, $p = .036$, $d = 1.25$; Argumentativo Normal, $p = .029$, $d = 1.56$; Narrativo Excelente, $p < .001$, $d = .70$; Narrativo Normal, $p = .046$, $d = 1.43$].

También se analizó la precisión de varias IAs en replicar la capacidad evaluativa de los docentes en formación. El método utilizado para calcular la precisión de la IA consistió en contar el número de textos evaluados por la IA con una desviación de menos de 1 punto de la evaluación dada por los docentes en formación y dividirlo por el número total de textos (12). La desviación de 1 punto se seleccionó de manera arbitraria, ya que se entiende que puede haber un buen acuerdo entre la evaluación proporcionada por la IA y la evaluación dada por los docentes en formación cuando existe una desviación de menos de 1 punto de 10 entre las dos puntuaciones. Los resultados revelaron que ChatGPT fue la IA entre las analizadas que mejor replicó las calificaciones proporcionadas por los docentes en formación ($8/12 = 66.66\%$ de precisión), seguida de Gemini ($7/12 = 58.33\%$ de precisión), y Copilot de Bing en la peor posición ($6/12 = 50\%$ de precisión).

Tabla 2

Resultados principales de las pruebas T

Tipo	Calidad	Inteligencia Artificial				Docentes en Formación	p	d
		ChatGPT	Bing	Gemini	Total IA			
Descriptivo	Excelente	8+	9.25+	9+	8.75 (.661)+	8.95 (1.46)	(ns)	-
	Normal	6+	8.5	7+	7.16 (1.25)+	6.75 (1.77)	(ns)	-
	Bajo	4+	4.5+	6	4.83 (1.04)+	4.95 (1.43)	(ns)	-
Argumentativo	Excelente	9	9	8+	8.66 (.577)	7.08 (1.68)	*	1.25
	Normal	7	8.5	7	7.5 (.866)	5.52 (1.56)	*	1.56
	Bajo	6	4+	7	5.66 (1.52)	4.52 (1.32)	(ns)	-
Instructivo	Excelente	9+	8.75+	9+	8.91 (.144)+	9.08 (1.37)	(ns)	-
	Normal	5	8.75	8+	7.25 (1.98)+	7.11 (1.63)	(ns)	-
	Bajo	4+	3.5	5+	4.16 (.763)+	4.61 (1.44)	(ns)	-
Narrativo	Excelente	9+	8.75+	9+	8.91 (.144)+	8.08 (1.65)	***	.70
	Normal	7+	8.75	8	7.91 (.877)	6.07 (1.59)	*	1.43
	Bajo	3+	2.75+	5	3.58 (1.23)+	3.18 (1.09)	(ns)	-
Precisión		66.66%	50%	58.33%	66.66%			

Nota. +, La puntuación de las IAs está por debajo a 1 punto de desviación de la media de la evaluación proporcionada por los docentes en formación; * p < .05; ** p < .01; *** p < .001; (ns), no significativo.

Posteriormente, en relación con el objetivo 2, se realizaron diversas pruebas ANOVA de un factor, considerando tres grupos diferentes: evaluaciones proporcionadas por la IA, evaluaciones proporcionadas por docentes en formación masculinos y evaluaciones proporcionadas por docentes en formación femeninas. Como se observa en la Tabla 3, las diferencias entre géneros fueron mínimas, ya que solo en 3 de los 12 textos se encontraron diferencias significativas basadas en el género. De estas 3 diferencias por género, en 2 de ellas, las mujeres tuvieron una puntuación más similar a la proporcionada por la IA, mientras que en solo un caso los hombres tuvieron una puntuación más similar a la proporcionada por la IA. En los casos restantes, no hubo diferencias estadísticamente significativas basadas en el género.

A partir de este análisis, se calculó la precisión como el número de textos de los docentes en formación que tuvieron una media aritmética inferior a 1 punto en comparación con la media aritmética proporcionada por las IAs generativas. En este caso, la precisión para ambos géneros respecto a las evaluaciones proporcionadas por las IAs generativas fue del 66.6% (8/12 textos).

Tabla 3

Principales resultados de la ANOVA de un factor en base al género de los docentes en formación

Tipo	Calidad	Total IA (3)	Género		p	Post-Hoc
			Masculino (1)	Femenino (2)		
Descriptivo	Excelente	8.75 (.661)	8.72 (1.54) ₊	9.04 (1.42) ₊	(ns)	-
	Normal	7.16 (1.25)	6.62 (1.64) ₊	6.81 (1.82) ₊	(ns)	-
	Bajo	4.83 (1.04)	5.06 (1.49) ₊	4.90 (1.41) ₊	(ns)	-
Discutidor	Excelente	8.66 (.577)	6.87 (1.50)	7.17 (1.75)	(ns)	-
	Normal	7.5 (.866)	5.78 (1.63)	5.41 (1.51)	**	2<1; 1<3; 2<3
	Bajo	5.66 (1.52)	4.56 (1.37) ₊	4.50 (1.30)	(ns)	-
Instructivo	Excelente	8.91 (.144)	8.98 (1.36) ₊	9.13 (1.38) ₊	(ns)	-
	Normal	7.25 (1.98)	6.89 (1.34) ₊	7.21 (1.74) ₊	(ns)	-
	Bajo	4.16 (.763)	4.56 (1.43) ₊	4.63 (1.45) ₊	(ns)	-
Narrativo	Excelente	8.91 (.144)	7.81 (1.55)	8.20 (1.68) ₊	*	1<2
	Normal	7.91 (.877)	5.89 (1.47)	6.15 (1.64)	*	(ns)
	Bajo	3.58 (1.23)	3.29 (1.21) ₊	3.13 (1.04) ₊	(ns)	-
Precisión			8/12 (66.6%)	8/12 (66.6%)		

Nota. +, La puntuación de las IAs está por debajo a 1 punto de desviación de la media de la evaluación proporcionada por los docentes en formación; * p < .05; ** p < .01; *** p < .001; (ns), no significativo. Post-Hoc realizado a través de la prueba de Tukey.

A continuación, en relación con el objetivo 3, se realizaron varias pruebas ANOVA de un solo factor, considerando 4 grupos: las IAs, docentes en formación que comienzan su carrera universitaria (año 1), docentes en formación en medio de su carrera universitaria (año 2) y docentes en formación que concluyen su carrera universitaria (años 3 y 4). En términos generales, los análisis recogidos en la Tabla 4 revelan que las diferencias observadas entre los diversos grupos son mínimas, siendo en todos los casos 8 de las 12 evaluaciones (66.6% de precisión) similares a las proporcionadas por las IAs generativas.

Estos resultados pueden confirmar que, independientemente de la etapa en la que se encuentre el docente en formación, su capacidad para calificar y evaluar no será muy diferente. Las únicas excepciones fueron los textos descriptivos de baja calidad, donde los docentes en formación que concluían sus estudios sobrestimaron la puntuación proporcionada por sus compañeros y por la IA (p < .001), así como en los textos instruccionales excelentes, donde también sobreestimaron la puntuación proporcionada (p = .004).

Tabla 4

Principales resultados de la ANOVA de un factor en base al nivel de formación de los docentes en formación

Tipo	Calidad	IA Total (4)	Nivel de formación			p	Post-Hoc
			Inicial (1)	Medio (2)	Finalizando (3)		
Descriptivo	Excelente	8.75 (.661)	8.82 (1.53) ₊	9.01 (1.20) ₊	9.01 (1.56) ₊	(ns)	-
	Normal	7.16 (1.25)	6.52 (1.73) ₊	6.78 (1.75) ₊	6.92 (1.81) ₊	(ns)	-
	Bajo	4.83 (1.04)	4.76 (1.35) ₊	4.61 (1.19) ₊	5.34 (1.57) ₊	***	1<3; 2<3
Argumentativo	Excelente	8.66 (.577)	7.04 (1.65)	7.18 (1.63)	7.06 (1.76)	(ns)	-
	Normal	7.5 (.866)	5.69 (1.54)	5.48 (1.48)	5.41 (1.62)	*	(ns)
	Bajo	5.66 (1.52)	4.44 (1.28)	4.47 (1.23)	4.62 (1.40)	(ns)	-
Instructivo	Excelente	8.91 (.144)	8.79 (1.57) ₊	9.34 (1.00) ₊	9.16 (1.37) ₊	**	1<2; 1<3
	Normal	7.25 (1.98)	7.16 (1.74) ₊	7.07 (1.57) ₊	7.11 (1.59) ₊	(ns)	-
	Bajo	4.16 (.763)	4.81 (1.47) ₊	4.42 (1.23) ₊	4.57 (1.53) ₊	(ns)	-
Narrativo	Excelente	8.91 (.144)	8.04 (1.61) ₊	8.10 (1.70) ₊	8.10 (1.70) ₊	(ns)	-
	Normal	7.91 (.877)	6.16 (1.63)	5.96 (1.46)	6.07 (1.64)	(ns)	-
	Bajo	3.58 (1.23)	3.16 (1.10) ₊	3.07 (1.05) ₊	3.26 (1.11) ₊	(ns)	-
Precisión			8/12 (66.6%)	8/12 (66.6%)	8/12 (66.6%)		

Nota. +, La puntuación de las IAs está por debajo a 1 punto de desviación de la media de la evaluación proporcionada por los docentes en formación; * p < .05; ** p < .01; *** p < .001; (ns), no significativo. Post-Hoc realizado a través de la prueba de Tukey.

Finalmente, en relación con el objetivo 4, se realizaron varias pruebas ANOVA de un solo factor, considerando 4 grupos: las IAs, docentes en formación con bajo rendimiento académico (percentil < 33), docentes en formación con rendimiento académico medio (percentil 34 a 66) y docentes en formación con alto rendimiento académico (percentil > 67). Los resultados, como se observa en la Tabla 5, mostraron que los docentes en formación con un rendimiento académico más alto fueron evaluadores más precisos (9 de 12 textos, 75%) de las calificaciones proporcionadas por la IA, en contraste con los docentes en formación con rendimiento medio (8 de 12 textos, 66.6%) o bajo rendimiento académico (7 de 12 textos, 58.3%).

Asimismo, como se puede observar en las diferencias de medias presentadas en la Tabla 5, el número de textos con puntuaciones más altas aumenta con un mayor rendimiento académico. Por lo tanto, se puede observar que los estudiantes con un rendimiento más alto tienden a evaluar los textos escritos de manera más favorable en comparación con los estudiantes con un rendimiento académico más bajo.

Además, se puede ver que las IAs analizadas tendieron a sobreestimar las puntuaciones en la mayoría de los textos, independientemente del rendimiento académico de los estudiantes. Así, el número de textos sobreestimados por la IA en comparación con las evaluaciones de los estudiantes con bajo rendimiento fue de 10 de 12 textos (5 de ellos con más de 1 punto de

desviación de la media), en el caso de estudiantes con rendimiento medio fue de 8 de 12 textos (4 de ellos con más de 1 punto de desviación de la media), y en el caso de estudiantes con alto rendimiento fue de 7 de 12 textos (3 de ellos con más de 1 punto de desviación de la media). Estos datos demuestran cómo el rendimiento académico podría ser un factor importante en la generación de evaluaciones más alineadas con las proporcionadas por la IA.

Tabla 5

Principales resultados de la ANOVA de un factor en base al rendimiento académico de los docentes en formación

Tipo	Calidad	IA Total (4)	Rendimiento académico						p	Post-Hoc
			Bajo P<33 (1)		Medio P 34-66 (2)		Alto P>67 (3)			
			M (DT)	Dif M ¹	M (DT)	Dif M ¹	M (DT)	Dif M ¹		
Descriptivo	Excelente	8.75 (.661)	8.72 (1.52) ₊	.03	8.88 (1.50) ₊	-.13	9.35 (1.24) ₊	-.60	**	1<3; 2<3
	Normal	7.16 (1.25)	6.41 (1.73) ₊	.75	6.84 (1.81) ₊	.32	7.08 (1.69) ₊	.08	**	1<3
	Bajo	4.83 (1.04)	4.74 (1.39) ₊	.09	4.84 (1.39) ₊	-.01	5.39 (1.48) ₊	-.56	***	1<3; 2<3
Argumentativo	Excelente	8.66 (.577)	6.87 (1.64)	1.79	7.11 (1.75)	1.55	7.34 (1.61)	1.32	*	(ns)
	Normal	7.5 (.866)	5.48 (1.54)	2.02	5.43 (1.56)	2.07	5.71 (1.56)	1.79	(ns)	-
	Bajo	5.66 (1.52)	4.40 (1.36)	1.26	4.44 (1.19)	1.22	4.80 (1.40) ₊	.86	*	1<3
Instructivo	Excelente	8.91 (.144)	8.94 (1.40) ₊	-.03	9.06 (1.43) ₊	-.15	9.31 (1.24) ₊	-.40	(ns)	-
	Normal	7.25 (1.98)	7.05 (1.59) ₊	.20	6.99 (1.68) ₊	.26	7.37 (1.60) ₊	-.12	(ns)	-
	Bajo	4.16 (.763)	4.48 (1.47) ₊	-.32	4.60 (1.37) ₊	-.44	4.79 (1.49) ₊	-.63	(ns)	-
Narrativo	Excelente	8.91 (.144)	7.68 (1.68)	1.23	8.07 (1.64) ₊	.84	8.65 (1.46) ₊	.26	***	1<3; 2<3
	Normal	7.91 (.877)	5.78 (1.50)	2.13	6.11 (1.64)	1.8	6.40 (1.57)	1.51	***	1<3
	Bajo	3.58 (1.23)	3.20 (1.11) ₊	.38	3.10 (1.04) ₊	.48	3.25 (1.15) ₊	.33	(ns)	-
Precisión			7/12 (58.3%)		8/12 (66.6%)		9/12 (75%)			

Nota. +, La puntuación de las IAs está por debajo a 1 punto de desviación de la media de la evaluación proporcionada por los docentes en formación; Dif M, La diferencia entre la media aritmética de la IA y la media aritmética del grupo. Un valor positivo más alto indica una mayor subestimación por parte del grupo en comparación con la evaluación de la IA, mientras que un valor negativo más alto indica una mayor sobreestimación por parte del grupo en comparación con la evaluación de la IA.; * p < .05; ** p < .01; *** p < .001; (ns), no significativo. Post-Hoc realizado a través de la prueba de Tukey.

4. DISCUSIÓN Y CONCLUSIONES

La evaluación es una de las principales tareas que cualquier docente realiza en sus funciones profesionales (Ley Orgánica 3/2020). Esta tarea, a pesar de ser compleja y consumir mucho tiempo debido a las ratios en el aula (por ejemplo, Ramesh y Kumar, 2022), puede aportar mejoras significativas en los procesos de aprendizaje de los estudiantes (Mellati y Khademi, 2018; Xu y Brown, 2016). Ante este dilema y con la introducción de nuevas tecnologías en las aulas, una alternativa posible para abordar este problema es emplear herramientas basadas en IA que permitan una replicación válida y confiable de las evaluaciones realizadas por los docentes.

Tomando esta idea como punto de partida, el objetivo principal de este estudio ha sido determinar si las evaluaciones proporcionadas por diferentes IAs generativas son fieles a las dadas por los docentes en formación, así como entender si existe alguna variable, como el género, el nivel de formación o el rendimiento académico, que pueda influir en la fidelidad de una evaluación en comparación con la proporcionada por una IA generativa.

Los resultados mostraron que las diferentes IAs analizadas son capaces de replicar bastante bien los patrones de los docentes en formación al evaluar tareas escritas, siendo ChatGPT la IA que obtuvo la mayor precisión (cerca del 70% de acuerdo con la evaluación de los docentes en formación) y Copilot de Bing la que presentó la menor precisión (50% de acuerdo con la evaluación de los docentes en formación). Estos resultados son consistentes con la limitada literatura previa sobre este tema, que generalmente revela una concordancia entre la retroalimentación proporcionada por la IA y la retroalimentación dada por los docentes en servicio, encontrando pequeñas diferencias entre ambos grupos (Grivokostopoulou et al., 2017; Houtao et al., 2022).

Además, los resultados han mostrado que este grado de acuerdo, en términos generales, es idéntico a lo que indica la literatura limitada sobre este tema, independientemente del género y del nivel de formación del docente en formación. Como excepción, se encontraron diferencias significativas claras solo en las evaluaciones dadas según el rendimiento académico de los docentes en formación, donde aquellos con un rendimiento académico más alto proporcionaron evaluaciones más alineadas con las ofrecidas por la IA en comparación con aquellos con un rendimiento académico más bajo. Estos resultados también son parcialmente consistentes con la limitada literatura sobre este tema. Específicamente, de acuerdo con el presente estudio, Salama y Subahi (2020) también observaron que el género y el nivel de formación eran variables que tenían poca influencia en el conocimiento y habilidades de evaluación, mientras que trabajos como los de Deneen y Brown (2016) mostraron cómo el rendimiento académico de los docentes en formación influía significativamente en la profundidad de la evaluación realizada. Sin embargo, los hallazgos del presente estudio contradicen los de Salama y Subahi (2020), quienes observaron que el rendimiento académico no era una variable influyente en la evaluación de conocimientos y habilidades.

4.1. Implicación teóricas y prácticas

Estos resultados tienen importantes implicaciones teóricas y prácticas que deben ser discutidas. En primer lugar, los hallazgos de este estudio pueden ser valiosos para la comunidad científica,

ya que podrían contribuir a ampliar la comprensión actual sobre el grado de concordancia entre las evaluaciones proporcionadas por educadores y las realizadas por sistemas basados en IA.

En segundo lugar, estos resultados pueden ser relevantes para los docentes universitarios en áreas relacionadas con la educación, ya que subrayan el interés potencial en capacitar a los futuros docentes en tecnologías digitales, como el uso de inteligencia artificial, big data o analíticas de aprendizaje, para optimizar los recursos temporales y proporcionar asistencia y monitoreo personalizados a sus estudiantes.

En Europa, se han realizado esfuerzos significativos para desarrollar un marco para la conceptualización y desarrollo de la competencia digital de los docentes, siendo el modelo DigCompEdu el principal referente (Redecker, 2017). Este marco destaca la relevancia de emplear tecnologías digitales en la formación docente para llevar a cabo tareas evaluativas dentro de la cuarta competencia "Evaluación". Esta competencia está orientada a mejorar las estrategias de evaluación, analizar la evidencia de aprendizaje y proporcionar retroalimentación mediante tecnologías digitales. En este sentido, el uso de tecnologías basadas en IA podría contribuir parcialmente a abordar esta competencia. Asimismo, estudios previos muestran cómo la mejora en la competencia digital docente podría tener efectos importantes en la gestión del tiempo y la autoeficacia docente, variables clave para reducir el estrés y la incomodidad generados por las tareas profesionales (Galindo-Domínguez y Bezanilla, 2021).

Finalmente, estos resultados pueden ser útiles para los docentes en ejercicio, ya que destacan el creciente potencial de los sistemas basados en inteligencia artificial como herramientas para evaluar el trabajo escrito de los estudiantes. Si bien estudios previos han demostrado que los docentes de educación primaria y secundaria actualmente utilizan herramientas basadas en IA principalmente para fines de creación de contenido, como textos o imágenes (Galindo-Domínguez et al., 2024), podría ser interesante añadir formación específica sobre cómo emplear la IA para la evaluación y el monitoreo de estudiantes como parte del proceso de desarrollo profesional continuo en las instituciones educativas. Como mencionan diferentes autores (por ejemplo, Kasneci et al., 2023; Owan et al., 2023), el uso de la IA en el ámbito educativo requiere que tanto docentes como estudiantes desarrollen un conjunto de competencias necesarias para comprender la tecnología, aprovechar sus potencialidades y reconocer sus limitaciones.

4.2. Limitaciones y prospectiva

El presente estudio tiene varias limitaciones que deben tenerse en cuenta al interpretar los resultados. La primera limitación se refiere a la muestra. Aunque esta es relativamente grande, los resultados podrían variar si los mismos participantes se evaluaran en unos años cuando estén en activo. Por ello, futuros estudios podrían replicar la metodología utilizada en este estudio, empleando a docentes en servicio de diferentes etapas educativas como participantes, para observar si los resultados obtenidos son similares o no. Además, podría ser interesante comparar las evaluaciones de docentes novatos, educadores de media carrera y docentes veteranos, ya que esto podría revelar importantes perspectivas sobre cómo la experiencia docente puede influir en los procedimientos de evaluación. Esta idea se justifica en que, mientras algunos estudios sugieren que la cantidad de experiencia docente podría influir en la alfabetización evaluativa (por ejemplo, Spear-Swerling et al., 2005), otros apuntan en la dirección opuesta (por ejemplo, Bagsao y Peckley, 2020; Salama y Subahi, 2020).

Asimismo, el presente estudio solo consideró la evaluación de trabajos escritos, lo que imposibilita determinar si el nivel de acuerdo entre las evaluaciones proporcionadas por la IA y los docentes en formación también podría ocurrir en trabajos presentados en otros formatos, como audio, video, imágenes o ecuaciones matemáticas. Aunque es más complejo para el estado actual de los sistemas de IA, futuros estudios podrían intentar replicar la metodología utilizada, pero evaluando tareas en formatos diferentes al escrito.

Finalmente, los textos escritos fueron generados por la IA bajo la instrucción de asumir el papel de un estudiante de 10 años. Sin embargo, las diferencias en los resultados obtenidos podrían surgir si se utilizaran textos reales escritos por estudiantes de 10 años. Por esta razón, futuros estudios podrían replicar la metodología del presente trabajo, pero utilizando textos redactados por estudiantes reales.

5. REFERENCIAS

- Atjonen, P. (2017). Development of teacher assessment literacy in comprehensive schools – Views from the curriculum analysis. *Kriteerit Puntarissa*, 74, 132–169.
- Atjonen, P., Pöntinen, S., Kontkanen, S., & Ruotsalainen, P. (2022). In Enhancing Preservice Teachers' Assessment Literacy: Focus on Knowledge Base, Conceptions of Assessment, and Teacher Learning. *Frontiers in Education*, 7, 1-12. <https://doi.org/10.3389/feduc.2022.891391>
- Bagsao, J., & Peckley, M.K. (2020). Assessment Literacy of Public Elementary School Teachers in the Indigenous Communities in Northern Philippines. *Universal Journal of Educational Research*, 8(11b), 5693-5703. <http://dx.doi.org/10.13189/ujer.2020.082203>
- Choi, Y., & McClenen, C. (2020). Development of adaptive formative assessment system using computerized adaptive testing and dynamic bayesian networks. *Applied Sciences*, 10(22), 8196. <https://www.mdpi.com/2076-3417/10/22/8196#>
- Contreras, J.O., Hilles, S.M., & Abubakar, Z.B. (2018) Automated essay scoring with ontology based on text mining and NLTK tools. In I. Zen (Pres.), *2018 International Conference on Smart Computing and Electronic Enterprise* (pp. 1-6). IEEEExplore.
- Coppock, A., Leeper, T.J., Mullinix, K.J. (2018). Generalizability of heterogeneous treatment effect estimates across samples. *PNAS*, 115(49), 12441-12446. <http://www.pnas.org/cgi/doi/10.1073/pnas.1808083115>
- Cummins, R., Zhang, M., & Briscoe, E. (2016). *Constrained multi-task learning for automated essay scoring*. Association for Computational Linguistics.
- Darwish, S.M., & Mohamed, S.K. (2019) Automated essay evaluation based on fusion of fuzzy ontology and latent semantic analysis. In A.E. Hassanien, A.T. Azar, T. Gaber, R. Bhatnagar, & M.F. Tolba (Eds.), *The International Conference on Advanced Machine Learning Technologies and Applications* (pp. 566-575). Springer.
- DeLuca, D., Willis, J., Cowie, B., Harrison, C., Coombs, A., Gibson, A., et al. (2019). Policies, programs, and practices: exploring the complex dynamics of assessment education in teacher education across four countries. *Frontiers in Education*, 4, 1-19. <https://doi.org/10.3389/feduc.2019.00132>

- Deneen, C.C., & Brown, G.T.L (2016). The impact of conceptions of assessment on assessment literacy in a teacher education program. *Cogent Education*, 3(1), 1225380. <https://doi.org/10.1080/2331186X.2016.1225380>
- Dillenbourg, P. (2016). The evolution of research on digital education. *International Journal of Artificial Intelligence in Education*, 26(2), 544-560. <https://doi.org/10.1007/s40593-016-0106-z>
- Dong, F., Zhang, Y., Yang, J. (2017). Attention-based recurrent convolutional neural network for automatic essay scoring. In R. Levy & L. Specia (Eds.), *Proceedings of the 21st Conference on Computational Natural Language Learning* (pp. 153–162). Association for Computational Linguistics.
- Eklöf, H. (2010). Skill and will: test-taking motivation and assessment quality. *Assessment in Education: Principles, Policy & Practice*, 17(4), 345-356. <https://doi.org/10.1080/0969594X.2010.516569>
- Galindo-Domínguez, H., & Bezanilla, M.J. (2021). Promoting Time Management and Self-Efficacy Through Digital Competence in University Students: A Mediation Model. *Contemporary Educational Technology*, 13(2), ep294. <https://doi.org/10.30935/cedtech/9607>
- Galindo-Domínguez, H., Delgado, N., Losada, D., & Etxabe, J.M. (2024). An analysis of the use of artificial intelligence in education in Spain: The in-service teacher's perspective. *Journal of Digital Learning in Teacher Education*, 40(1), 41-56. <https://doi.org/10.1080/21532974.2023.2284726>
- González-Calatayud, V., Prendes-Espinosa, P., & Roig-Vila, R. (2021). Artificial Intelligence for student assessment: a systematic review. *Applied Sciences*, 11, 5467. <https://doi.org/10.3390/app11125467>
- Government of Newfoundland and Labrador (2014). *English Language Arts Grade 6. Appendix D: Sample Elementary Classroom Rubrics and Checklists*. Department of Education of the Government of Newfoundland and Labrador. https://www.gov.nl.ca/education/files/k12_curriculum_guides_english_grade6_300614_g6_ela.pdf
- Grivokostopoulou, F., Perikos, I., Hatzilygeroudis, I. (2017). An Educational System for Learning Search Algorithms and Automatically Assessing Student Performance. *International Journal of Artificial Intelligence in Education*, 27, 207–240. <http://dx.doi.org/10.1007/s40593-016-0116-x>
- Hill, M., Ell, F., & Evers, G. (2017). Assessment capability and student self-regulation: the challenge of preparing teachers. *Frontiers in Education*, 2, 1-15. <https://doi.org/10.3389/feduc.2017.00021>
- Houtao, L., Wenjia, M., Tingting, W., & Chuanhua, X. (2022). The Study of Feedback in Writing from College English Teachers and Artificial Intelligence Platform Based on Mixed Method Teaching. *Pacific International Journal*, 5(4), 147-154. <https://doi.org/10.55014/pij.v5i4.270>
- Hrastinski, S., Olofsson, A. D., Arkenback, C., Ekström, S., Ericsson, E., Fransson, G., Jaldemark, J., Ryberg, T., Öberg, L.-M., Fuentes, A., Gustafsson, U., Humble, N., Mozelius, P., Sundgren, M., & Utterberg, M. (2019). Critical imaginaries and reflections on artificial intelligence and robots in post-digital K-12 education. *Post-Digital Science and Education*, 1(2), 427-445. <https://doi.org/10.1007/s42438-019-00046-x>
- Jani, K.H., Jones, K.A., Jones, G.W., Amiel, J., Barron, B., & Elhadad, N. (2020). Machine learning to extract communication and historytaking skills in OSCE transcripts. *Medical Education*, 54, 1159–1170. <https://doi.org/10.1111/medu.14347>

- Kasneji, E., Sessler, K., Küchemann, S., ..., Kasneji, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Ke, Z., Inamdar, H., Lin, H., & Ng, V. (2019). Give me more feedback II: Annotating thesis strength and related attributes in student essays. In A. Korhonen, D. Traum & L. Márquez (Eds.), *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (pp. 3994-4004). Association for Computational Linguistics.
- Kumar, Y., Aggarwal, S., Mahata, D., Shah, R. R., Kumaraguru, P., & Zimmermann, R. (2019). Get it scored using autosas—an automated system for scoring short answers. In B. Williams, Y. Chen, & J. Neville (Eds.), *Proceedings of the AAAI Conference on Artificial Intelligence* (pp. 9662–9669). AAAI Press.
- Liu, M., Wang, Y., Xu, W., & Liu, L. (2017). Automated Scoring of Chinese Engineering Students' English Essays. *International Journal of Distance Education Technologies*, 15(1), 52–68.
- Lovorn, M.G., Reza, A. (2011). Assessing the Assessment: Rubrics Training for Pre-service and New In-service Teachers. *Practical Assessment, Research, and Evaluation*, 16(1), 16. <https://doi.org/10.7275/sjt6-5k13>
- Mathias, S., & Bhattacharyya, P. (2018). Thank “Goodness”! A Way to Measure Style in Student Essays. In Y. Tseng, H. Chen, V. Ng. & M. Komachi (Eds.), *Proceedings of the 5th Workshop on Natural Language Processing Techniques for Educational Applications* (pp. 35–41). Association for Computational Linguistics.
- Mellati, M., & Khademi, M. (2018). Exploring teachers' assessment literacy: Impact on learners' writing achievements and implications for teacher development. *Australian Journal of Teacher Education*, 43(6), 1-18. <http://dx.doi.org/10.14221/ajte.2018v43n6.1>
- Mirchi, N., Bissonnette, V., Yilmaz, R., Ledwos, N., Winkler-Schwartz, A., & Del Maestro, R.F. (2020). The Virtual Operative Assistant: An explainable artificial intelligence tool for simulation-based training in surgery and medicine. *PLoS ONE* 15, e0229596. <https://doi.org/10.1371/journal.pone.0229596>
- Ocaña-Fernández, Y., Valenzuela-Fernández, L.A., & Garro-Aburto, L.L. (2019). Inteligencia artificial y sus implicaciones en la educación superior. *Propósitos y Representaciones*, 7(2), 536-568. <https://doi.org/10.20511/pyr2019.v7n2.274>
- Organic Law 3/2020, of December 29th, amending Organic Law 2/2006, of May 3rd, on Education. *Official State Gazette*, 340, 122868-122953. <https://www.boe.es/eli/es/lo/2020/12/29/3>
- Ouguengay, Y.A., El Faddouli, N.-E., & Bennani, S. (2015). A neuro-fuzzy inference system for the evaluation of reading/writing competencies acquisition in an e-learning environment. *Journal of Theoretical and Applied Information Technology*, 81(3), 600–608.
- Owan, V.J., Bekom, K., Emoji, D., Onor, E., & Asuquo, B. (2023). Exploring the potential of artificial intelligence tools in educational measurement and assessment. *Modestum. Eurasia Journal of Mathematics, Science and Technology Education*, 19(8), em2307. <https://doi.org/10.29333/eimste/13428>
- Ramesh, D., & Kumar, S. (2022). An automated essay scoring systems: a systematic literature review. *Artificial Intelligence Review*, 55, 2495-2527. <https://doi.org/10.1007/s10462-021-10068-2>

- Redecker, C. (2017). *European Framework for the Digital Competence of Educators: DigCompEdu*. Joint Research Centre. <http://dx.doi.org/10.2760/159770>
- Rhienmora, P., Haddawy, P., Suebnukarn, S., Dailey, M.N. (2011). Intelligent dental training simulator with objective skill assessment and feedback. *Artificial Intelligence in Medicine*, 52(2), 115–121. <https://doi.org/10.1016/j.artmed.2011.04.003>
- Salama, S., & Subahi, A. M. (2020). The Impact of Specialty, Sex, Qualification, and Experience on Teachers' Assessment Literacy at Saudi Higher Education. *International Journal of Learning, Teaching and Educational Research*, 19(5), 200-216. <https://doi.org/10.26803/ijlter.19.5.12>
- Samarakou, M., Fylladitakis, E.D., Karolidis, D., Früh, W.-G., Hatzia Apostolou, A., Athinaios, S.S., & Grigoriadou, M. (2016). Evaluation of an intelligent open learning system for engineering education. *Knowledge Management & E-Learning: An International Journal*, 8(3), 496–513.
- Spear-Swerling, L., Owen, P., & Alfano, M.P. (2005). Teachers' literacy-related knowledge and self-perceptions in relation to preparation and experience. *Annals of Dyslexia*, 55, 266-296. <https://doi.org/10.1007/s11881-005-0014-7>
- Stiggins, R. (2014). Improve assessment literacy outside of schools too. *Phi Delta Kappan*, 96, 65–72.
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53-55. <https://doi.org/10.5116/ijme.4dfb.8dfd>
- Ulum, Ö.G. (2020). A critical deconstruction of computer-based test application in Turkish State University. *Education and Information Technologies*, 25, 4883–4896. <https://doi.org/10.1007/s10639-020-10199-z>
- Vij, S., Tayal, D., & Jain, A. (2020). A machine learning approach for automated evaluation of short answers using text similarity based on WordNet graphs. *Wireless Personal Communications*, 111(2), 1271–1282. <https://doi.org/10.1007/s11277-019-06913-x>
- Wang, P. (2019). On Defining Artificial Intelligence. *Journal of Artificial General Intelligence*, 10(2), 1-37. <https://doi.org/10.2478/jagi-2019-0002>
- Xu, Y., & Brown, G.T.L. (2016). Teacher assessment literacy in practice: a reconceptualization. *Teaching and Teacher Education*, 58, 149-162. <http://dx.doi.org/10.1016/j.tate.2016.05.010>

Para citar este artículo:

Galindo-Domínguez, H., Delgado, N., Sainz de la Maza, M., y Expósito, E. Un análisis experimental de la relación entre las evaluaciones proporcionadas por la inteligencia artificial y las proporcionadas por los docentes en formación [An experimental analysis of the relationship between the evaluations of artificial intelligence and pre-service teachers]. *EduTEC, Revista Electrónica de Tecnología Educativa*, (89), 84-104. <https://doi.org/10.21556/edutec.2024.89.3509>

6. ANEXOS

6.1. ANEXO 1: Textos generados

Tipo	Nivel	Texto
Texto descriptivo	Alto	En el tranquilo bosque detrás de mi casa, el sol de la tarde se filtra a través de las hojas de los altos árboles, creando un mágico juego de luces y sombras en el suelo. Los pájaros cantan melodías alegres mientras exploran entre las ramas, y el aire está lleno del aroma fresco de la naturaleza. A medida que camino por el sendero cubierto de hojas crujientes, el murmullo del arroyo cercano me guía hacia su orilla, donde el agua fluye suavemente sobre las piedras redondas y pulidas. Cada paso que doy me sumerge aún más en este misterioso mundo, donde la paz y la belleza se entrelazan en perfecta armonía."
	Medio	Me gusta mucho el parque cerca de mi casa. Cuando llego allí, siento una emoción especial. El parque es un lugar divertido y emocionante para mí. Hay tantas cosas que hacer y explorar. Los columpios son mi parte favorita. Me balanceo alto y siento como si volara. También hay un tobogán grande y emocionante en el que deslizo rápidamente. Mis amigos y yo nos divertimos mucho compitiendo a ver quién llega primero. Además de los juegos, el parque está yeno de árboles altos que proporcionan sombra en los días calurosos. Me encanta sentarme bajo un árbol con mis amigos y hablar de nuestras aventuras. También hay un pequeño estanque en el que a veces vemos patitos nadando. Cuando tenemos hambre, a veces traemos helado y nos sentamos en los bancos del parque para disfrutarlo. Es tan refrescante y delicioso.
	Bajo	Aier fui a un lugar bonito. Estaba afuera y hacia mucha calor. Me gustó mucho por qué era un día ermoso. Había columpios, un tovogán grande y un par de vancos para sentarse. Jugamos en los columpios, subimos al tovogán y reimos mucho. Yo y mis amigos pasamos un buen rato. Después de jugar, fuimos a una tienda cercana y comimos elado. Yo pedí un elado de fresa por qué es el más rico. La parte más dibertida fue ver cómo se deretía bajo el sol caliente. En general, el día estuvo guay, aunque no recuerdo el nombre esacto del lugar. Haber si volvemos otro dia.
Texto argumentativo	Alto	Creo que debemos cuidar a los animales. Los animales son nuestros amigos y necesitan amor y protección. Si cuidamos de ellos, podemos ayudar a que estén felices y saludables. También es importante no tirar basura en el suelo porque puede lastimar a los animales y contaminar su hogar. Debemos mantener limpio nuestro entorno para que los animales vivan en un lugar seguro. En resumen, debemos ser buenos amigos de los animales y cuidar de nuestro planeta para que todos podamos vivir en armonía.
	Medio	Creo que los videojuegos son divertidos. Muchos niños juegan a los videojuegos y se divierten mucho. A veces, es mejor jugar afuera, pero los videojuegos son buenos para cuando hace frío o llueve. También es importante hacer la tarea antes de jugar a los videojuegos. La tarea es aburrida, pero es importante hacerla para aprender cosas. En resumen, los videojuegos son divertidos, pero la tarea es importante. Los dos pueden ser buenos si los hacemos en el momento adecuado.
	Bajo	Los bideojuegos son guais. Me gusta jugar porque es dibertido. Tengo muchos juegos en mi consola. Algunos son de carreras y otros de abenturas. Me emociona mucho cuando gano. Mis amigos tanbién juegan a bideojuegos. Jugamos juntos y nos dibertimos mucho. Abeces, nos retamos haber quién gana más partidas. Son guais, pero tanvién se que es inportante acer otras cosas, como leer o acer los deveres. Asique, aunque me gusten mucho los bideojuegos, sé que no puedo jugar todo el tiempo.
Texto instructivo	Alto	Cómo Hacer una Cometa Materiales: Papel de seda o papel de envolver Palos de madera Tijeras Hilo fuerte Pegamento

	<p>Cinta adhesiva Pinturas y pinceles (opcional) Instrucciones:</p> <p>Primero, corta cuatro palos de madera de igual longitud para hacer el armazón de la cometa. Dos de ellos serán los lados verticales y los otros dos, los lados horizontales.</p> <p>Une los palos para formar un marco en forma de cruz. Utiliza cinta adhesiva o hilo fuerte para asegurar las uniones.</p> <p>A continuación, corta el papel de seda o papel de envolver en forma de diamante, dejando unos centímetros extra en los bordes para doblarlos y pegarlos al armazón.</p> <p>Pega con cuidado el papel al marco de madera, asegurándote de que quede estirado y sin arrugas.</p> <p>Ahora, corta un trozo de hilo fuerte de aproximadamente un metro de largo. Amarra un extremo al centro del palo horizontal y deja el otro extremo libre.</p> <p>Decorar la cometa con pinturas y pinceles si lo deseas. Puedes hacer diseños creativos en el papel.</p> <p>Cuando hayas terminado de decorar la cometa y el pegamento esté seco, ¡está lista para volar!</p> <p>Encuentra un día con viento suave, ve a un campo abierto y suelta tu cometa al viento sujetando el extremo del hilo. ¡Disfruta viendo cómo se eleva y vuela alto en el cielo!"</p>
	<p>Cómo Hacer un Sándwich de PByJ</p> <p>Materiales:</p> <p>Pan de molde Mantequilla de maní Mermelada de fresa Cuchillo</p> <p>Medio</p> <p>Instrucciones:</p> <p>Coge dos rebanadas de pan de molde.</p> <p>Abre la mantequilla de maní y la mermelada de fresa.</p> <p>Unta mantequilla de maní en una de las rebanadas de pan.</p> <p>Luego, pon mermelada de fresa en la otra rebanada.</p> <p>Presiona las dos rebanadas juntas para que la mantequilla de maní y la mermelada se mezclen.</p>
	<p>Como Hacer Un Pastel:</p> <p>Bajo</p> <p>Conprar una mescla para pastel.</p> <p>Agregar uevos y leche.</p> <p>Mezclarlo todo.</p> <p>Poner en un molde.</p> <p>Meter el molde en el orno.</p> <p>Sacarlo cuando este listo.</p>
Texto Narrativo	<p>Alto</p> <p>Ayer, junto a mis amigos, pasé un emocionante día en el parque. Juntos, construimos un inmenso castillo de arena y nos sumergimos en un emocionante juego de escondidas. Posteriormente, comimos deliciosos helados mientras admirábamos el colorido arco iris que se formó en el cielo. Sin duda, fue uno de los días más sorprendentes que he vivido.</p>
	<p>Medio</p> <p>Un día soleado, fui al parque con mis amigos. Corrimos y jugamos en los columpios. Después, decidimos explorar el bosque cercano. Seguido, encontramos un arrollo y lanzamos piedras al agua. Después, nos sentamos bajo un árbol a comer sándwiches. Para acabar el día, regresamos a casa, cansados pero felices.</p>
	<p>Bajo</p> <p>Un día, fui al parke. Jugamos mucho y comimos elado. Luego, fuimos a casa. Fin.</p>

6.2. Anexo II: Rúbrica empleada por los docentes en formación para evaluar los diferentes textos.

Criterio	Excelente (4)	Bueno (3)	Regular (2)	Pobre (1)
Contenido	El texto presenta información detallada y rica del tema, con una variedad de detalles y ejemplos relevantes que enriquecen la comprensión del lector.	El texto presenta una cantidad de información adecuada del tema, con detalles y ejemplos que hacen que el tema sea claro para el lector.	El texto presenta una cantidad de información limitada del tema, con detalles insuficientes o poco claros que dificultan la comprensión del lector.	El texto tiene una cantidad de información tan pobre que no se puede entender el tema.
Organización	El texto tiene una estructura clara y lógica, con una introducción, desarrollo y conclusión bien definidos. Las ideas están organizadas de manera efectiva.	El texto tiene una estructura generalmente clara y lógica, aunque la organización podría mejorar en algunos lugares. Las ideas están organizadas de manera adecuada.	El texto tiene una estructura poco clara o desorganizada, lo que dificulta la comprensión de las ideas.	El texto carece de estructura y organización, lo que hace que sea difícil seguir las ideas presentadas.
Vocabulario y Lenguaje	El texto utiliza un vocabulario variado y preciso. Las oraciones son complejas y están bien construidas.	El texto utiliza un vocabulario adecuado, aunque podría incorporar más variedad y precisión. Las oraciones son en su mayoría correctas.	El texto utiliza un vocabulario limitado y repetitivo, lo que afecta la calidad de la producción. Las oraciones son simples y pueden contener errores gramaticales.	El texto utiliza un vocabulario muy limitado y/o inadecuado, lo que dificulta la comprensión. Las oraciones son incorrectas y confusas.
Coherencia y Cohesión	El texto muestra una alta coherencia y cohesión, con conexiones claras entre las ideas y un uso efectivo de conectores y referencias.	El texto muestra coherencia y cohesión en general, aunque algunas conexiones entre las ideas pueden ser más claras. Se utilizan algunos conectores y referencias.	El texto carece de coherencia y cohesión, lo que dificulta la transición entre las ideas. Los conectores y referencias son escasos o inapropiados.	El texto es incoherente y carece de cualquier forma de cohesión, lo que hace que sea difícil seguir la narración.



AI Tutors vs. Human Instructors: Perceptions of Higher Education Students in Hungary and Spain

Tutores de IA frente a instructores humanos: Percepciones de los estudiantes de educación superior en Hungría y España

 Ahmad Hajeer; hajeer.ahmad@uni-bge.hu; Budapest Business School (Hungary)

 Árpád Papp-Váry; papp-vary.arpad@uni-bge.hu; Budapest Business University (Hungary)

 Éva Pólya; polya.eva@uni-bge.hu; Budapest Business University (Hungary)

Abstract

Integrating AI-powered tutoring systems in higher education represents a significant advancement in educational technology, offering personalized and adaptive learning experiences. This study investigates the perceptions and expectations of higher education students in Hungary and Spain regarding AI tutors. Despite extensive research on the technological efficacy of AI systems, there is limited understanding of student attitudes in these specific cultural contexts. This research aims to fill this gap by exploring student expectations, satisfaction levels, and perceived benefits of AI tutors compared to human instructors. To achieve this, a validated questionnaire was administered to 184 higher education students from Hungary and Spain, capturing data on various dimensions of their expectations. The study's findings indicate that students appreciate the adaptability and continuous guidance provided by AI tutors, with Hungarian students showing higher overall expectations compared to their Spanish counterparts. These insights suggest that AI tutoring systems can enhance the learning experience by addressing individual student needs more effectively. The implications of this study are significant for higher education institutions seeking to integrate AI technologies.

Keywords: Higher education, AI tutoring systems, adaptive learning, educational technology, Student perceptions

Resumen

La integración de sistemas de tutoría impulsados por IA en la educación superior representa un avance significativo en la tecnología educativa, ofreciendo experiencias de aprendizaje personalizadas y adaptativas. Este estudio investiga las percepciones y expectativas de los estudiantes de educación superior en Hungría y España respecto a los tutores de IA. A pesar de la extensa investigación sobre la eficacia tecnológica de los sistemas de IA, existe un entendimiento limitado sobre las actitudes de los estudiantes en estos contextos culturales específicos. Esta investigación pretende llenar este vacío explorando las expectativas de los estudiantes, sus niveles de satisfacción y los beneficios percibidos de los tutores de IA en comparación con los instructores humanos. Para lograr esto, se administró un cuestionario validado a 184 estudiantes de educación superior de Hungría y España, capturando datos sobre diversas dimensiones de sus expectativas. Los hallazgos del estudio indican que los estudiantes valoran la adaptabilidad y la orientación continua proporcionada por los tutores de IA, con los estudiantes húngaros mostrando expectativas más altas en comparación con sus homólogos españoles. Estos conocimientos sugieren que los sistemas de tutoría de IA pueden mejorar la experiencia de aprendizaje al abordar las necesidades individuales de los estudiantes de manera más efectiva.

Palabras clave: Educación superior, sistemas de tutoría con IA, aprendizaje adaptativo, tecnología educativa, percepciones de los estudiantes.



1. INTRODUCTION

Higher education institutions are encountering significant challenges as they strive to meet the varied needs of learners in the world (Hajeer et al., 2023). Traditional pedagogical approaches often fall short in delivering personalized instruction and accommodating individual learning preferences (Hajeer, 2024). The integration of Artificial Intelligence (AI) in higher education has emerged as a transformative force impacting education. AI-powered tutoring systems, in particular, have received attention for their potential to revolutionize teaching and learning processes by providing personalized learning experiences (Dede, 2014). The importance of this topic in higher education is undeniable, as these systems promise to address key challenges such as improving student engagement, personalizing education, and optimizing learning outcomes (Kim, 2020; Basri, 2024). In the context of higher education management and marketing, AI tutoring systems offer several advantages. These systems can enhance the learning experience, making institutions more attractive to prospective students by showcasing their commitment to innovative educational practices. Moreover, AI tutoring systems can provide valuable data insights, enabling better resource allocation and more effective marketing strategies aimed at prospective students and their parents (Nguyen et al., 2024). This technological advancement is particularly relevant as higher education institutions strive to differentiate themselves in a competitive market (Chan, 2023).

Studies have shown that these systems can adapt to individual student needs, provide immediate feedback, and support skill development in various subjects (Chounta et al., 2022; Nguyen et al., 2024). However, despite the promising results, there is a notable gap in research concerning the perceptions of students in specific cultural contexts, such as Hungary and Spain. Existing studies have primarily focused on the technological efficacy of AI systems, with limited exploration of student attitudes and expectations in these regions (Basri, 2024). To address this gap, the current study aims to investigate the perceptions of higher education students in Hungary and Spain regarding AI tutoring systems. Utilizing a validated questionnaire, this research seeks to capture student expectations, regarding AI tutors compared to human instructors. The primary research question guiding this study is: "What are the perceptions and expectations of higher education students in Hungary and Spain towards AI-powered tutoring systems?"

1.1. Evolution and applications of intelligent tutoring systems in higher education

AI tutoring refers to the use of artificial intelligence technologies to simulate human tutoring and provide personalized learning experiences for students. This involves using machine learning algorithms to analyze student data, identify knowledge gaps, and tailor instruction and feedback accordingly (Mittra et al., 2021). Natural language processing allows AI tutors to interact with students in a natural and engaging way, providing explanations and answering questions similarly to human tutors. Intelligent tutoring systems (ITS), as defined by Woolf (2009), are "computer systems that provide immediate and customized instruction or feedback to learners," serving as a prominent example of AI tutoring applications by providing customized instruction and feedback, thereby simulating a human tutor's role.

The concept of AI in education dates back several decades, with early systems focusing on simple programmed instruction and rule-based tutoring. For instance, the SCHOLAR system developed in the 1970s (Carbonell, 1970) was an early attempt to use AI for tutoring in

geography. Over time, advancements in AI technologies have led to the development of more sophisticated and adaptive tutoring systems. Woolf (2009) discusses the evolution of intelligent tutors from simple feedback systems to complex, student-centered learning environments. The integration of AI in education has progressively moved from theoretical research to practical applications, marked by milestones in the development of adaptive learning platforms and AI-driven educational tools (Dede, 2014).

AI tutoring systems are employed to provide personalized learning experiences, adapting content and instructional methods to meet individual student needs. These systems are designed to identify students' strengths and weaknesses, offering tailored support to improve learning outcomes (Kim, 2020). For example, de Baker and Inventado (2014) discuss the use of AI-powered tutoring systems to address diverse learning needs and provide personalized feedback and guidance. Studies have shown that AI tutors can significantly enhance learning outcomes by providing adaptive learning experiences tailored to individual student needs (Chounta et al., 2022; Al-Shanfari et al., 2023). The use of AI in higher education also addresses the need for innovative solutions to enhance student engagement and motivation. For example, AI tutors can provide interactive and personalized learning experiences that are crucial in maintaining student interest and promoting active learning (Dede, 2014). Furthermore, AI technologies like chatbots and virtual assistants are increasingly used in student support services, aiding in administrative tasks and providing round-the-clock assistance to students, thereby enhancing the overall student experience (Chaudhry & Kazim, 2022). Despite the numerous advantages, the implementation of AI in higher education also raises important ethical considerations. Issues such as data privacy, potential biases in AI algorithms, and the need for transparency and fairness in AI-driven systems are critical areas that require careful attention. Ensuring that AI technologies are used responsibly and ethically is paramount to gaining the trust of educators and students alike (Nguyen et al., 2024).

1.1. Machine teachers: benefits, challenges, and student perceptions

Machine teachers, which include both embodied and disembodied agents, represent an innovative shift in educational technology. Embodied machine teachers necessitate a physical form, whether tangible, virtual, or a hybrid of both (Pfeifer & Scheier, 1999). Physical embodiments, such as robots made from materials like plastic, wood or metal could be used in face-to-face teaching scenarios to engage students directly (Li et al., 2015). Conversely, virtually embodied machine teachers are computer-generated and exist visually on screens (Li, 2015). Disembodied machine teachers. These include chatbots, software agents, and interface agents that communicate through text or voice. The rise in online education has amplified the necessity for machine teachers. Both disembodied and virtually embodied teachers are well-suited for online learning environments, providing continuous support without physical presence (Allen & Seaman, 2017). Despite their growing importance, there remains a paucity of research regarding student reactions to and acceptance of machine teachers (Kim, 2020).

Some studies suggest that while students may prefer human instructors, knowledge retention is notably better when instruction is provided by robots (Li et al., 2015). Conversely, concerns about the costs, the necessity for specialized teacher training, and the practical applicability of robots in diverse educational contexts persist (Edwards et al., 2016). Despite these concerns, a substantial body of literature highlights the beneficial impacts of social robots on educational

outcomes. For example, researchers have found that social robots are perceived as credible and effective conveyors of educational content (Chaudhry & Kazim, 2022). Students have responded positively to robots that provide affirmative feedback, perceiving them as both attractive and acceptable in the learning environment. Moreover, social robots have facilitated the inclusion of homebound students, enabling them to participate in real-time classroom interactions with their peers and teachers (Double Robotics, 2017).

The positive effects of social robots in education are well-documented. For instance, Edwards et al. (2016) reported that students view social robots as reliable sources of information, capable of delivering educational content effectively. Additionally, research by Park and Whang et al. (2022) emphasizes the importance of empathy and engagement in human-robot interactions, noting that robots capable of exhibiting empathic behaviors and social cues can significantly enhance the learning experience. Li (2015) also discovered that physically present robots tend to be more persuasive and receive more attention compared to their virtual counterparts. Moreover, studies have shown varying impacts of robotic technologies on knowledge recall. For instance, Li et al. (2015) found that videos featuring human instructors and animated robots have similar effects on student knowledge recall, whereas videos of real robots resulted in weaker recall performances. These findings underscore the complex dynamics of robot-assisted learning and the necessity of tailoring robot design and deployment to specific educational contexts.

1.2. How are students' expectations affected?

The integration of AI-powered tutoring systems in higher education has generated significant interest and varying expectations among students. Several factors influence these expectations, including technological familiarity, perceived usefulness, and cultural context. Understanding these factors is essential for effectively implementing AI tutors in educational settings. One of the primary factors influencing student expectations of AI tutors is their familiarity with technology. Students who are more accustomed to using advanced technologies in their daily lives tend to have higher expectations of AI tutors. They anticipate that these systems will provide seamless, efficient, and personalized learning experiences (Dede, 2014). Additionally, the perceived usefulness of AI tutors plays a critical role. Students expect AI tutors to offer immediate feedback, personalized learning paths, and adaptive content that caters to their individual learning needs. This expectation is rooted in the potential of AI to address diverse learning styles and pace, thereby enhancing the overall learning experience (Chounta et al., 2022).

Empirical findings suggest that students generally perceive AI tutors positively, particularly in terms of their ability to provide personalized and adaptive learning experiences. Studies have shown that AI tutors can significantly improve student engagement and motivation by offering tailored support and feedback (Kim, 2020). For instance, Chounta et al. (2022) found that students appreciate the immediate feedback provided by AI tutors, which helps them identify and correct mistakes in real-time. This instant response mechanism is seen as a key advantage over traditional learning methods, where feedback may be delayed.

Moreover, AI tutors' ability to adapt to individual student needs is highly valued. Students perceive AI tutors as effective tools for personalized learning, as these systems can analyze student performance data and adjust the learning content accordingly (Al-Shanfari et al., 2023).

This adaptability not only helps in addressing specific knowledge gaps but also promotes a deeper understanding of the subject matter. Nguyen et al. (2024) highlight that such tailored learning experiences can lead to improved academic performance and higher satisfaction levels among students.

When comparing student perceptions of AI tutors with human tutors, several distinct differences and similarities emerge. One notable difference is the level of personalized attention. While human tutors provide personalized interaction, they are limited by time and availability. AI tutors, on the other hand, can offer continuous, individualized support without such constraints. This continuous availability is particularly appreciated by students who seek help outside of regular class hours (Nguyen et al., 2024). However, despite the benefits of AI tutors, students often express concerns about the lack of human empathy and emotional understanding in AI interactions. Human tutors are perceived as more capable of providing emotional support and understanding the nuances of student emotions, which AI systems currently struggle to replicate (Al-Shanfari et al., 2023). This human element is crucial for creating a supportive and motivating learning environment, which is often cited as a limitation of AI tutors. Furthermore, trust and reliability are critical factors in student perceptions. While many students trust AI tutors to deliver accurate and consistent information, there is still a degree of skepticism about the technology's reliability and potential biases (Nguyen et al., 2024). Ensuring transparency in AI algorithms and addressing ethical concerns are essential steps to build and maintain student trust in AI-powered educational tools.

Student Perceptions of AI tutors are influenced by factors such as technological familiarity, perceived usefulness, and the ability to deliver personalized learning. While students appreciate the adaptability and availability of AI tutors, concerns persist about the absence of human empathy. AI tutors excel in personalized learning but cannot replicate the emotional support human instructors provide. Overcoming these limitations with ethical and transparent AI implementation is essential for their successful integration in higher education.

2. METHODS

This study used a quantitative research design with a cross-sectional survey to examine higher education students' views on AI tutoring systems in Hungary and Spain. A questionnaire was designed and used to gather data on students' expectations, satisfaction, and perceived benefits of AI tutors compared to human instructors. Statistical tests were conducted to ensure that the questionnaire was both reliable and valid. First, Exploratory Factor Analysis (EFA) was conducted to identify the main factors and understand the structure of the questionnaire. Then, Confirmatory Factor Analysis (CFA) was used to confirm this structure and assess how well the model fits the data. To check the reliability of the questionnaire, Cronbach's Alpha and Omega coefficients were calculated.

2.1. Participants and data collection

This study aimed to explore higher education students' perceptions of AI tutoring systems in Hungary and Spain, two countries with distinct educational and cultural contexts. These countries were chosen because they offer a comparative look at how AI is perceived in educational settings where the adoption of new technologies is increasingly relevant. The

participants were selected from institutions in both countries to represent diverse academic backgrounds in an attempt to ensure a comprehensive view of student attitudes.

The sample included 184 higher education students, with age distribution ranging from 18 to 25 years old. The largest age groups were 21 (35.33%) and 22 years old (32.61%), with smaller proportions aged 20 (10.87%), 23 (11.41%), and a minimal number aged 18, 19, 24, and 25. In terms of gender, the sample was nearly evenly split, with 52.17% male and 47.83% female participants. Students were drawn from different academic years, with first-year students making up 9.78% of the sample, second-year students 14.13%, third-year students 30.43%, and fourth-year students the largest group at 45.65%. Most participants were from Spain (72.28%), followed by Hungary (19.02%), with 8.70% international students from regions outside Europe, including Latin America and the Middle East. Regarding academic specialization, 29.35% of the students were from social sciences and humanities, 47.83% were from business and economics, and 22.83% were studying STEM subjects (Science, Technology, Engineering, Mathematics, and Medicine). The diverse demographics and academic representation of the sample is to ensure that various perspectives were captured in the study, making the findings possibly applicable to a broader student population.

The data were collected via an online survey created using the Google Forms platform, which was distributed through a learning management platform at the participating academic institutions. This platform served as the primary method for reaching students in Hungary and Spain. After the initial distribution via the learning platform, students were encouraged to share the survey link with their peers to increase participation. Informed consent was obtained from all participants, who were assured of the anonymity and confidentiality of their responses. While formal approval from an Ethics Committee was not sought for this study, the research followed ethical guidelines for data collection and participant consent. This includes ensuring voluntary participation, confidentiality, and transparency about the purpose of the study.

2.1. Instrument design and validation

The survey instrument was designed to assess students' perceptions and expectations regarding AI tutoring systems. It included 24 statements (see Appendix A) aimed at comparing the perceived benefits of AI tutors with those of human instructors. The key focus areas of the survey were skill development, information retention, adaptability to student needs, and personalized assessments. To ensure the validity and reliability of the questionnaire, Exploratory Factor Analysis (EFA) was conducted to uncover the underlying factor structure, followed by Confirmatory Factor Analysis (CFA) to confirm the structure and assess how well the model fit the data.

As for the Exploratory Factor Analysis, the scree plot, combined with parallel analysis as seen in Figure 1, indicates that there are three significant factors in the data. The first factor explains the most variance, followed by the second and third factors, each having eigenvalues slightly above 1. Factors beyond the third show eigenvalues less than or equal to 1, suggesting they do not significantly contribute to the explained variance and should not be retained. This conclusion is supported by the observed data points falling below the simulated data from the parallel analysis after the third factor.

Figure 1

Scree Plot of Eigenvalues from Exploratory Factor Analysis

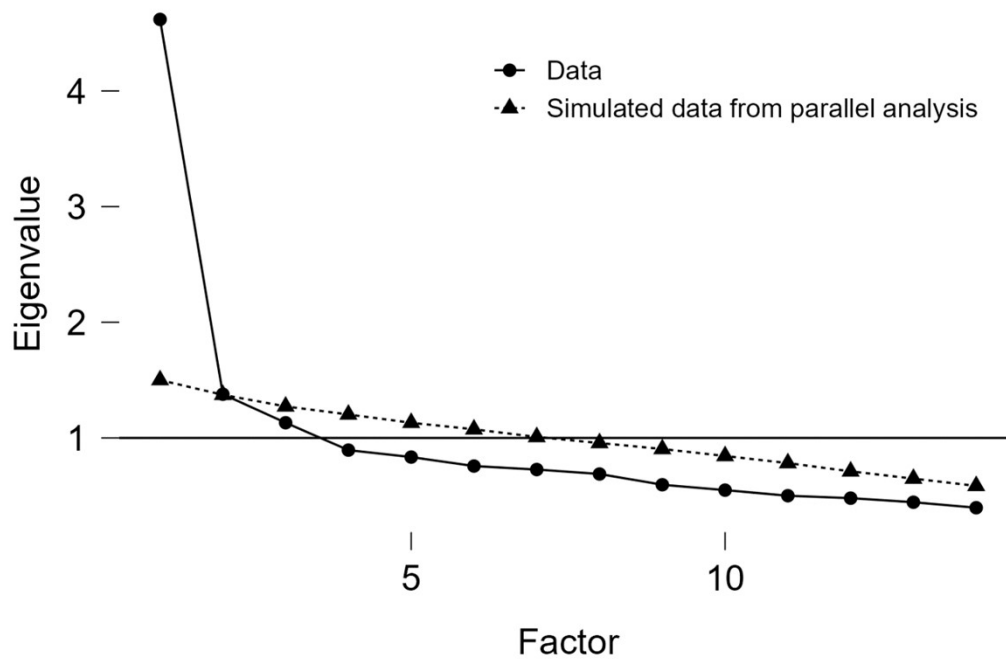


Table 1 presents the factor loadings for various items on three factors: Gains, Adaptability, and Guidance, with the "Uniqueness" column indicating the uniqueness scores of the items. Factor loadings indicate how strongly each item is associated with a particular factor. Here, the items (e.g., gains_2, gain_10) are listed in rows, and the factors are listed in columns. The values show the strength and direction of the association between each item and the factors. A cutoff of 0.4 is used, meaning only loadings of 0.4 or higher are considered significant.

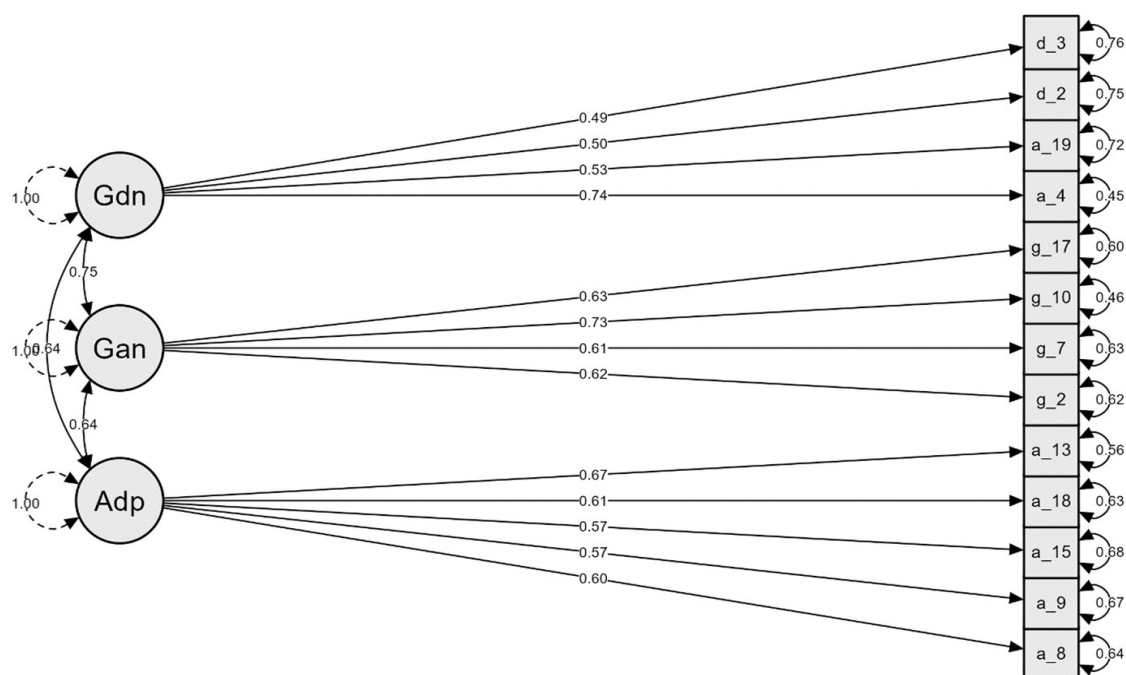
Table 1
 Factor loadings for the EFA

Items	Gains	Adaptability	Guidance	Uniqueness
gain_2	0.726			0.549
gain_10	0.602			0.526
gain_17	0.576			0.591
gain_7	0.576			0.562
adap_18		0.69		0.57
adap_15		0.584		0.654
adap_8		0.55		0.627
asse_9		0.537		0.673
asse_13		0.49		0.598
asse_19			0.741	0.584
asse_4			0.505	0.532
deve_24			0.465	0.729
deve_3			0.415	0.769

Regarding the Confirmatory Factor Analysis, the (CFA) path diagram (see Figure 2) illustrates the relationships between three latent factors—Guidance (Gdn), Gains (Gan), and Adaptability (Adp)—and their associated observed variables. Each latent factor is connected to several observed variables, with factor loadings indicating the strength of these relationships. For the Guidance factor, d_3 (0.74) and d_2 (0.75) show the strongest associations, followed by a_19 (0.50) and a_4 (0.49). The Gains factor is linked to g_17 (0.60), g_10 (0.46), g_7 (0.63), and g_2 (0.62), with notable strength in g_7 and g_2. Adaptability shows robust connections across all its observed variables: a_13 (0.61), a_18 (0.67), a_15 (0.68), a_9 (0.67), and a_8 (0.64). The model also indicates significant interrelationships among the latent factors, with correlations of 0.75 between Guidance and Gains, 0.64 between Gains and Adaptability, and 0.64 between Guidance and Adaptability. Measurement errors are depicted next to the observed variables, representing variance in the measurements. Overall, the CFA model demonstrates good convergent validity, with high factor loadings, and discriminant validity, supported by distinct correlations among different factors. This implies that the model represents the underlying constructs and their relationships with observed variables.

Figure 2

CFA Model Showing Four Factors of Intercultural Sensitivity



The adequacy of this CFA model was assessed using several fit indices, as summarized in Table 2. The chi-square to degrees of freedom ratio (χ^2/df) was comfortably below the value of 3.0, indicating an acceptable fit with a calculated χ^2/df of 1.039. This falls within the good-fit guidelines as recommended by Drasgow et al. (1995). The Comparative Fit Index (CFI) stood at .995, which, as suggested by Hu and Bentler (1999), suggests a model that fits the data well. The Tucker-Lewis Index (TLI) was at .994, above the recommended .95 indicating a good model fit. The Root Mean Square Error of Approximation (RMSEA) was well within acceptable limits at

.015, not exceeding the threshold of .06 set forth by Hu and Bentler (1999). Lastly, the Standardized Root Mean Square Residual (SRMR) was .04, which is below the .05 cutoff, as proposed by Byrne (2001), reflecting a good fit between the hypothesized model and the observed data. Overall, these fit indices suggest that the model is a good representation of the data.

Table 2

Fit indices

Index	Value
Comparative Fit Index (CFI)	0.995
Tucker-Lewis Index (TLI)	0.994
χ^2/df	1.039
RMSEA	0.015
SRMR	0.04

2.2.1. Reliability measures for the three factors

Table 3 presents reliability measures for three factors—Adaptability, Gains, and Guidance—and the total reliability across all factors, using Omega (ω) and Cronbach's Alpha (α) coefficients. Adaptability shows good internal consistency with $\omega = 0.738$ and $\alpha = 0.741$. Gains also demonstrates strong reliability, indicated by $\omega = 0.745$ and $\alpha = 0.744$. Guidance, while slightly lower, still maintains acceptable internal consistency with $\omega = 0.660$ and $\alpha = 0.655$. The overall reliability across all factors is high, with an Omega coefficient of 0.858 and a Cronbach's Alpha of 0.835, indicating that the combined items reliably measure the overall construct. These results suggest that the items within each factor and the total scale have good to excellent internal consistency, ensuring reliable measurement of the underlying constructs.

Table 3

Reliability measures

Factor	ω	α
Adaptability	0.738	0.741
Gains	0.745	0.744
Guidance	0.66	0.655
Total	0.858	0.835

3. RESULTS AND DISCUSSIONS

3.1. Comparing the Scales

Given the background of the study, which aims to measure students' expectations about AI tutoring in higher education, the results of the paired samples t-test, as seen in Table 4, provide insight into how students perceive these different aspects of AI tutoring. Adaptability scores are significantly higher than Gains scores ($t(183) = 6.256, p < .001, \text{Cohen's } d = 0.461$), indicating that students believe AI tutors are better at adapting to their individual needs compared to human tutors. The moderate effect size suggests higher expectations for adaptability over gains. Additionally, there is no significant difference between Adaptability and Guidance scores ($t(183) = 1.460, p = 0.146, \text{Cohen's } d = 0.108$), suggesting that students view the adaptability of AI tutors and the guidance provided by them similarly.

Furthermore, Guidance scores are significantly higher than Gains scores ($t(183) = -5.121, p < .001, \text{Cohen's } d = -0.378$), indicating that students perceive the guidance provided by AI tutors to be significantly better than the overall gains they expect to achieve from AI tutoring. The moderate effect size indicates higher expectations for the guidance aspect over the gains one.

Table 4

Comparison among scales using paired samples t-test

Measure 1	Measure 2	t	df	p	Cohen's d	SE Cohen's d
Adaptability	Gains	6.256	183	< .001	0.461	0.08
Adaptability	Guidance	1.46	183	0.146	0.108	0.077
Gains	Guidance	-5.121	183	< .001	-0.378	0.075

The findings reveal that students highly value AI tutors for their adaptability and guidance, seeing them as crucial for personalized learning. However, the lower scores for expected gains suggest skepticism about AI's ability to deliver tangible academic outcomes like skill development or knowledge retention. This could indicate that while students appreciate AI's ability to tailor learning and offer support, they remain unsure of its effectiveness in driving deeper learning or critical thinking. To meet student expectations, AI tutors must not only personalize learning but also clearly demonstrate their impact on measurable academic success, addressing these doubts.

3.1. Adaptability Across Groups

To compare the adaptability scores among different groups an ANOVA analysis was conducted. The results indicated a significant effect of the country of origin on adaptability scores, $F(2, 181) = 6.438, p = 0.002$, suggesting that the mean adaptability scores differ significantly between these groups. Post hoc comparisons using Tukey's HSD test revealed that Hungarian students have significantly higher adaptability scores than Spanish students, with a mean difference of 0.493 ($p < .001$). Descriptive statistics further showed that Hungarian students had the highest mean adaptability score (3.743) with relatively low variability ($SD = 0.535$), indicating more consistent perceptions. In contrast, Spanish students had the lowest mean

adaptability score (3.271) with higher variability (SD = 0.737), suggesting more diverse opinions within this group. Additionally, the analysis found no significant differences in adaptability scores across students from different fields of study, genders, or years of study. This may indicate that perceptions of AI tutoring adaptability are not influenced by these demographic factors, highlighting the specific influence of regional differences. The results show that Hungarian students perceive AI tutors as significantly more adaptable to their needs compared to Spanish students, likely due to cultural or educational differences. The consistency in Hungarian students' scores suggests a more uniform perception, while the greater variability among Spanish students indicates a wider range of views. No significant differences were found across gender, academic field, or year of study, suggesting regional context plays a larger role in shaping perceptions of AI adaptability. These findings emphasize the importance of considering cultural differences when implementing AI tutoring systems to better meet the expectations of diverse student populations.

3.2. Gains Scores Among Groups

The ANOVA results for the Gains scale indicate a significant interaction effect between gender and field of study, while the main effects of study and gender alone are not significant. Specifically, the significant interaction effect ($F(2, 178) = 5.706, p = 0.004$) suggests that the effect of field of study on Gains scores differs by gender. Post hoc comparisons reveal that female students in Social Sciences and Humanities have significantly higher Gains scores compared to male students in the same field (Mean Difference = 0.667, $p = 0.017$). Additionally, female students in Social Sciences and Humanities have significantly lower Gains scores compared to female students in Business and Economics (Mean Difference = -0.537, $p = 0.047$). These findings highlight the importance of considering both gender and field of study when assessing the perceived benefits of AI tutoring. Female students in Social Sciences and Humanities perceive greater gains compared to their male counterparts, while they perceive fewer gains compared to female students in Business and Economics. No significant main effects of study or gender alone were found, nor were there significant differences in Gains scores among students from different countries or years of study. This suggests that perceptions of gains from AI tutoring are specifically influenced by the interaction between gender and field of study rather than by these demographic factors individually.

3.3. Overall Expectations Scores

The ANOVA results for overall expectations scores, which combine the factors of Adaptability, Gains, and Guidance, show a significant effect of country, $F(2, 181) = 6.137, p = 0.003$. This indicates that the mean overall expectations scores differ significantly between the two country groups (Spain and Hungary). Post hoc comparisons using Tukey's HSD test reveal that Hungarian students have significantly higher overall expectations scores compared to Spanish students, with a mean difference of 0.370 ($p = 0.002$). The effect size (Cohen's d) of 0.643 suggests a substantial difference (Cohen, 1988), indicating that Hungarian students perceive AI tutoring more positively overall than their Spanish counterparts. Additionally, the analysis showed no significant differences in overall expectations scores across gender, year of study, or study specialization. This suggests that these demographic factors may not influence students' overall expectations of AI tutoring. The significant difference observed is specifically tied to the country

of the students, emphasizing the need to consider regional differences when evaluating and implementing AI tutoring systems in higher education.

The significant difference in overall expectations between Hungarian and Spanish students highlights the influence of cultural and educational contexts on how AI tutoring is perceived. Hungarian students' higher expectations may reflect greater familiarity or a stronger focus on adaptability and guidance in their education, while lower expectations in Spain may indicate skepticism or less exposure to AI in learning. The large effect size emphasizes that these differences are driven by regional factors rather than individual demographics. This suggests that AI tutoring systems must be adapted to align with the specific educational and cultural contexts of each country to meet student expectations effectively.

4. CONCLUSIONS

This study sought to understand the perceptions and expectations of higher education students in Hungary and Spain towards AI-powered tutoring systems. The research problem addressed the gap in literature concerning student attitudes and expectations in these specific cultural contexts. Using a validated questionnaire, we captured data on student expectations, satisfaction levels, and perceived benefits of AI tutors compared to human instructors. The key findings indicate that students generally perceive AI tutors positively, particularly appreciating their adaptability and guidance. Hungarian students showed higher overall expectations and adaptability scores compared to Spanish students. The study achieved its objective of providing an understanding of student attitudes towards AI tutoring, contributing valuable insights to the discourse on educational technology adoption.

The implications of these findings are present both theoretically and practically. Theoretically, this study enriches the existing knowledge on AI in education by highlighting regional differences in student perceptions, which had not been extensively explored before. It confirms the adaptability and guidance provided by AI tutors as critical factors in student satisfaction, aligning with existing research on AI's potential to personalize learning experiences (Chounta et al., 2022; Nguyen et al., 2024). Practically, the findings suggest that higher education institutions can enhance their attractiveness by integrating AI tutoring systems, which are viewed positively by students for their ability to provide continuous, personalized support. This can inform marketing strategies and resource allocation, making institutions more competitive (Chan, 2023). Furthermore, the study contributes uniquely by focusing on Hungarian and Spanish students, offering insights into cultural and/or regional influences on the acceptance of AI technologies. This cultural specificity is crucial for tailoring AI implementations to better meet the needs of diverse student populations. By addressing the expectations and satisfaction of students in these countries, the research provides actionable data that can guide the development and deployment of AI tutoring systems in higher education.

While this study offers valuable insights, it is essential to acknowledge its limitations. The sample was limited to students from Hungary and Spain, which may not represent the broader student population. Future research should consider expanding the study to include students from other countries and regions to provide a deeper understanding of global student perceptions. Additionally, the emotional and empathetic aspects of AI tutors were not deeply explored in this study. Furthermore, future research could investigate the cultural factors that

contribute to differences in perceptions between Hungary and Spain. A deeper exploration of these cultural dimensions would provide a more comprehensive understanding of how local values and educational norms shape students' expectations of AI tutoring systems.

Moreover, exploring the potential applications of AI tutoring systems beyond academic support could be beneficial. For instance, AI tutors could be utilized for career guidance, mental health support, and administrative assistance, broadening their impact on the student experience. Implementers of AI in education, managers of higher education institutions, and marketers could perhaps utilize these insights to optimize the integration and promotion of AI technologies, ensuring they meet the evolving needs of students effectively.

5. REFERENCES

- Allen, I. E., & Seaman, J. (2017). Digital learning compass: Distance education enrollment report 2017. Babson Survey Research Group.
- Al-Shanfari, L., Abdullah, S., Fstnassi, T., & Al-Kharusi, S. (2023). Instructors' perceptions of intelligent tutoring systems and their implications for studying computer programming in Omani higher education institutions. *International Journal of Membrane Science and Technology*, 10(2), 947-967. <https://doi.org/10.15379/ijmst.v10i2.1395>
- Basri, W. S. (2024). Effectiveness of AI-powered tutoring systems in enhancing learning outcomes. *Eurasian Journal of Educational Research*, 110, 33-52. 10.14689/ejer.2024.110.003
- Byrne, B. M. (2001). Structural equation modeling with AMOS, EQS, and LISREL: Comparative approaches to testing for the factorial validity of a measuring instrument. *International Journal of Testing*, 1(1), 55-86.
- Carbonell, J. R. (1970). AI in CAI: An artificial-intelligence approach to computer-assisted instruction. *IEEE Transactions on Man-Machine Systems*, 11(4), 190-202.
- Chan, C. (2023). A comprehensive AI policy education framework for higher education institutions. *Journal of Educational Technology*, 32(4), 215-230. <https://doi.org/10.1080/10494820.2023.1963527>
- Chaudhry, M. A., & Kazim, E. (2022). Artificial Intelligence in Education (AIEd): A high-level academic and industry note 2021. *AI and Ethics*, 2(1), 157-165. <https://doi.org/10.1007/s43681-021-00074-z>
- Chounta, I. A., Bardone, E., Raudsep, A., & Pedaste, M. (2022). Exploring teachers' perceptions of artificial intelligence as a tool to support their practice in Estonian K-12 education. *International Journal of Artificial Intelligence in Education*, 32(3), 725-755. <https://doi.org/10.1007/s40593-021-00243-5>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Baker, R.S., Inventado, P.S. (2014). Educational Data Mining and Learning Analytics. In: Larusson, J., White, B. (eds) Learning Analytics. Springer, New York, NY. https://doi.org/10.1007/978-1-4614-3305-7_4
- Dede, C. (2014). The Role of Digital Technologies in Deeper Learning. Students at the Center: Deeper Learning Research Series. *Jobs for the Future*.
- Double Robotics. (2017). Double Robotics: Homebound students connect to classrooms. Retrieved from <https://www.doublerobotics.com>

- Drasgow, F., Levine, M. V., Tsien, S., Williams, B. A., & Mead, A. D. (1995). Fitting polytomous item response models to multiple-choice tests. *Applied Psychological Measurement, 19*, 145-165. <https://doi.org/10.1177/014662169501900203>
- Edwards, A., Edwards, C., Spence, P. R., Harris, C., & Gambino, A. (2016). Robots in the classroom: Differences in students' perceptions of credibility and learning between "teacher as robot" and "robot as teacher". *Computers in Human Behavior, 65*, 627-634. <https://doi.org/10.1016/j.chb.2016.06.005>
- Hajeer, A. (2024). Teaching ESP in the digital age: Implications for crafting effective course descriptions for online learning. *Journal of Teaching English for Specific and Academic Purposes, 255*–267. <https://doi.org/10.22190/JTESAP231116021H>
- Hajeer, A., Tóptsi, J., & Horváth-Csikós, G. (2023). Validating the intercultural sensitivity scale in the Hungarian university context. *Cultural Management: Science and Education, 7*(2), 79-95. <https://doi.org/10.30819/cmse.7-2.05>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1-55. <https://doi.org/10.1080/10705519909540118>
- Kim, J., Merrill, K., Xu, K., & Sellnow, D. D. (2020). My teacher is a machine: Understanding students' perceptions of AI teaching assistants in online education. *International Journal of Human-Computer Interaction, 36*(20), 1902-1911. <https://doi.org/10.1080/10447318.2020.1801227>
- Li, J. (2015). The benefit of being physically present: A survey of experimental works comparing copresent robots, telepresent robots and virtual agents. *International Journal of Human-Computer Studies, 77*, 23-37. <https://doi.org/10.1016/j.ijhcs.2015.01.001>
- Mitra, P. P. (2021). Fitting elephants in modern machine learning by statistically consistent interpolation. *Nature Machine Intelligence, 3*(5), 378-386.
- Nguyen, A., Kremantzis, M., Essien, A., Petrounias, I., & Hosseini, S. (2024). Enhancing student engagement through artificial intelligence: Understanding the basics, opportunities, and challenges. *Journal of University Teaching and Learning Practice, 21*(6). <https://doi.org/10.53761/caraaq92>
- Park, S., & Whang, M. (2022). Empathy in human-robot interaction: Designing for social robots. *International journal of environmental research and public health, 19*(3), 1889. <https://doi.org/10.3390/ijerph19031889>
- Pfeifer, R., & Scheier, C. (1999). *Understanding intelligence*. MIT Press.
- Woolf, B. (2009). *Building intelligent interactive tutors: Student-centered strategies for revolutionizing e-learning*. Morgan Kaufmann.

To cite this work:

Hajeer, A., Papp-Váry, Árpád, & Pólya, Éva. Tutores de IA frente a instructores humanos: Percepciones de los estudiantes de educación superior en Hungría y España. *Edutec, Revista Electrónica de Tecnología Educativa, (89)*, XX-XX. <https://doi.org/10.21556/edutec.2024.89.3523>

Appendix A: Questionnaire

Survey: Attitudes towards AI tutoring

Please be assured that your participation in this survey is completely voluntary, and you can choose to withdraw at any time without any consequences. All responses will be kept strictly confidential and will be used for research purposes only. Your answers will be anonymized, ensuring that no personally identifiable information is collected. We do not share any data with third parties, and the results will be reported in aggregate form only.

By participating, you are contributing valuable insights that may help improve understanding AI tutoring higher education. The survey should take approximately 5-10 minutes to complete.

Thank you for your time and valuable contribution to this research!

Age :

Gender:

Country of origin:

Year of study:

Major field of study:

1. I expect an AI tutor to be able to adjust the difficulty level of the learning material based on my understanding of the topic. (Strongly disagree, Disagree, Neutral, Agree, Strongly agree)
2. I believe explanations provided by AI tutors would be clearer and easier to understand than explanations from a human tutor. (Strongly disagree, Disagree, Neutral, Agree, Strongly agree)
3. I believe an AI tutor could effectively guide me through the process of breaking down complex problems (problem-solving) into smaller, more manageable steps.
4. Compared to a human teacher, I believe an AI tutor could better explain the reasoning behind its assessments, such as why an answer is incorrect. (Strongly disagree, Disagree, Neutral, Agree, Strongly agree)
5. I think using an AI tutor would encourage me to think more critically about the course material (analyze information, identify biases, etc.) than traditional tutoring methods. (Strongly disagree, Disagree, Neutral, Agree, Strongly agree)
6. I believe AI tutors would be better at identifying my learning gaps compared to traditional tutoring methods. (Strongly disagree, Disagree, Neutral, Agree, Strongly agree)
7. I expect using an AI tutor would help me understand complex topics better than a human tutor. (Strongly disagree, Disagree, Neutral, Agree, Strongly agree)
8. I expect an AI tutor to be able to personalize the learning experience based on my preferred learning style. (Strongly disagree, Disagree, Neutral, Agree, Strongly agree)
9. Compared to human teacher feedback, I expect AI tutors to offer more frequent feedback on my learning progress. (Strongly disagree, Disagree, Neutral, Agree, Strongly agree)
10. I expect explanations provided by AI tutors would be clearer and easier to understand than explanations from a human tutor. (Strongly disagree, Disagree, Neutral, Agree, Strongly agree)
11. Compared to traditional tutoring methods, I expect AI tutors would be more helpful in teaching me to evaluate the strengths and weaknesses of different approaches to solving problems. (Strongly disagree, Disagree, Neutral, Agree, Strongly agree)
12. I believe an AI tutor could track my progress over time and provide personalized feedback better than a human teacher. (Strongly disagree, Disagree, Neutral, Agree, Strongly agree)


13. Compared with human teachers, I expect feedback from an AI tutor to be more specific and actionable, helping me understand how to improve my learning.
14. I expect AI tutors could improve my long-term knowledge of a subject more effectively than traditional tutoring methods. (Strongly disagree, Disagree, Neutral, Agree, Strongly agree)
15. I believe AI tutors would be better at focusing on areas where I need the most help compared to traditional tutoring methods. (Strongly disagree, Disagree, Neutral, Agree, Strongly agree)
16. I expect AI tutors would be better at providing me with opportunities to practice critical thinking skills (e.g., analyzing evidence, forming arguments) compared to traditional tutoring methods. (Strongly disagree, Disagree, Neutral, Agree, Strongly agree)
17. I believe I would learn and remember information faster if I used an AI tutor instead of traditional tutoring methods. (Strongly disagree, Disagree, Neutral, Agree, Strongly agree)
18. I believe an AI tutor could adjust the pace of the tutoring session based on how quickly I grasp new information. (Strongly disagree, Disagree, Neutral, Agree, Strongly agree)
19. Compared to human teacher, I believe an AI tutor could more effectively identify areas where I made mistakes in my work and suggest improvements. (Strongly disagree, Disagree, Neutral, Agree, Strongly agree)
20. After working with an AI tutor (hypothetically), I would feel more confident in my ability to apply my knowledge to solve new and unfamiliar problems. (Strongly disagree, Disagree, Neutral, Agree, Strongly agree)
21. I think using an AI tutor would provide me with various assessment methods (e.g., quizzes, exercises) to gauge my understanding. (Strongly disagree, Disagree, Neutral, Agree, Strongly agree)
22. I expect AI tutors would be more effective in providing alternative explanations if I don't understand a concept the first time. (Strongly disagree, Disagree, Neutral, Agree, Strongly agree)
23. Compared to traditional tutoring methods, I expect AI tutors would be more effective in helping me remember key facts and details. (Strongly disagree, Disagree, Neutral, Agree, Strongly agree)
24. I think using an AI tutor would encourage me to be better at identifying different sources of information when approaching problems, compared to traditional tutoring methods. (Strongly disagree, Disagree, Neutral, Agree, Strongly agree)


Scales: Adaptability(1, 6, 8, 15, 18, 22), Learning Gains (2, 7, 10, 14, 17, 23) Skills Development (3, 5, 11, 16, 20, 24), Assessment (4, 9, 12, 13, 19, 21).



Impacto de los Sistemas de Tutoría Inteligente. Una revisión sistemática

Impact of Intelligent Tutoring Systems. A systematic review

 Noelia Carbonell Bernal; noelia.carbonell@unir.net
Universidad Internacional de la Rioja (España)

 M^a Ángeles Hernández Prados; mangeles@um.es;
Universidad de Murcia (España)

Resumen

La Inteligencia Artificial desarrolla herramientas eficientes en diversas esferas, incluida la educación, donde los Sistemas de Tutoría Inteligente (STI) tienen gran relevancia. Este estudio tiene como objetivo analizar la producción científica sobre los STI en los procesos de enseñanza-aprendizaje, utilizando los criterios PRISMA como metodología. Se emplearon las bases de datos Scopus, Web of Science y ERIC, seleccionando 20 artículos publicados entre 2013 y 2023 mediante estrategias de búsqueda y criterios de inclusión/exclusión. El análisis aborda la tendencia cronológica, país de publicación, características de la muestra, ámbito de aplicación, los STI y los resultados obtenidos. Los hallazgos destacan el impacto del COVID-19 en la producción científica, las mejoras en el rendimiento académico de los estudiantes gracias a los STI y la importancia de una coordinación interdisciplinar. Las conclusiones subrayan que los STI influyen positivamente en los procesos de enseñanza-aprendizaje al mejorar la participación y el compromiso de los estudiantes, personalizar los contenidos y apoyar la autorregulación, adaptando la dificultad a las capacidades individuales de los estudiantes.

Palabras clave: Inteligencia Artificial, Procesos de Enseñanza-Aprendizaje, Revisión Sistemática, Sistemas de Tutoría Inteligente

Abstract

Artificial Intelligence develops efficient tools in various fields, including education, where Intelligent Tutoring Systems (ITS) play a significant role. This study aims to analyze the scientific output on ITS in teaching-learning processes, using PRISMA criteria as the methodological framework. The databases Scopus, Web of Science, and ERIC were used, selecting 20 articles published between 2013 and 2023 through search strategies and inclusion/exclusion criteria. The analysis focuses on the chronological trend, country of publication, sample characteristics, application scope, ITS features, and study outcomes. The findings highlight the impact of COVID-19 on scientific output, improvements in students' academic performance through ITS use, and the importance of interdisciplinary coordination. The conclusions emphasize that ITS positively influence teaching-learning processes by enhancing student engagement and participation, personalizing content, supporting self-regulation, and adapting difficulty to individual students' abilities.

Keywords: Artificial Intelligence, Intelligent Tutoring System, Teaching Learning Processes, Systematic Review



1. INTRODUCCIÓN

Las Tecnologías de la Información y Comunicación (TIC) han transformado profundamente la sociedad, afectando la forma de aprender, pensar y actuar. Los contextos educativos también se han adaptado a estas demandas, generando oportunidades e incertidumbres (Alahi et al., 2023), siendo esencial redefinir el rol docente para que sea competente digitalmente, capaz de guiar a los estudiantes, adaptarse e integrar las TIC en el aula de manera efectiva (Achard, 2020). Deben ser proactivos y reflexivos en el uso de estas tecnologías, que siguen vinculadas principalmente a la gestión de datos, calificación rápida y motivación de los estudiantes, a pesar de las amplias aplicaciones que ofrecen (Gros-Salvat y Cano-García, 2021). Esto subraya la necesidad de un enfoque más profundo y creativo en la incorporación de las TIC en la educación.

A pesar de que la competencia digital cuenta con el respaldo de instituciones como la Comisión Europea (Redecker, 2020) y la UNESCO (2019), los avances en las Tecnologías de la Información y Comunicación (TIC) siguen generando debate, especialmente entre sectores más conservadores. La llegada de la Inteligencia Artificial (IA) no ha sido la excepción, provocando un impacto revolucionario en diversos sectores sociales. A pesar de las preocupaciones éticas y sociales relacionadas con la privacidad, seguridad y justicia social, la IA ha demostrado su potencial para mejorar la eficiencia y calidad de vida en ciudades inteligentes, optimizando la gestión de tráfico, energía, residuos y seguridad pública (Alahi et al., 2023).

En el ámbito educativo, la IA está transformando los métodos pedagógicos. Herramientas como ChatGPT han alterado significativamente el panorama de la enseñanza formal (García-Peñalvo, 2023). La IA permite desarrollar sistemas adaptados a las necesidades individuales de los estudiantes, brindando retroalimentación instantánea y personalizada, además de facilitar el acceso a recursos educativos diversos. También ayuda a los educadores a procesar grandes volúmenes de información y automatizar tareas administrativas, optimizando la enseñanza (García-Peña et al., 2020; Troncoso et al., 2023). En este contexto, los Sistemas Tutoriales Inteligentes (STI) juegan un papel clave. Surgidos en los años sesenta con el programa "Tutor", los STI han evolucionado hacia plataformas más interactivas y personalizadas, como "Algebra Tutor", "Andes Physics Tutor", y los más recientes "ALEKS" y "ASSISTments", que ofrecen tutorías adaptativas en diversas materias (Urquilla, 2022). Un ejemplo destacado es ArtiBos, un sistema de tutoría inteligente que mejora las habilidades de resolución de problemas, impulsando el pensamiento crítico y analítico de los estudiantes, permitiéndoles resolver problemas con mayor precisión y rapidez (Çetin et al., 2023).

Los Sistemas Tutoriales Inteligentes (STI) son plataformas diseñadas para emular la tutoría humana, ofreciendo apoyo personalizado a los estudiantes mediante orientación y retroalimentación adaptadas a sus necesidades (Mathew et al., 2021). Estos sistemas mejoran la eficiencia de la enseñanza, permitiendo que los estudiantes avancen a su ritmo y se concentren en áreas donde necesitan más ayuda (Troncoso et al., 2023; Vera, 2023). Además, superan limitaciones como la falta de recursos y ratios altos de alumnos, proporcionando un acompañamiento individualizado (Gupta y Chen, 2022). Los STI permiten crear perfiles de aprendizaje detallados mediante el seguimiento del progreso de los estudiantes (Ponce et al., 2019).

El potencial de los STI ha sido demostrado en diversas áreas educativas. Por ejemplo, Heck et al. (2022) destacan su efectividad en la enseñanza de idiomas extranjeros, donde las herramientas digitales permiten itinerarios personalizados y adaptativos en función de las necesidades del estudiante.

Sin embargo, los STI también enfrentan desafíos como la excesiva personalización, que puede reducir las interacciones sociales (Moreno, 2019), o la falta de empatía y aspectos emocionales en la tutoría, aunque se están desarrollando STI afectivos para abordar este problema (Ponce et al., 2019). Además, se plantean desafíos como la dependencia tecnológica, la falta de una base pedagógica sólida, sesgos, y resistencia al cambio (Selwin, 2020; Tirochi, 2022). La computación afectiva podría mitigar estos problemas al incorporar la detección y transmisión de estados emocionales, lo que ayudaría a fortalecer el compromiso con el aprendizaje (Carini et al., 2006; Webster, 1997).

La inteligencia artificial (IA) está influyendo en los procesos de enseñanza-aprendizaje, entendidos como interdependientes y complementarios, manteniendo una interacción dialéctica constante (Osorio et al., 2022; De la Rosa et al., 2019). Aunque el aprendizaje puede ocurrir sin enseñanza, esta solo tiene sentido si acompaña al aprendizaje (Moreira, 2004). La enseñanza es una actividad de interacción social, y el aprendizaje depende de variables como la experiencia-interacción del sujeto con su entorno sociocultural, por lo que la tecnología no debe dominar estos procesos.

Para que el aprendizaje sea eficaz, es fundamental que existan motivación, concentración, estrategias pedagógicas adecuadas y una correcta evaluación del proceso (Vera, 2023). Algunos estudios recientes van más allá de los aspectos técnicos y logros académicos, analizando factores emocionales. Ni y Cheung (2023) proponen un modelo que predice el uso de sistemas de tutoría inteligente (STI) en estudiantes secundarios chinos, destacando factores como la utilidad percibida, el disfrute y la ansiedad. Por otro lado, Cao et al. (2021) señalan que la actitud de los estudiantes hacia las plataformas de tutoría está influenciada por el impacto social, y Zuo et al. (2022) subrayan que la facilidad de uso percibida y la autoeficacia son claves para su aceptación.

Diversas revisiones han mostrado la eficacia potencial de los STI en comparación con la tutoría humana. VanLehn (2011) destaca su efectividad, mientras que Ocaña Fernández et al. (2019) indican que estos sistemas optimizan la enseñanza-aprendizaje, aunque requieren una adecuada formación docente (Martín-Marchante, 2022). González-Palacios y Avelino-Rubio (2016) señalan la necesidad de reflexionar sobre el propósito pedagógico de los STI y cómo aprovecharlos para enriquecer el aprendizaje. Investigaciones recientes, como la de Pezzini (2022), también analizan su adaptación en secundaria, y otros estudios críticos destacan la falta de familiaridad con los STI como un obstáculo, agravando problemas de autoeficacia y ansiedad que pueden reducir la satisfacción y aumentar la deserción (Cao et al., 2021; Zuo et al., 2022).

Estas investigaciones han identificado oportunidades y desafíos clave en el ámbito de la IA y la tecnología educativa. Proporcionan una base crítica para guiar enfoques que beneficien a los estudiantes y mejoren la calidad del aprendizaje, además de enriquecer la comprensión sobre su implementación efectiva reflexiva. No cabe duda, que un mayor conocimiento de la producción científica acerca de esta temática resulta útil para acallar los temores e inseguridades que aparecen reflejados en los estudios comentados anteriormente, así como facilitar la comprensión de cómo estos sistemas pueden adaptarse y beneficiar a diferentes niveles educativos, proporcionando una base sólida para la investigación actual sobre los STI y su impacto en el proceso de enseñanza-aprendizaje. Concretamente, el problema de investigación que se plantea, atendiendo al método PICOT (Participantes, Intervención, Comparación, Objetivos y Tiempo) es: ¿Qué producción científica existe sobre la influencia de Sistemas de Tutoría Inteligente (I) en los procesos de aprendizaje (O) de estudiantes adolescentes (P) en el período temporal de 2013-2023 (T)? En el mismo se exponen dos núcleos principales: inteligencia artificial y procesos de aprendizaje. Atendiendo a lo expuesto, y al problema de investigación, se formula el objetivo general consistente en: analizar, siguiendo la metodología PRISMA, la producción científica existente sobre la influencia de los Sistemas de Tutoría Inteligente en los procesos de enseñanza-aprendizaje de los alumnos, el cual se concreta en los siguientes objetivos específicos:

1. Explorar la evolución cronológica y el contexto geográfico de la producción científica sobre la influencia de los STI en los procesos de enseñanza-aprendizaje, centrándose en los artículos publicados entre 2013 y 2023.
2. Clasificar la investigación existente sobre la influencia de los STI en los procesos de aprendizaje en función a los aspectos metodológicos, específicamente, mediante el análisis de las muestras utilizadas en los estudios, teniendo en cuenta elementos como aspectos metodológicos del estudio, concretamente, el tamaño de la muestra, la etapa educativa y el diseño del estudio.
3. Analizar la producción científica sobre STI en el aprendizaje de los estudiantes, considerando diversos contextos de aplicación como entornos escolares, universitarios y otros ámbitos educativos analizando las fortalezas y debilidades de estos.
4. Analizar los hallazgos más pertinentes obtenidos en la investigación científica relacionada con la repercusión de las STI en los procesos de enseñanza-aprendizaje del alumnado.

Con esta revisión sistemática, se pretende aportar información sobre los Sistemas de Tutoría Inteligente y su influencia en el aprendizaje de los estudiantes, que será de utilidad para futuros desarrollos y mejoras en el ámbito educativo.

2. MÉTODO

Para identificar, evaluar e interpretar la investigación disponible en el ámbito de interés, se ha empleado el enfoque de revisión sistemática, siguiendo los lineamientos del protocolo PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses), el cual establece un conjunto de criterios para garantizar la calidad y transparencia en la realización de revisiones

sistemáticas y meta-análisis (Urrútia y Bonfill, 2010), así como las fases establecidas por Ramírez y Lugo (2020).

En primer lugar, se formulan las siguientes preguntas de investigación (fase 1): ¿Cuál es la definición y funcionamiento de los Sistemas de Tutoría Inteligente? ¿Qué impacto generan los STI en los procesos de enseñanza y aprendizaje de los estudiantes? ¿Qué aspectos positivos y negativos se asocian con la incorporación de estos sistemas en la educación?, que actuarán como eje conductor de todo el proceso de investigación.

A continuación, se inicia el proceso de búsqueda (fase 2), empleando “Joven, adolescente, estudiante, inteligencia artificial, sistema de tutoría inteligente” en castellano, y “Young, teenager, student, artificial intelligence, Intelligent Tutoring System” en inglés, como descriptores de búsqueda en las siguientes bases de datos: Scopus, Web of Science y ERIC. De forma complementaria, para afinar más en la búsqueda, se han utilizado los operadores booleanos lógicos AND y OR, los booleanos de proximidad () para buscar con prioridad, "" para buscar los conceptos completos y el operador booleano de truncamiento * y encontrar conceptos en todas sus formas, quedando las frases de búsqueda de la siguiente manera: (“joven” OR “adolescente*” OR “estudiante*”) AND “Inteligencia Artificial” AND “Sistema de Tutoría Inteligente”, y (“young” OR “teenager*” OR “student*”) AND "Artificial Intelligence" AND "Intelligent tutoring system". Finalmente, se delimita el período temporal (2013-2023) y el tipo de documento (artículo), arrojando un resultado de 1.602 artículos, de los cuales se eliminaron los repetidos quedando un total de 1302.

Luego, se establecen los criterios de inclusión y exclusión, tal y como se muestran en la tabla 1, que permitirán ajustar los documentos convenientemente al objeto de estudio (fase 3). Como consecuencia de la aplicación de estos criterios, el volumen de artículos se redujo a 118.

Tabla 1

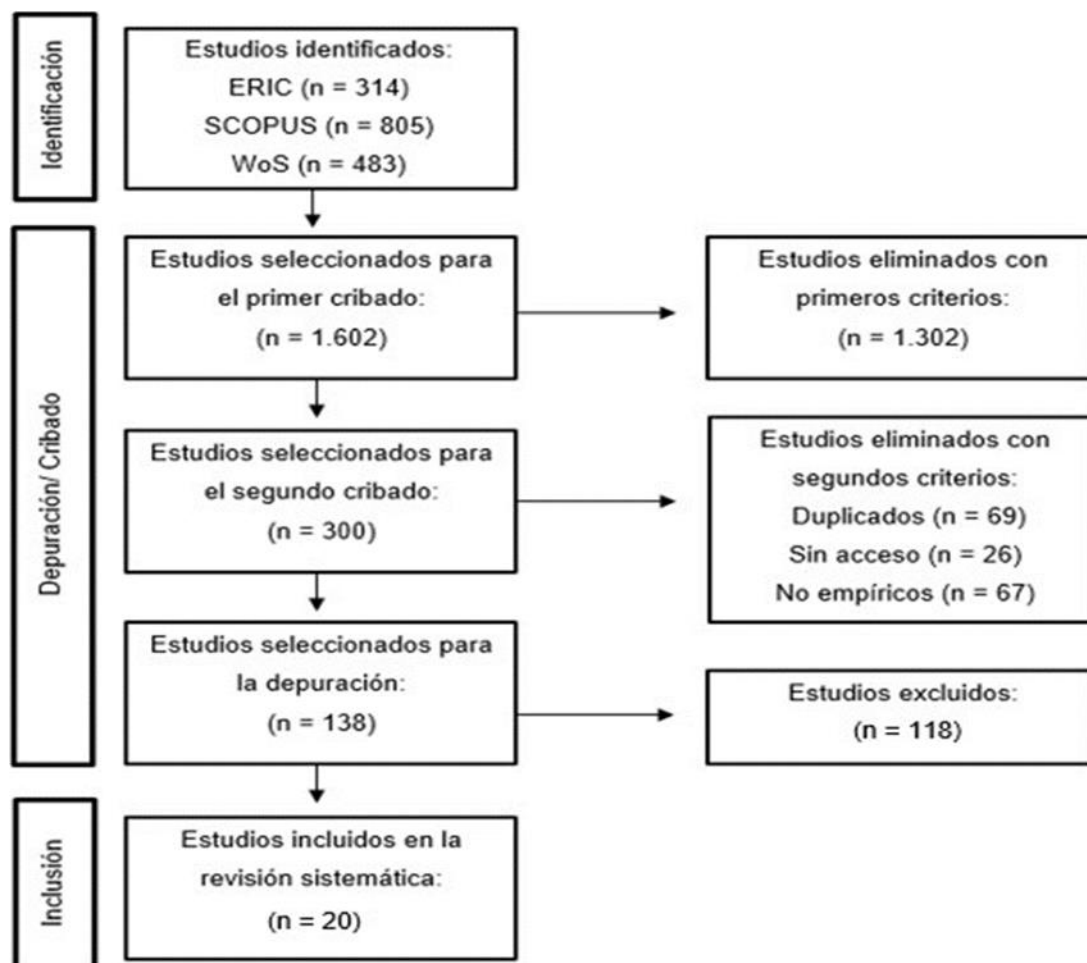
Criterios de inclusión y exclusión

<i>Criterios de inclusión</i>	<i>Criterios de exclusión</i>
Artículos publicados en revistas científicas.	Documentos que no sean artículos científicos
Investigaciones de carácter empírico.	Investigaciones que no sean de carácter empírico
Artículos con acceso al texto completo.	Textos completos no disponibles
Publicaciones comprendidas entre 2013 y 2024.	Publicaciones anteriores a 2013
Estudios de cualquier país.	Redactados en otro idioma que no sea inglés ni español
Redactados en inglés o español.	
Estudios que incluyan como muestra a estudiantes de secundaria.	

Para garantizar el ajuste a los objetivos, se procedió a la lectura de los resúmenes, y aquellos que no aportan información relevante relativo a los STI en la etapa que abarca de los 12 a los 20 años fueron descartados, reduciéndose a 54 artículos, de los cuales se seleccionaron los que incluyen en el título o en las palabras clave “Sistema tutoría inteligentes”, quedando 20 artículos.

Figura 1

Diagrama de flujo



Seguidamente, se seleccionaron los estudios pertinentes (Tabla 2) y se empezaron a extraer los datos relevantes (fase 4). Por último, realizamos la síntesis de los datos obtenidos exponiéndose en función de los objetivos de investigación, tal y como se muestra en el apartado siguiente (fase 5).

3. RESULTADOS

Atendiendo al primer objetivo, se exponen los datos bibliométricos que nos permiten conocer la tendencia de investigación sobre la temática. En lo que respecta a la cronología, no existe una tendencia claramente definida al incremento o descenso de la producción, ya que se dan grandes variaciones. Tan solo es destacable el punto máximo alcanzado en 2020, año en el que la producción científica se incrementó extremadamente a causa del confinamiento por la COVID, y al empuje tan evidente que experimentaron las TIC ante el cierre de los centros educativos, tal y como señalan Sanz y López, (2021). Este patrón podría ser congruente con la observación realizada por Alemán-Díaz. (2023) en su estudio sobre las dinámicas de publicación en áreas afines, donde también se apreciaron fluctuaciones en la producción científica en

relación con eventos de relevancia global, como la pandemia. Aunque es preciso señalar que esta variabilidad podría atribuirse al alcance limitado de la muestra analizada.

Otro de los aspectos contemplados, el país de publicación, evidencia que Estados Unidos se muestra altamente interesado en esta temática, ya que publicó un total de siete artículos, la cifra más alta alcanzada, seguido muy de lejos por España quien editó tres de los 20 artículos muestrales. También es relevante destacar la notoria manifestación de producción científica en naciones de América Latina, específicamente Brasil, así como en naciones del sudeste asiático, tales como Indonesia y Malasia. Estas manifestaciones podrían indicar un signo prometedor de la democratización de las aplicaciones de Ciencia, Tecnología e Innovación en el contexto de promover la equidad en el acceso a la educación. No obstante, es imperativo señalar la omisión de publicaciones en el contexto africano, lo cual podría servir como un indicador tangible de la disparidad digital que prevalece en la actualidad, así como la carencia de recursos e infraestructura que obstaculizan la realización de investigaciones y proyectos en el ámbito de las STI (Castillo et al, 2021, Rodríguez-Málaga et al, 2019).

El segundo objetivo se centra en los aspectos metodológicos del estudio, concretamente, en el tamaño de la muestra, la etapa educativa y el diseño del estudio (diseño experimental, cuasi-experimental). Se debe considerar, en primer lugar, la relación intrínseca que mantiene el tamaño de la muestra con la calidad de los resultados y la capacidad de generalización y extrapolación de las conclusiones a la población general (Vanh Lehn et al., 2023), por lo que, según Johnson y Christensen (2019), si el tamaño de la muestra es adecuado mejora la representatividad de la población estudiada, la robustez y aplicabilidad de las conclusiones, y, por ende, la validez externa de los hallazgos, contribuyendo así a un proceso investigativo más sólido y confiable.

Al examinar los datos recopilados, plasmados en la Tabla 2, se constata en primer lugar, que se han recogido valores que oscilan entre 52 y 537 participantes en los estudios. Esta disparidad se debe a la variedad de estudios, siendo menor la muestra en evaluaciones de intervenciones educativas con STI, tal y como se comprueba en la experiencia educativa reflejada en los diferentes estudios revisados. Por ejemplo, en la intervención pre-post para evaluar cómo la implementación de STI afectó el aprendizaje de Arroyo (2014), participaron un total de 94 estudiantes de secundaria. Sin embargo, con una finalidad similar, Grivokostopoulou et al. (2021) contó con la participación de un total de 396 estudiantes de secundaria. La diferencia se explica por varias casuísticas como las peculiaridades de aplicabilidad en diferentes contextos educativos, de disponibilidad de participantes y de preguntas de investigación propias.

Los estudios sobre Sistemas Tutoriales Inteligentes (STI) muestran una variabilidad en el tamaño de las muestras y los enfoques metodológicos, influyendo en la generalización de los resultados. Investigaciones como la de VanLehn et al. (2020) y Grivokostopoulou et al. (2021) usaron muestras grandes de 537 y 522 participantes, respectivamente, lo que aporta robustez a los resultados. Otros estudios, como los de Subirats et al. (2021) y VanLehn et al. (2016), emplearon muestras menores (52 y 55 participantes), ofreciendo una visión más detallada del aprendizaje pero limitando la generalización de los hallazgos. Además, estudios como Ramos de Melo et al. (2014) y VanLehn et al. (2023) proporcionan información valiosa sobre cómo los STI afectan la motivación y el aprendizaje, aunque la falta de detalles sobre el tamaño de las muestras puede complicar la evaluación de la validez externa de los resultados.

La Tabla 2 clasifica los estudios según la etapa educativa y el diseño de investigación. Los artículos analizados se centran mayoritariamente en la educación secundaria (55%) y universitaria (45%), con una notable ausencia de estudios en educación Infantil y Primaria. Este sesgo hacia la educación superior podría reflejar una mayor disponibilidad de recursos tecnológicos y su impacto en la formación universitaria. Sin embargo, esto sugiere también la necesidad de adaptar y evaluar los STI en niveles educativos inferiores.

La mayoría de los estudios en secundaria utilizan diseños pre-post para evaluar la eficacia de los STI en la mejora del rendimiento académico. Algunos emplean grupos de control (McCarthy et al., 2018) o diseños de discontinuidad de regresión (VanLehn et al., 2020). Los resultados indican que los STI tienen un impacto positivo en el rendimiento académico, con un 68.3% de los estudios mostrando mayor participación y compromiso de los estudiantes. Factores como la personalización del contenido y el apoyo en la autorregulación contribuyen a estos resultados positivos.

En el ámbito universitario, los STI también demuestran beneficios significativos, como la mejora en el rendimiento y el aumento del compromiso estudiantil (Grivokostopoulou et al., 2021; Subirats et al., 2021). La personalización y la atención individualizada son cruciales, dadas las diversas necesidades de los estudiantes universitarios (Eryilmaz & Adabashi, 2020). Los estudios muestran que la implementación de STI en contextos universitarios puede mejorar el rendimiento académico y el compromiso de los estudiantes.

La mayoría de los estudios utilizan enfoques experimentales (60%) y cuasi-experimentales (40%). Los diseños experimentales comparan grupos con y sin STI (Rahim et al., 2020), mientras que los diseños pre-post evalúan cambios dentro de un mismo grupo antes y después de la intervención (Arroyo et al., 2014). Algunos estudios incluyen más de dos grupos para comparar diferentes versiones del STI o herramientas digitales (Ramos de Melo et al., 2014; Weston-Sementelli et al., 2016). Los estudios universitarios utilizan tanto enfoques experimentales controlados (Subirats et al., 2021) como metodologías avanzadas como el diseño de discontinuidad de regresión (VanLehn et al., 2020), proporcionando una comprensión completa del impacto de los STI en diferentes contextos educativos.

En conclusión, la investigación analizada revela una amplia gama de enfoques metodológicos utilizados para examinar la influencia de los Sistemas de Tutoría Inteligente en los procesos de enseñanza-aprendizaje. Estos enfoques permiten una comprensión más profunda de los resultados en distintos contextos educativos y contribuyen significativamente a la evolución del campo al brindar perspectivas diversas sobre la efectividad y aplicabilidad de estos sistemas.

Tabla 2

Etapa educativa y tipo de diseño de los estudios

Artículo	N	Etapa	Diseño	Artículo	N	Etapa	Diseño
1.Arroyo (2014)	94	Secundaria	Pre post	11. Frontino de Medeiros et al. (2020)	100	Secundaria	Experimental
2.Ramos de Melo et al. (2014)	NA	Secundaria	Pre post	12.Eryilmaz & Adabashi (2020)	105	Secundaria	Experimental
3.Wolfe et al. (2014)	410	Universidad	Pre post	13.De Chiusole et al. (2020)	76	Secundaria	Experimental
4.Guerrero-Roldán et al. (2016)	400	Un iversidad	Experiment al	14.Gunawan et al. (2020)	152	Secundaria	Experimental
5.eBahçeci & Gürol (2016)	162	Secundaria	pretest-posttest control group	15.Eryilmaz & Adabashi (2020)	136	Un iversidad	Experimental
6.VanLehn et al. (2016)	55	secundaria	Experiment al	16.Furlan et al. (2021)	103	Un iversidad	Experimental
7.McCarthy et al. (2018)	234	Secundaria	pretest-posttest control group	17.Grivokostopou lou et al. (2021)	522	Un iversidad	Experimental
8.Weston Sementelli et al. (2018)	175	Secundaria	Experiment al	18.Grivokostopou lou et al. (2021)	396	Secundaria	Experimento de campo
9. VanLehn et al. (2020)	537	Secundaria	regression discontinuit y design	19.Subirats et al. (2021)	52	Un iversidad	Experimental
10. Tacoma et al. (2020)	151	Secundaria	pretest-posttest control group	20.VanLehn et al. (2023)	151	Secundaria	Experimental

En cuanto a la aplicabilidad de los Sistemas de Tutoría Inteligente (STI), se ha considerado su integración en materias curriculares y las características de las herramientas empleadas. La mayoría de los artículos analizados han implementado los STI en materias relacionadas con Informática (100%), como Ingeniería, Computación e Inteligencia Artificial. También se ha subrayado su relevancia en Matemáticas (40%) y Ciencias, como Biología, Física y Biotecnología, así como en Comprensión Lectora, Economía y Medicina, aunque solo

representaron un 10% cada una. No se han encontrado experiencias en disciplinas más sociales y humanísticas, al menos en los estudios analizados, lo que refleja una limitación en su aplicación.

Independientemente de la materia, la unicidad observada en su implementación podría indicar una resistencia al cambio en la comunidad docente, a pesar de las ventajas de estas tecnologías. Factores como la falta de familiaridad con las nuevas herramientas, la percepción de complejidad en su implementación y la tradición educativa más convencional podrían influir en esta reticencia.

Además, la Tabla 3 destaca la diversidad de herramientas STI disponibles, con un total de 20 sistemas registrados, cada uno configurado para contextos particulares. Estas herramientas se caracterizan por su capacidad de personalización, presentación multinivel de contenidos, retroalimentación constante y evaluación del desempeño. Su adaptabilidad y flexibilidad permiten enriquecer los procesos de enseñanza-aprendizaje en diversos entornos educativos, promoviendo la innovación en la educación mediante tecnología inteligente.

Tabla 3

Herramientas Utilizadas en los Sistemas de Tutoría Inteligente

Artículo	Herramientas Utilizadas
Arroyo et al. (2014)	Multimedia Adaptativo, Cognición, Metacognición, Emoción. Emplearon una combinación de herramientas que abordaron aspectos cognitivos y emocionales, utilizando contenido multimedia adaptativo para promover una comprensión profunda y estrategias metacognitivas para mejorar la autorregulación del aprendizaje.
Ramos de Melo et al.(2014)	Organización Computacional, Aprendizaje Personalizado Se centraron en la personalización de la experiencia de aprendizaje, adaptando la tutoría y las actividades según las necesidades individuales de los estudiantes.
Wolfe et al. (2014)	Sistema de Tutoría Web, Teoría de Trazo Difuso Utilizaron un sistema de tutoría web y la teoría del trazo difuso para comunicar el riesgo genético de cáncer, buscando mejorar la comprensión del riesgo a través de la web.
Guerrero-Roldán et al. (2016)	Sistema Educativo, Algoritmos de Búsqueda Utilizaron un sistema educativo centrado en el aprendizaje de algoritmos de búsqueda, brindando una oportunidad para desarrollar habilidades en algoritmos y evaluación automática de estudiantes.
eBahçeci y Gürol (2016)	Instrucción Individualizada, Aprendizaje Personalizado Implementaron instrucción individualizada y enfoques de aprendizaje personalizado para potencialmente mejorar el rendimiento académico a través de la adaptación a las necesidades de los estudiantes.
VanLehn et al. (2016)	Dragoon (Sistema), Ciencias y Modelos Utilizaron el sistema Dragoon para el aprendizaje de ciencias a través de la construcción de modelos, con el objetivo de aumentar el aprendizaje sin requerir más tiempo dedicado.
McCarthy et al. (2018)	Funciones Metacognitivas, Retroalimentación Incluyeron funciones metacognitivas y retroalimentación en un sistema de tutoría, mejorando la reflexión sobre el proceso de aprendizaje y la autorregulación.
Weston Sementelli et al. (2018)	Estrategias de Comprensión y Escritura

Artículo	Herramientas Utilizadas
	Se enfocaron en el entrenamiento en estrategias de comprensión y escritura, mejorando las habilidades de escritura específicas.
VanLehn et al. (2020)	Construcción de Modelos Algebraicos, Tutoría Utilizaron la tutoría para la enseñanza de construcción de modelos algebraicos, con el objetivo de mejorar la habilidad de los estudiantes en esta área matemática.
Tacoma et al. (2020)	Retroalimentación Inteligente, Prueba de Hipótesis Implementaron retroalimentación inteligente en la prueba de hipótesis para mejorar la comprensión y el rendimiento en este tipo de evaluaciones.
Frontino de Medeiros et al. (2020)	Sistema de Tutoría Bayesian, Asistente Cognitivo Utilizaron un sistema de tutoría Bayesian con asistente cognitivo para mejorar la interacción a través de "small talk" y asistencia cognitiva.
Eryilmaz y Adabashi (2020)	Redes, Lógica Difusa, Mejora del Rendimiento Se centraron en el uso de redes y lógica difusa para mejorar el rendimiento estudiantil, explorando enfoques avanzados de análisis de datos.
De Chiusole et al. (2020)	Stat-Knowlab (Plataforma), Evaluación de Estadísticas Utilizaron la plataforma Stat-Knowlab para la enseñanza y evaluación de estadísticas, buscando mejorar la comprensión y dominio de este tema.
Gunawan et al. (2020)	Sistemas de Tutoría en Entornos de Aprendizaje Mixto Se centraron en sistemas de tutoría en entornos de aprendizaje mixto, explorando la mejora de los planes de lecciones en la educación de ciencias.
Rahim et al. (2020)	Medición y Predicción del Rendimiento a través de la función predictiva La función predictiva involucra el análisis de datos y el uso de algoritmos para anticipar el rendimiento futuro de los estudiantes en función de sus interacciones y desempeño previos en el sistema.
Furlan et al. (2021)	Procesamiento de Lenguaje Natural, Simulador de Paciente Virtual La función predictiva involucra el análisis de datos y el uso de algoritmos para anticipar el rendimiento futuro de los estudiantes en función de sus interacciones y desempeño previos en el sistema.
Grivokostopoulou et al. (2021)	Sistema Inteligente Adaptativo, Mejora del Rendimiento Utilizaron el procesamiento de lenguaje natural y un simulador de paciente virtual para mejorar el aprendizaje clínico mediante la simulación y la tutoría.
Subirats et al. (2021)	Inteligencia Artificial, Aprendizaje online en Pandemia Utilizaron inteligencia artificial para contrarrestar los efectos de la pandemia en el aprendizaje, adaptando enfoques de enseñanza a las circunstancias cambiantes.
VanLehn et al. (2023)	Sistema de Tutoría Basado en Pasos, Modelos Algebraicos Se centraron en la tutoría basada en pasos para la construcción de modelos algebraicos, con el objetivo de desarrollar habilidades en estudiantes con bajo rendimiento.

En el cuarto objetivo se analiza el impacto de los Sistemas de Tutoría Inteligente (STI) en el proceso de enseñanza-aprendizaje, evaluando tanto los efectos positivos como los desafíos asociados a su implementación. Según los datos analizados, los estudios muestran que los STI pueden tener un impacto significativo en varios aspectos educativos. Por ejemplo, Arroyo et al. (2014) destacan cómo los STI mejoran la comprensión de conceptos matemáticos al abordar aspectos cognitivos y emocionales. Ramos de Melo et al. (2014) enfocan su atención en la personalización de la experiencia educativa, lo que está relacionado con un mayor

rendimiento y participación estudiantil. Wolfe et al. (2014) subrayan cómo los STI contribuyen a una mejor comprensión del riesgo genético mediante plataformas en línea.

Además de mejorar la comprensión de conceptos complejos y habilidades específicas, como la construcción de modelos matemáticos y la resolución de problemas, se observa un aumento en la participación estudiantil. Sin embargo, también se señalan desafíos, como la necesidad de una infraestructura tecnológica adecuada, la capacitación docente y preocupaciones sobre la privacidad y seguridad de los datos de los estudiantes (Guerrero-Roldán et al., 2016; Eryilmaz & Adabashi, 2020; De Chiusole et al., 2020). Otros estudios, como McCarthy et al. (2018), exploran cómo los STI pueden fomentar la autorregulación y la reflexión metacognitiva, mientras que VanLehn et al. (2023) destacan su capacidad para cerrar brechas de aprendizaje en estudiantes con bajo rendimiento mediante tutorías basadas en pasos. En conjunto, los estudios sugieren que el éxito de los STI depende de la disposición de los estudiantes para interactuar activamente y de la adecuada integración tecnológica en el aula.

Tabla 4

Impacto de los STI

Autor	Enfoque Principal	Impacto en el Proceso de Enseñanza-Aprendizaje
Arroyo et al. (2014)	Tutoría adaptativa multimedia en matemáticas	Mejora la comprensión matemática, abordando cognición, metacognición y emoción.
Ramos de Melo et al. (2014)	Organización computacional de contenidos didácticos	Personalización de entornos de aprendizaje virtual.
Wolfe et al. (2014)	Comunicación genética de riesgo de cáncer	Aumento en la comprensión del riesgo genético a través de la web.
Guerrero-Roldán et al. (2016)	Aprendizaje de algoritmos de búsqueda	Desarrollo de habilidades en algoritmos y evaluación automática de estudiantes.
eBahçeci & Gürol (2016)	Instrucción individualizada y rendimiento académico	Potencial mejora en resultados con instrucción personalizada.
VanLehn et al. (2016)	Aprendizaje de ciencias a través de la construcción de modelos	Aumento de aprendizaje sin incrementar el tiempo dedicado.
McCarthy et al. (2018)	Funciones metacognitivas en un STI	Mejora en la reflexión sobre el proceso de aprendizaje y autorregulación.
Weston Sementelli et al. (2018)	Entrenamiento en estrategias de comprensión y escritura	Mejora en tareas de escritura con fuentes específicas.
VanLehn et al. (2020)	Enseñanza de construcción de modelos algebraicos	Habilidad mejorada en la construcción de modelos matemáticos.
Tacoma et al. (2020)	Retroalimentación inteligente en la prueba de hipótesis	Mejora en comprensión y rendimiento en pruebas de hipótesis.
Frontino de Medeiros et al. (2020)	Asistente cognitivo en interacciones de tutoría	Mejora en la interacción a través de "small talk" y asistencia cognitiva.
Eryilmaz & Adabashi (2020)	Redes y lógica difusa para rendimiento estudiantil	Potencial mejora en el rendimiento estudiantil con enfoque en redes y lógica.

Autor	Enfoque Principal	Impacto en el Proceso de Enseñanza-Aprendizaje
De Chiusole et al. (2020)	Aprendizaje y evaluación de estadísticas con teoría de Espacios de Conocimiento	Mejora en la enseñanza y evaluación de estadísticas.
Gunawan et al. (2020)	Uso de STI en entornos de aprendizaje mixto	Exploración de planes de lecciones en educación de ciencias.
Rahim et al. (2020)	Medición y predicción del rendimiento de STI	Utilización de función predictiva para evaluación.
Furlan et al. (2021)	Simulador de paciente virtual basado en procesamiento de lenguaje natural y STI	Aprendizaje clínico mejorado mediante simulación y tutoría.
Grivokostopoulou et al. (2021)	Sistema inteligente adaptativo en cursos de Economía y Negocios	Mejora en el rendimiento en cursos en línea.
Subirats et al. (2021)	IA para contrarrestar efectos de COVID-19 en aprendizaje	Utilización de IA para mitigar impacto de pandemia en la educación.
VanLehn et al. (2023)	Tutoría basada en pasos para construcción de modelos algebraicos	Desarrollo de habilidades en estudiantes con bajo rendimiento.

En relación con los resultados particulares, Bahçeci y Gürol (2016) y Guerrero-Roldán et al. (2016) evidencian diferencias significativas en el rendimiento académico en comparación con el método tradicional. Concretamente, Arroyo et al. (2014) identifica un aumento del 12% en aprobación de pruebas y 7% en niveles avanzados, y Eryilmaz & Adabashi (2020) un menor tiempo empleado en las pruebas de evaluación. Por su parte, Weston Sementelli et al. (2018) y Wolfe et al. (2014) destacan mejoras en la elaboración de ensayos y adquisición de contenidos respectivamente, y Subirats et al. (2021) detectan resultados significativos al analizar la categoría excelente y en tareas de seguimiento. Finalmente, mientras VanLehn et al. (2023) evidencian mejoras en las puntuaciones de todos los alumnos, De Chiusole et al. (2020) muestran un aumento en el dominio de habilidades, especialmente entre estudiantes con menor nivel previo de conocimientos.

4. DISCUSIÓN Y CONCLUSIONES

Esta revisión sistemática muestra un discreto progreso de la producción sobre el impacto de los STI en el proceso de enseñanza aprendizaje en la última década, especialmente asociado con el protagonismo que las TIC han alcanzado durante el periodo pandémico, aunque no se mantiene constante. La covid-19 ha impulsado significativamente la adopción acelerada de las TIC en la educación, incluyendo los STI, como una respuesta eficaz a los desafíos planteados por la enseñanza remota y la educación en línea (Arana, 2021), facilitando una experiencia educativa más personalizada y efectiva, como subrayan Castillo et al. (2021). Pero su impacto en el proceso de enseñanza-aprendizaje no sigue una progresión uniforme, como se refleja en las fluctuaciones observadas en la producción científica, hay periodos de mayor énfasis y atención, seguidos por etapas de relativa quietud investigativa. Esto puede estar influenciado por diversos factores, como la evolución de las políticas educativas, los cambios en las prioridades de investigación y las transformaciones en la percepción y adopción de las tecnologías educativas por parte de educadores, estudiantes y otras partes interesadas (Arana,

2021; Cervantes et al, 2021). La producción científica sobre STI ha experimentado un notable incremento en los últimos años, especialmente a partir de 2019, coincidiendo con el impacto del COVID-19 en la educación y el auge de la inteligencia artificial generativa, como la herramienta Chat GPT3.5. La mayoría de los estudios se concentran en países con alta inversión en tecnología educativa, como Estados Unidos, China y varios países de Europa.

Vivir como muchos tecnólogos deseaban, al 100% de la virtualidad, fue una experiencia única, inesperada en un momento concreto y con una capacitación insuficiente del profesorado, que despertó muchas inquietudes en la totalidad de la comunidad educativa, especialmente en las familias (Hernández-Prados y Alvarez- Muñoz, 2021). En palabras de Vera (2023) presentó desafíos inesperados y complejidades que no fueron anticipadas en su totalidad. La falta de interacción cara a cara, la adaptación de los contenidos para la plataforma digital y la creación de un ambiente de aprendizaje colaborativo fueron solo algunos de los obstáculos que el profesorado enfrentó de manera repentina y a menudo sin disponibilidad de recursos (Lima & Silva, 2022). Además, autores como Ponce et al, (2019) han planteado interrogantes sobre la extrapolación de los modelos educativos tradicionales al entorno virtual, señalando que no todas las dinámicas de aprendizaje son fácilmente transferibles y que se requiere una reevaluación profunda de las estrategias pedagógicas.

La eficacia de los STI en este nuevo panorama también se ha cuestionado, a pesar de haber demostrado valía en diversos contextos educativos. Los estudios revisados presentan una diversidad en términos de tamaño de muestra, etapas educativas y diseños metodológicos. Se observa una tendencia hacia la utilización de muestras grandes y heterogéneas, lo que permite una mayor generalización de los resultados. Autores como Rodríguez et al. (2020) han señalado que la personalización y adaptación de los STI pueden ser afectadas por la falta de interacción directa entre el tutor y el estudiante, lo que podría influir en la calidad de la retroalimentación y la comprensión profunda de las necesidades del estudiante. Asimismo, la capacitación insuficiente del profesorado para aprovechar al máximo las capacidades de los STI en el entorno virtual es un factor que no debe pasarse por alto (Johnston et al., 2018). La tecnología avanzada de los STI exige una comprensión profunda por parte de los profesores para su integración efectiva en la pedagogía, y la falta de esta preparación puede limitar su potencial (Ponce y Biset, 2019). Por tanto, se requiere una preparación docente más exhaustiva y de una reflexión profunda sobre la aplicabilidad y limitaciones de los STI en este nuevo paradigma educativo tendente hacia la virtualidad, así como una mayor indagación teórica.

Al respecto, el presente estudio permite concluir que los STI se presentan como elementos cruciales en el proceso educativo, con el potencial de transformar y mejorar diversas etapas y contextos de aprendizaje. Los STI se han implementado exitosamente en una variedad de entornos educativos, desde escuelas primarias hasta universidades. Estos sistemas han demostrado ser efectivos en mejorar el rendimiento académico y en fomentar la autorregulación y la motivación de los estudiantes. Por ejemplo, Çetin et al. (2023) investigaron el efecto de un sistema de tutoría inteligente gamificado en las habilidades de resolución de problemas, encontrando mejoras significativas. Además, Rebolledo-Mendez et al. (2022) exploraron el comportamiento meta-afectivo dentro de un sistema de tutoría inteligente para matemáticas, destacando su impacto positivo en el aprendizaje. Al explorar las múltiples dimensiones de los STI y su influencia en el aprendizaje, este estudio ha profundizado en cómo la tecnología puede desempeñar un papel catalizador en la educación contemporánea. Sin

embargo, también destaca la importancia de un enfoque crítico-reflexivo en la implementación de los STI y en la consideración del entorno y variables contextuales que pueden impactar en la dinámica de enseñanza-aprendizaje. Las lecciones extraídas de investigaciones anteriores actúan como guías valiosas para abordar de manera efectiva y eficiente el desafío de implementar los STI de manera beneficiosa.

En medio de esta evolución constante, es esencial abordar desafíos como la generalización y reutilización de los STI. Una implicación significativa es el potencial de los STI en las etapas iniciales de la educación. La muestra analizada sugiere que los STI podrían desempeñar un papel crucial en la formación de habilidades fundamentales desde una edad temprana, sentando bases sólidas para un aprendizaje continuo a lo largo de la vida. Por lo tanto, es crucial que futuras investigaciones se centren en la creación y evaluación de STI diseñados específicamente para niveles de Educación Infantil y Primaria. Esto podría abrir nuevas vías para enriquecer la educación desde sus primeros momentos y establecer un fundamento sólido para el desarrollo educativo a largo plazo.

Finalmente, aunque los avances actuales son prometedores, es necesario avanzar hacia la creación de dominios de aprendizaje más amplios y aplicables. Superar la limitación de enfoque en áreas particulares permitirá que los STI se conviertan en recursos educativos más versátiles y efectivos, beneficiando a un espectro más amplio de estudiantes y disciplinas. Los estudios destacan que los STI no solo mejoran el rendimiento académico, sino que también personalizan el aprendizaje, adaptando la dificultad a las capacidades individuales de los estudiantes y apoyando la autorregulación. Sin embargo, es necesario un enfoque interdisciplinar para maximizar su eficacia y abordar las limitaciones tecnológicas y pedagógicas. Ni y Cheung (2023) analizaron la intención de los estudiantes de secundaria de continuar utilizando un STI potenciado por IA para el aprendizaje del inglés, destacando la importancia de la personalización y el apoyo continuo. Asimismo, Heck et al (2022) discutieron la generación automática de ejercicios para apoyar la macro-adaptabilidad en los sistemas de tutoría de idiomas, subrayando la capacidad de estos sistemas para ajustarse a las necesidades individuales. En esta dirección, este estudio también abre nuevos interrogantes, cómo se puede fomentar una mentalidad abierta hacia la tecnología en la educación y cómo se pueden proporcionar a los docentes la formación y el apoyo necesarios para integrar de manera efectiva las tutorías inteligentes en distintos contextos disciplinares, e invita a reflexionar sobre cómo se pueden abordar las preocupaciones y percepciones que podrían estar contribuyendo al hermetismo docente. Explorar estrategias para sensibilizar a los educadores sobre los beneficios de los STI y cómo estos pueden enriquecer la enseñanza y el aprendizaje en diversas áreas puede ser crucial para fomentar una mayor adopción y apertura. En última instancia, este patrón señala la importancia de continuar investigando y promoviendo la implementación STI en una variedad de materias, enfocándose no solo en los beneficios tecnológicos, sino también en abordar las percepciones y actitudes que pueden estar frenando su adopción en múltiples disciplinas.

5. REFERENCIAS

- Achard, I. (2020). *¿ Nuevo rol o nueva identidad docente en la era digital?*. Educación y Tecnología.
- Alahi, M. E. E., Sukkuea, A., Tina, F. W., Nag, A., Kurdthongmee, W., Suwannarat, K., & Mukhopadhyay, S. C. (2023). Integration of IoT-enabled technologies and artificial intelligence (AI) for smart city scenario: recent advancements and future trends. *Sensors*, 23(11), 5206. <https://doi.org/10.3390/s23115206>
- Alemán-Díaz, A. Y. (2023). *Motivations guiding public research funding in science, technology and innovation (STI) policy: a synthesis*. In Handbook of Public Funding of Research (pp. 38-54). Edward Elgar Publishing.
- Arana, C. (2021). Inteligencia Artificial Aplicada a la Educación: Logros, Tendencias y Perspectivas. INNOVA UNTREF. *Revista Argentina de Ciencia y Tecnología*, 1-22. <https://revistas.untref.edu.ar/index.php/innova/article/download/1107/917>
- Batista, L. H. L., & Maffu, N. F. (2007). Web y tutoriales como herramientas en el proceso de enseñanza y aprendizaje. *Revista Información Científica*, 54(2). <https://pesquisa.bvsalud.org/portal/resource/pt/cum-37552>
- Cao, J., Yang, T., Lai, I. K. W., & Wu, J. (2021). Is online education more welcomed during COVID-19? An empirical study of social impact theory on online tutoring platforms. *The International Journal of Electrical Engineering & Education*, 0020720920984001. <https://doi.org/10.1177/0020720920984001>
- Carini, R. M., Kuh, G. D., & Klein, S. P. (2006). Student engagement and student learning: Testing the linkages. *Research in higher education*, 47, 1-32. <https://link.springer.com/article/10.1007/s11162-005-8150-9>
- CASTILLO AVILA, A. A.; GONZÁLEZ CALLEROS, J. M.; GUERRERO GARCÍA, J.(2021) Revisión sistemática de la literatura sobre los sistemas tutores afectivos: 2001-2020. *Revista Brasileira de Informática na Educação, [S. l.]*, v. 29, p. 928–956, 2021. DOI: 10.5753/rbie.2021.29.0.928.
- Cervantes-Pérez, F., Navarro-Perales, J., Franzoni-Velázquez, A.L. y de la Fuente Valentín, L. (2021). Bayesian Knowledge Tracing for Navigation through Marzano's Taxonomy. *International Journal of Interactive Multimedia and Artificial Intelligence*, 6(6), 234-239. <https://dx.doi.org/10.9781/ijimai.2021.05.006>
- Çetin, I., Erumit, A. K., Nabiyev, V., Karal, H., Kosa, T., & Kokoc, M. (2023). The Effect of Gamified Adaptive Intelligent Tutoring System Artibos on Problem-Solving Skills. *Participatory Educational Research*, 10(1), 344–374. <https://doi.org/10.17275/per.23.19.10.1>
- De La Rosa Valdiviezo, A., Toro Girón, K., Jaén Armijos, K., & Espinoza Freire, E. E. (2019). El proceso de enseñanza-aprendizaje en las ciencias naturales: las estrategias didácticas como alternativa. *Revista Científica Agroecosistemas*, 7(1), 58-62. <https://aes.ucf.edu.cu/index.php/aes/article/view/243>

- Incio Flores, F. A., Capuñay Sanchez, D. L. ., Estela Urbina, R. O. ., Valles Coral, M. Ángel ., Vergara Medrano, S. E. ., & Elera Gonzales, D. G. . (2021). Inteligencia artificial en educación: una revisión de la literatura en revistas científicas internacionales. *Apuntes Universitarios*, 12(1), 353–372. <https://doi.org/10.17162/au.v12i1.974>
- García-Peña, V. R., Mora-Marcillo, A. B., & Ávila-Ramírez, J. A. (2020). La inteligencia artificial en la educación. *Domino De Las Ciencias*, 6(3), 648–666. <https://doi.org/10.23857/dc.v6i3.1421>
- García-Peñalvo, F.J. (2023). La percepción de la Inteligencia Artificial en contextos educativos tras el lanzamiento de ChatGPT: Disrupción o Pánico. *Education in the Knowledge Society*, 24. <https://doi.org/10.14201/eks.31279>
- González-Palacios, A., & Avelino-Rubio, I. (2016). Tutoría: una revisión conceptual. *Revista de Educación y desarrollo*, 38, 57-68. http://148.202.248.26/sites/default/files/38_gonzalez_palacios.pdf
- Gros Salvat, B., y Cano García, E. (2021). Procesos de feedback para fomentar la autorregulación con soporte tecnológico en la educación superior: Revisión sistemática. *RIED. Revista Iberoamericana de Educación a Distancia*, 24(2), 107-125. <https://doi.org/10.5944/ried.24.2.28886>
- Gupta, Sambhav and Chen, Yu (2022). Supporting Inclusive Learning Using Chatbots? A Chatbot-Led Interview Study. *Journal of Information Systems Education*, 33(1), 98-108. <https://aisel.aisnet.org/jise/vol33/iss1/11>
- Heck, T., Meurers, D., & Nuxoll, F. (2022). Automatic Exercise Generation to Support Macro-Adaptivity in Intelligent Language Tutoring Systems. *Research-Publishing.Net*.
- Hernández-Prados, M. Á. y Álvarez-Muñoz, J.S. (2021). Familia-escuela-COVID19. Una triada emergente. En Hernández-Prados, M. Á. y Belmonte, M.L. (Coord.) *La nueva normalidad educativa, Educando en tiempos de pandemia* (26-38). Dykinson
- Hockenberry, M. J., & Warschauer, M. (2016). Laptops and fourth-grade literacy: Assisting the jump over the fourth-grade slump. *Journal of Research on Technology in Education*, 48(3), 160-175.
- Johnson, R. B., y Christensen, L. B. (2019). *Educational research: Quantitative, qualitative, and mixed methods approaches*. SAGE Publications
- Johnston, J. H., Burke, C. S., Milham, L. A., Ross, W. M., & Salas, E. (2018). Challenges and propositions for developing effective team training with adaptive tutors. In *Building intelligent tutoring systems for teams: What matters* (pp. 75-97). Emerald Publishing Limited.
- Lima Júnior, A. B. D., & Silva, L. T. G. (2022). Os sistemas tutores inteligentes e a adaptação do ensino aos perfis de aprendizagem do usuário. *ETD Educação Temática Digital*, 24(3), 618-632. DOI 10.20396/etd.v24i3.8663707

- Martín-Marchante, B.(2022).TIC e inteligencia artificial en la revisión del proceso de escritura: su uso en las universidades públicas valencianas.*Research in Education and Learning Innovation Archives*, 28,16--31.10.7203/realia.28.2062
- Mathew, A.N., Rohini, V. & Paulose, J. (2021). NLP-based personal learning assistant for school education. *International Journal of Electrical and Computer Engineering*, 11(5), 4522-4530. <http://doi.org/10.11591/ijece.v11i5.pp4522-4530>
- Moreira, M. A. (2004). Máquinas que enseñan. Una revisión de los métodos de enseñanza aprendizaje con ordenadores. *Bordón. Revista de pedagogía*, 56(3), 483-491.efaidnbnmnnibpcajpcglcfindmkaj/http://manarea.webs.ull.es/articulos/art12_BORDON.pdf
- Moreno Padilla, R. D. (2019). La llegada de la inteligencia artificial a la educación. *Revista De Investigación En Tecnologías De La Información*, 7(14), 260–270. <https://doi.org/10.36825/RITI.07.14.022>
- Ni, A., & Cheung, A. (2023). Understanding Secondary Students' Continuance Intention to Adopt AI-Powered Intelligent Tutoring System for English Learning. *Education and Information Technologies*, 28(3), 3191–3216. <https://doi.org/10.1007/s10639-022-11305-z>
- Ocaña-Fernández, Y., Valenzuela-Fernández, L.A. y Garro-Aburto, L.L. (2019). Inteligencia artificial y sus implicaciones en la educación superior. *Propósitos y Representaciones*, 7(2), 536-568. <http://dx.doi.org/10.20511/pyr2019.v7n2.274>
- Organización de Naciones Unidas para la Educación, la Ciencia y la Cultura. (2019). *Educación Inclusiva*.UNESCO
- Osorio, L., Vidanovic, A. y Finol, M. (2021). Elementos del proceso de enseñanza aprendizaje y su interacción en el ámbito educativo. *Qualitas*, 23(23). <https://doi.org/10.55867/qual23.01>
- Pezzini, M. C (2022). Tutores inteligentes en la enseñanza: una revisión y análisis en la educación secundaria. In *XXVIII Congreso Argentino de Ciencias de la Computación (CACIC)(La Rioja, 3 al 6 de octubre de 2022)*.
- Pincay Ponce, J. I., Pintado Zumba, P. F., & Biset, J. C. (2019). ANÁLISIS DE IMPLEMENTACIONES DE SISTEMAS TUTORES INTELIGENTES Y AFECTIVOS. REVISIÓN SISTEMÁTICA. *REFCaIE: Revista Electrónica Formación Y Calidad Educativa*. ISSN 1390-9010, 7(2), 218–234. <https://refcale.uleam.edu.ec/index.php/refcale/article/view/3019>
- Ponce, P., Pintado, P., y Biset, J. (2019). Análisis de implementaciones de sistemas tutores inteligentes y afectivos. Revisión sistemática. *REFCaIE: Revista Electrónica Formación y Calidad Educativa*, 7(2), 218-234. <https://observatorioturisticobahia.uleam.edu.ec/index.php/refcale/article/view/3019>
- Ramírez-Montoya, M., & Lugo-Ocando, J. (2020). Systematic review of mixed methods in the framework of educational innovation. [Revisión sistemática de métodos mixtos en el

- marco de la innovación educativa]. *Comunicar*, 65, 9-20. <https://doi.org/10.3916/C65-2020-01>
- Rebolledo-Mendez, G., Huerta-Pacheco, N. S., Baker, R. S., & du Boulay, B. (2022). Meta-Affective Behaviour within an Intelligent Tutoring System for Mathematics. *International Journal of Artificial Intelligence in Education*, 32(1), 174–195. <https://doi.org/10.1007/s40593-021-00247-1>
- Rodríguez Málaga, L; Rodríguez Pérez, C; Fidalgo Redondo, R (2019) NUEVOS ENTORNOS DE APRENDIZAJE PARA LA ESCRITURA: SISTEMAS DE TUTORÍA INTELIGENTE *Papeles del Psicólogo*,40,(2) <https://doi.org/10.23923/pap.psicol2019.2895>
- Sanz Ponce, J. R., & López Luján, E. (2021). Consecuencias pedagógicas entre el alumnado de enseñanza básica derivadas de la COVID-19. Una reflexión en torno a los grandes olvidados de la pandemia. *Teoría de La Educación. Revista Interuniversitaria*, 33(2). <https://doi.org/10.14201/teri.25471>
- Selwyn, N. (2020). *¿Deberían los robots sustituir al profesorado?: La IA y el futuro de la educación*. Ediciones Morata.
- Tirocchi, S. (2022). *Educación y plataformas digitales. Implicaciones para los estudiantes. In Redes sociales y ciudadanía: ciberculturas para el aprendizaje*. (pp. 159-166). Grupo Comunicar Ediciones.
- Troncoso, M.O., Dueñas, Y.K. y Verdecia, E. (2023). Inteligencia artificial y educación: nuevas relaciones en un mundo interconectado. *Estudios del Desarrollo Social: Cuba y América Latina*, 11(2), 312-328. <https://revistas.uh.cu/revflaco/article/view/4815>
- Urquilla, A. (2022). Un viaje hacia la inteligencia artificial en la educación. *Realidad y Reflexión*, 2022 Año. 22, N° 56, p 121-136. <http://ri.ufg.edu.sv/jspui/handle/11592/9810>
- Urrútia, G. y Bonfill, X. (2010). Declaración PRISMA: Una propuesta para mejorar la publicación de revisiones sistemáticas y metaanálisis. *Medicina clínica*, 135(11), 507-511. <https://doi.org/doi:10.1016/j.medcli.2010.01.015>
- VanLehn, K., Milner, F., Banerjee, C. & Wetzel, J. (2023). A Step-Based Tutoring System to Teach Underachieving Students How to Construct Algebraic Models. *International Journal of Artificial Intelligence in Education*. <https://doi.org/10.1007/s40593-023-00328-3>
- Vera, F. (2023). Integración de la Inteligencia Artificial en la Educación superior: Desafíos y oportunidades. *Transformar*, 4(1), 17–34. <https://www.revistatransformar.cl/index.php/transformar/article/view/84>
- Webster, C. (1997). *Adaptive depression, affective computing, and intelligent processing. In Intelligent Processing Systems, 1997. ICIPS'97. 1997 IEEE International Conference on* (Vol. 2, pp. 1181–1184). IEEE

- Xiangjie, Q., Zhiliang, W., Jun, Y., & Xiuyan, M. (2006, August). An affective intelligent tutoring system based on artificial psychology. In *First International Conference on Innovative Computing, Information and Control-Volume I (ICICIC'06)* (Vol. 3, pp. 402-405). IEEE.
- Zuo, M., Hu, Y., Luo, H., Ouyang, H., & Zhang, Y. (2022). K-12 students' online learning motivation in China: An integrated model based on community of inquiry and technology acceptance theory. *Education and Information Technologies*, 27(4), 4599-4620. <https://doi.org/10.1007/s10639-021-10791-x>

Para citar este artículo:

Carbonell Bernal, N., y Hernández Prados, M. Ángeles. Impacto de los sistemas de tutoría inteligente. Una revisión sistemática. *Eduotec, Revista Electrónica de Tecnología Educativa*, (89), 121-143. <https://doi.org/10.21556/edutec.2024.89.3025>

Anexo I. Breve descripción de los artículos de investigación seleccionados para el estudio.

Autores	Título	Contenido
Arroyo et al. (2014)	A Multimedia Adaptive Tutoring System for Mathematics that Addresses Cognition, Metacognition and Affect	El artículo se centra en un sistema de tutoría adaptativo multimedia diseñado para enseñar matemáticas y abordar aspectos relacionados con la cognición, la metacognición y la emoción en el proceso de aprendizaje.
Ramos de Melo et al.(2014)	Computational organization of didactic contents for personalized virtual learning environments	El artículo se centra en la organización computacional de contenidos didácticos para entornos de aprendizaje virtual personalizados.
Wolfe et al. (2014)	Efficacy of a Web-Based Intelligent Tutoring System for Communicating Genetic Risk of Breast Cancer: A Fuzzy-Trace Theory Approach	El artículo se enfoca en la eficacia de un sistema de tutoría inteligente basado en la web para comunicar el riesgo genético de cáncer de mama. El enfoque del estudio utiliza la teoría del trazo difuso (Fuzzy-Trace Theory) para abordar cómo se puede mejorar la comunicación y comprensión de la información sobre el riesgo genético de cáncer de mama a través de una plataforma en línea.
Guerrero-Roldán et al. (2016)	An Educational System for Learning Search Algorithms and Automatically Assessing Student	El artículo se centra en el desarrollo de un sistema educativo destinado a enseñar algoritmos de búsqueda y a realizar evaluaciones automáticas de los estudiantes en este campo.
eBahçeci & Gürol (2016)	The Effect of Individualized Instruction System on the Academic Achievement Scores of Students	El artículo busca comprender si la implementación de un sistema de instrucción individualizada puede conducir a mejoras en los resultados académicos en comparación con métodos de instrucción más tradicionales y generalizados.
VanLehn et al. (2016)	Learning Science by Constructing Models: Can Dragoon Increase Learning without Increasing the Time Required?	El artículo explora si el uso de un sistema llamado "Dragoon", puede aumentar el aprendizaje de los estudiantes en el ámbito de las ciencias sin requerir más tiempo del que normalmente se destina a la enseñanza.
McCarthy et al. (2018)	Metacognitive Overload: Positive and Negative Effects of Metacognitive Prompts in an Intelligent Tutoring System	El artículo se centra en examinar los efectos positivos y negativos de las funciones metacognitivas en un Sistema de Tutoría Inteligente. Estas señales metacognitivas pueden ser instrucciones o recordatorios diseñados para ayudar a los estudiantes a reflexionar sobre su propio proceso de aprendizaje y mejorar su autorregulación.


Autores	Título	Contenido
Weston Sementelli et al. (2018)	Comprehension and Writing Strategy Training Improves Performance on Content-Specific Source Based Writing Tasks	El artículo está relacionado con la investigación sobre la mejora del rendimiento en tareas de escritura basadas en fuentes específicas de contenido a través del entrenamiento en estrategias de comprensión y escritura. Explora cómo la formación en estrategias de comprensión y escritura puede impactar positivamente en la capacidad de los estudiantes para realizar tareas de escritura que requieren el uso de fuentes específicas.
VanLehn et al. (2020)	Teaching Algebraic Model Construction: A Tutoring System, Lessons Learned and an Evaluation	El artículo detalla cómo se desarrolló y evaluó un sistema de tutoría diseñado para ayudar a los estudiantes a aprender y mejorar sus habilidades en la construcción de modelos algebraicos.
Tacoma et al. (2020)	Intelligent Feedback on Hypothesis Testing	El artículo investiga cómo se puede implementar la retroalimentación inteligente en el contexto de la enseñanza y el aprendizaje de la prueba de hipótesis, con el objetivo de mejorar la comprensión y el rendimiento de los estudiantes.
Frontino de Medeiros et al. (2020)	Development of an Intelligent Tutoring System Using Bayesian	El artículo aborda el desarrollo de un sistema de tutoría inteligente utilizando el enfoque Bayesian, así como del asistente cognitivo que utiliza "small talk" (conversación casual) en las interacciones de tutoría.
Eryilmaz & Adabashi (2020)	Networks and Fuzzy Logic for a Higher Student Academic Performance	El artículo explora cómo se pueden utilizar conceptos de redes y lógica difusa para mejorar la efectividad de sistemas de soporte o intervenciones en el ámbito educativo con el objetivo de elevar el rendimiento estudiantil.
De Chiusole et al. (2020)	Stat-Knowlab. Assessment and Learning of Statistics with Competence-based Knowledge Space Theory	El artículo se enfoca en una plataforma llamada "Stat-Knowlab" que aborda la evaluación y el aprendizaje de estadísticas utilizando la teoría de Espacios de Conocimiento basados en Competencias.
Gunawan et al. (2020)	Exploring Science Teachers' Lesson Plans by the Implementation of Intelligent Tutoring Systems in Blended Learning Environments	El artículo se centra en explorar los planes de lecciones de profesores de ciencias a través de la implementación de Sistemas de Tutoría Inteligente en entornos de aprendizaje mixto (blended learning).
Rahim et al. (2020)	Intelligent Tutoring Systems' Measurement and Prediction of Students' Performance Using Predictive Function	El artículo se centra en la medición y predicción del rendimiento de los estudiantes en Sistemas de Tutoría Inteligente utilizando una función predictiva.

Autores	Título	Contenido
Furlan et al. (2021)	A Natural Language Processing- Based Virtual Patient Simulator and Intelligent Tutoring System for the Clinical Diagnostic Process: Simulator Development and Case Study Performance	El artículo se centra en el desarrollo de un simulador de paciente virtual basado en procesamiento de lenguaje natural y en un Sistema de Tutoría Inteligente para el proceso de diagnóstico clínico.
Grivokostopoulou et al. (2021)	Experiences in the use of an adaptive intelligent system to enhance online learner's performance: a case study in Economics and Business courses	El artículo se enfoca en las experiencias relacionadas con el uso de un sistema inteligente adaptativo para mejorar el rendimiento de los estudiantes en cursos en línea en el ámbito de Economía y Negocios.
Subirats et al. (2021)	Artificial Intelligence to Counterweight the Effect of COVID-19 on Learning in a Sustainable Environment	El artículo se centra en el uso de la Inteligencia Artificial para contrarrestar los efectos de la pandemia de COVID-19 en el aprendizaje en un entorno sostenible.
VanLehn et al. (2023)	A Step-Based Tutoring System to Teach Underachieving Students How to Construct Algebraic Models	El artículo aborda el desarrollo y la evaluación de un sistema de tutoría basado en pasos para enseñar a estudiantes con bajo rendimiento cómo construir modelos algebraicos. Este enfoque está diseñado para abordar las dificultades específicas que enfrentan los estudiantes que tienen dificultades para construir modelos matemáticos y ecuaciones algebraicas.



Integración de la tecnología y la pedagogía en los sistemas de tutoría inteligente

Integration of technology and pedagogy in intelligent tutoring systems

 Geovanny Francisco Ruiz Muñoz; geovanny.ruizm@ug.edu.ec

Universidad de Guayaquil (Ecuador)

Resumen

Los Sistemas de Tutoría Inteligente (STI) han demostrado su potencial para mejorar el proceso de enseñanza-aprendizaje al brindar instrucción personalizada y adaptativa. Sin embargo, la integración efectiva de la tecnología y la pedagogía en estos sistemas sigue siendo un desafío. Este estudio tuvo como objetivo analizar los enfoques y estrategias más efectivos para lograr esta integración. Se empleó una metodología mixta que combinó análisis bibliométricos, minería de textos, análisis de contenido y entrevistas a expertos. Los resultados destacan la importancia de utilizar modelos de estudiante precisos, incorporar enfoques constructivistas y de aprendizaje multimedia, aplicar retroalimentación adaptativa y andamiaje, y diseñar interfaces envolventes. También se identificaron desafíos relacionados con la accesibilidad, escalabilidad y consideraciones éticas. Se concluye que una combinación óptima de avances tecnológicos, principios pedagógicos sólidos y una comprensión profunda de las necesidades individuales de los estudiantes es crucial para el éxito de los STI.

Palabras clave: Sistemas de Tutoría Inteligente, integración tecnología-pedagogía, enfoques constructivistas, aprendizaje multimedia.

Abstract

Intelligent Tutoring Systems (ITS) have demonstrated their potential to enhance the teaching-learning process by providing personalized and adaptive instruction. However, effectively integrating technology and pedagogy within these systems remains a challenge. This study aimed to analyze the most effective approaches and strategies for achieving this integration. A mixed-method methodology was employed, combining bibliometric analysis, text mining, content analysis, and expert interviews. The results highlight the importance of using accurate student models, incorporating constructivist and multimedia learning approaches, applying adaptive feedback and scaffolding, and designing engaging interfaces. Challenges related to accessibility, scalability, and ethical considerations were also identified. The study concludes that an optimal combination of technological advancements, sound pedagogical principles, and a deep understanding of individual student needs is crucial for the success of ITS in improving learning outcomes.

Keywords: Intelligent Tutoring Systems, technology-pedagogy integration, constructivist approaches, multimedia learning.



1. INTRODUCCIÓN

Los Sistemas de Tutoría Inteligente (STI) son una aplicación de la Inteligencia Artificial en el ámbito educativo, cuyo objetivo principal es brindar una instrucción personalizada y adaptativa a los estudiantes (Essa, 2016; VanLehn, 2011). Los STI se basan en una combinación de principios pedagógicos, teorías del aprendizaje y técnicas de inteligencia artificial para crear entornos de aprendizaje interactivos y efectivos.

Estos sistemas han ganado una creciente relevancia en el ámbito educativo, ya que ofrecen una oportunidad única para personalizar el aprendizaje y adaptarse a las necesidades individuales de los estudiantes. Estos sistemas tienen el potencial de mejorar significativamente el proceso de enseñanza-aprendizaje al brindar una instrucción personalizada, retroalimentación inmediata y una experiencia de aprendizaje más atractiva y motivadora (Essa, 2016; VanLehn, 2011).

Sin embargo, a pesar de sus beneficios, la integración efectiva de la tecnología y la pedagogía en los STI ha sido un desafío persistente. Numerosos investigadores han explorado este tema, destacando la importancia de combinar de manera óptima los avances tecnológicos con principios pedagógicos sólidos (Koedinger et al., 2003; Mitrovic & Ohlsson, 1999; Roll et al., 2007).

El objetivo principal de esta investigación es analizar y evaluar las estrategias más efectivas para integrar la tecnología y la pedagogía en los STI, con el fin de mejorar su eficacia en el proceso de enseñanza-aprendizaje. Este estudio se centra en examinar los componentes clave de los STI: el modelo del estudiante, el modelo del dominio y el modelo pedagógico, y cómo su interacción puede optimizarse para crear experiencias de aprendizaje más personalizadas y efectivas.

Uno de los componentes clave de los STI es el modelo del estudiante, que representa el conocimiento, habilidades y características individuales del alumno (Nkambou et al., 2010). Este modelo se actualiza constantemente a medida que el estudiante interactúa con el sistema, permitiendo así una adaptación dinámica de la instrucción y la retroalimentación (Essa, 2016).

Otro elemento fundamental es el modelo del dominio, que representa el conocimiento y las competencias que se deben enseñar (Mitrovic & Ohlsson, 1999). Este modelo se basa en una representación formal del conocimiento del dominio, lo que permite al sistema evaluar las respuestas del estudiante y proporcionar retroalimentación precisa (Koedinger et al., 2003).

El modelo pedagógico es el componente que integra los principios de enseñanza y aprendizaje en el STI (Nkambou et al., 2010). Este modelo determina cómo se presentará el contenido, qué estrategias de enseñanza se utilizarán y cómo se proporcionará la retroalimentación (Soh et al., 2017).

Este estudio contribuirá significativamente al campo de la educación asistida por tecnología, proporcionando aportes para el desarrollo de la próxima generación de STI que puedan adaptarse de manera más eficaz a las necesidades individuales de los estudiantes y mejorar los resultados de aprendizaje.

La integración efectiva de la tecnología y la pedagogía en los STI implica la combinación óptima de estos componentes para crear una experiencia de aprendizaje personalizada y efectiva. A continuación, se analizan los principios teóricos y los enfoques que han demostrado ser efectivos en esta integración.

Teorías del Aprendizaje y Enfoques Pedagógicos

Uno de los enfoques teóricos más influyentes en el diseño de STI es el constructivismo, que enfatiza la construcción activa del conocimiento por parte del estudiante (Jonassen, 1991; Piaget, 1970). Los STI basados en el constructivismo promueven el aprendizaje a través de la exploración, la resolución de problemas y la interacción con el entorno (Graesser et al., 2018).

Otro enfoque clave es el aprendizaje por descubrimiento guiado, que combina la instrucción directa con la exploración y el descubrimiento guiado por el sistema (Mayer, 2004). Este enfoque ha demostrado ser efectivo para fomentar la comprensión profunda y la transferencia del conocimiento (Schwartz et al., 2011).

La teoría del aprendizaje situado también ha influido en el diseño de STI, enfatizando la importancia del contexto y la autenticidad de las actividades de aprendizaje (Brown et al., 1989). Los STI basados en esta teoría buscan crear entornos de aprendizaje inmersos en situaciones reales o simuladas (Koedinger & Corbett, 2006).

La teoría del aprendizaje multimedia de Mayer (2009) también ha sido ampliamente utilizada en el diseño de STI, ya que proporciona principios para la presentación efectiva de información mediante la integración de texto, gráficos, animaciones y otros medios.

Además, los enfoques basados en la autorregulación del aprendizaje han ganado relevancia en los STI, ya que promueven el desarrollo de habilidades metacognitivas y estrategias de aprendizaje efectivas (Azevedo & Alevan, 2013; Graesser et al., 2018).

Modelos de Estudiante y Adaptación

Un aspecto clave en la integración de la tecnología y la pedagogía en los STI es el desarrollo de modelos de estudiante precisos y adaptables. Estos modelos representan el conocimiento, las habilidades y las características individuales de cada estudiante, lo que permite al sistema ajustar la instrucción y la retroalimentación de manera personalizada (Nkambou et al., 2010).

Existen diferentes enfoques para modelar al estudiante, como los modelos basados en restricciones (Mitrovic & Ohlsson, 1999), los modelos basados en superposición (Koedinger et al., 2003) y los modelos bayesianos (Conati et al., 2002). Estos enfoques utilizan técnicas de inteligencia artificial, como la representación del conocimiento, el razonamiento basado en casos y las redes bayesianas, para inferir el estado de conocimiento del estudiante y adaptar la instrucción en consecuencia.

Otro aspecto importante es la adaptación de los STI a los estilos de aprendizaje y las preferencias de los estudiantes. Algunos enfoques utilizan cuestionarios o técnicas de minería de datos para identificar los estilos de aprendizaje y ajustar la presentación del contenido, las estrategias de enseñanza y la interfaz de usuario en consecuencia (Essa, 2016; Tsai et al., 2020).

Retroalimentación y Andamiaje

La retroalimentación efectiva es fundamental para el aprendizaje en los STI. La investigación ha demostrado que la retroalimentación oportuna, específica y basada en el progreso del estudiante mejora significativamente el rendimiento y la motivación (Shute, 2008; VanLehn, 2011).

Existen diferentes enfoques para proporcionar retroalimentación en los STI, como la retroalimentación basada en errores, la retroalimentación elaborada y la retroalimentación adaptativa (Essa, 2016; Roll et al., 2007). Estos enfoques se basan en principios pedagógicos sólidos y utilizan técnicas de inteligencia artificial para generar retroalimentación personalizada y relevante.

El andamiaje, o la provisión de apoyo temporal y ajustable, también es un componente clave en los STI. El andamiaje puede tomar diversas formas, como pistas, sugerencias, ejemplos o preguntas orientadoras, y su objetivo es guiar al estudiante a través del proceso de aprendizaje y fomentar su autonomía gradual (Graesser et al., 2018; Koedinger & Alevan, 2007).

Interacción y Agentes Pedagógicos

La interacción efectiva entre el estudiante y el sistema es crucial para el éxito de los STI. La investigación ha demostrado que las interfaces de usuario atractivas y fáciles de usar, que incorporan elementos multimedia y de gamificación, pueden mejorar la motivación y el compromiso de los estudiantes (Mayer, 2009; Schroeder & Adesope, 2014).

Uno de los enfoques más prometedores en este ámbito es el uso de agentes pedagógicos virtuales, que son representaciones antropomórficas o animadas que interactúan con los estudiantes y brindan instrucción y retroalimentación (Graesser et al., 2005; Schroeder & Adesope, 2014). Estos agentes pueden mejorar la experiencia de aprendizaje al proporcionar una interacción más natural y envolvente.

Evaluación y Métricas de Aprendizaje

La evaluación del aprendizaje y la efectividad de los STI es otro aspecto crucial en la integración de la tecnología y la pedagogía. Existen diferentes enfoques para evaluar el progreso de los estudiantes, como las pruebas de conocimiento, la resolución de problemas y la evaluación de habilidades de orden superior (Essa, 2016; Shute, 2008).

Además, es importante evaluar la efectividad general de los STI en comparación con otros métodos de enseñanza, como la instrucción en el aula tradicional o la tutoría humana (VanLehn, 2011). Estas evaluaciones pueden utilizar métricas como el rendimiento académico, la motivación, la autorregulación del aprendizaje y la transferencia de conocimientos.

Accesibilidad y Escalabilidad

A medida que los STI se vuelven más sofisticados y personalizados, es necesario abordar cuestiones relacionadas con la accesibilidad y la escalabilidad (Essa, 2016; Nkambou et al., 2010). Los STI deben ser accesibles para diferentes grupos de estudiantes, incluyendo aquellos

con discapacidades o necesidades especiales, y deben poder adaptarse a diferentes contextos educativos.

La escalabilidad también es un desafío, ya que los STI altamente personalizados requieren un gran esfuerzo de desarrollo y mantenimiento (Essa, 2016). Se han explorado enfoques como la inteligencia artificial conversacional, la generación automática de contenido y la minería de datos para abordar este desafío (Arroyo et al., 2014; Graesser et al., 2018).

Colaboración y Aprendizaje Social

Aunque los STI se han centrado tradicionalmente en la instrucción individual, existe un creciente interés en incorporar aspectos de colaboración y aprendizaje social (Nkambou et al., 2010; Soller et al., 2005). La investigación ha demostrado que el aprendizaje colaborativo puede mejorar la comprensión, la motivación y el desarrollo de habilidades sociales.

Algunos enfoques para integrar la colaboración en los STI incluyen el uso de entornos colaborativos virtuales, la incorporación de agentes pedagógicos colaborativos y el análisis de las interacciones entre los estudiantes para proporcionar retroalimentación y apoyo (Magnisalis et al., 2011; Soller et al., 2005).

Ética y Privacidad

A medida que los STI se vuelven más sofisticados y personalizados, surgen preocupaciones éticas y de privacidad (Essa, 2016; Nkambou et al., 2010). Es importante garantizar la privacidad de los datos de los estudiantes y evitar el uso indebido de la información personal.

Además, los STI deben estar diseñados de manera justa y no discriminatoria, evitando sesgos potenciales en los algoritmos o en el contenido (Arroyo et al., 2014; Essa, 2016). También es crucial fomentar la transparencia y la confianza en estos sistemas, permitiendo a los estudiantes y educadores comprender cómo funcionan y cómo se toman las decisiones.

2. MÉTODO

Este estudio empleó un enfoque de métodos mixtos, combinando una revisión sistemática de la literatura con entrevistas a expertos para analizar las estrategias de integración de tecnología y pedagogía en los Sistemas de Tutoría Inteligente (STI). Siguiendo el modelo PRISMA, se realizó una búsqueda exhaustiva en las bases de datos Web of Science, Scopus, ERIC y IEEE Xplore, utilizando términos clave como "intelligent tutoring systems", "technology integration" y "pedagogical approaches", junto con sus variantes y operadores booleanos. Se incluyeron artículos publicados entre 2010 y 2023, en inglés o español, que fueran estudios empíricos o revisiones sobre la integración de tecnología y pedagogía en STI. Se excluyeron estudios no centrados específicamente en STI, así como artículos de opinión o editoriales.

Tras el cribado inicial, se seleccionaron 230 artículos para el análisis cuantitativo, utilizando el software VOSviewer para realizar un análisis bibliométrico y de minería de textos. Este análisis se centró en la frecuencia de términos clave, redes de co-citación y tendencias temporales en temas de investigación. De estos 230 artículos, se seleccionaron 35 para un análisis cualitativo

en profundidad, basándose en su relevancia y número de citas. El análisis de contenido de estos artículos se enfocó en codificar enfoques pedagógicos, técnicas de modelado del estudiante, estrategias de retroalimentación y andamiaje, y métodos de evaluación.

Para complementar los hallazgos de la revisión de literatura, se realizaron entrevistas semiestructuradas con 15 expertos en el campo de los STI. Estos expertos, seleccionados mediante muestreo intencional, incluyeron 8 investigadores académicos con publicaciones destacadas, 4 desarrolladores senior de empresas líderes en STI, y 3 educadores con experiencia en la implementación de estos sistemas. Los expertos tenían un promedio de 12 años de experiencia en el campo, representaban 7 países diferentes y abarcaban diversas áreas de especialización, incluyendo inteligencia artificial, psicología educativa y diseño instruccional.

Las entrevistas, con una duración de 60 a 90 minutos, se centraron en explorar enfoques efectivos para integrar tecnología y pedagogía, desafíos en el desarrollo e implementación de STI, tendencias emergentes y direcciones futuras, así como consideraciones éticas y de accesibilidad.

Para integrar los hallazgos cuantitativos y cualitativos, se empleó una matriz de convergencia que permitió comparar las tendencias identificadas en el análisis bibliométrico con los temas emergentes del análisis cualitativo y las perspectivas de los expertos. Este proceso de triangulación facilitó la identificación de áreas de consenso, así como posibles discrepancias entre la literatura publicada y las opiniones de los expertos en el campo.

Esta metodología mixta, que combina el análisis riguroso de la literatura existente con las perspectivas actualizadas de expertos en el campo, permitió obtener una comprensión holística y matizada de las estrategias actuales y emergentes para la integración efectiva de tecnología y pedagogía en los Sistemas de Tutoría Inteligente.

3. RESULTADOS

Los resultados de este estudio se presentan en tres secciones principales: análisis bibliométrico, análisis cualitativo de la literatura, y análisis de las entrevistas a expertos.

Análisis bibliométrico

El análisis bibliométrico de los 230 artículos seleccionados reveló varias tendencias significativas en la investigación sobre la integración de tecnología y pedagogía en los Sistemas de Tutoría Inteligente (STI).

La distribución geográfica de las publicaciones mostró una concentración en ciertos países, como se muestra en la Tabla 1.

Tabla 1

Distribución geográfica de las publicaciones

País	Porcentaje de publicaciones
Estados Unidos	32%
China	18%
Reino Unido	12%
Canadá	8%
Alemania	7%
Otros	23%

Esta distribución sugiere una posible brecha en la investigación proveniente de otras regiones, especialmente de países en desarrollo.

En cuanto a las perspectivas teóricas, se identificaron cinco enfoques predominantes, como se muestra en la Tabla 2.

Tabla 2

Perspectivas teóricas predominantes

Perspectiva teórica	Porcentaje de artículos
Constructivismo	28%
Aprendizaje adaptativo	23%
Teoría de la carga cognitiva	18%
Aprendizaje situado	15%
Conectivismo	10%
Otras teorías	6%

El análisis de co-citación reveló tres clusters principales de investigación:

1. Modelado del estudiante y adaptación (35% de las co-citaciones)
2. Diseño de interfaces y experiencia de usuario (28%)
3. Evaluación de la efectividad de los STI (22%)

La evolución temporal de los temas de investigación mostró un creciente interés en la integración de tecnologías emergentes como la realidad virtual y aumentada (incremento del 180% en los últimos 5 años), y en consideraciones éticas y de privacidad (incremento del 150%).

Análisis cualitativo de la literatura

El análisis en profundidad de los 35 artículos seleccionados permitió identificar patrones y relaciones entre diferentes aspectos de la integración de tecnología y pedagogía en los STI.

Emergieron cuatro temas principales, con sus respectivas subcategorías, como se muestra en la Tabla 3.

Tabla 3

Temas principales y subcategorías emergentes del análisis cualitativo

Tema principal	Subcategorías
Personalización del aprendizaje	- Modelado cognitivo del estudiante - Adaptación en tiempo real - Consideración de factores emocionales
Estrategias pedagógicas innovadoras	- Aprendizaje basado en problemas - Gamificación - Aprendizaje colaborativo asistido por IA
Desafíos de implementación	- Escalabilidad - Interoperabilidad - Formación docente
Consideraciones éticas y sociales	- Privacidad de datos - Equidad y accesibilidad - Transparencia algorítmica

Se observó una fuerte relación entre las estrategias de personalización del aprendizaje y las consideraciones éticas, con el 75% de los artículos que abordaban la personalización también discutiendo implicaciones éticas.

La integración de estrategias pedagógicas innovadoras se correlacionó positivamente con los desafíos de implementación, especialmente en términos de escalabilidad y formación docente. El 80% de los artículos que proponían enfoques pedagógicos novedosos también abordaban estos desafíos.

Análisis de las entrevistas a expertos

Las entrevistas con los 15 expertos proporcionaron resultados que complementaron y, en algunos casos, contrastaron con los hallazgos de la literatura.

Los temas más frecuentemente mencionados por los expertos se presentan en la Tabla 4.

Tabla 4

Temas frecuentes en las entrevistas a expertos

Tema	Frecuencia de mención
Integración de IA avanzada	13/15
Necesidad de enfoques pedagógicos híbridos	12/15
Preocupaciones sobre la brecha digital	11/15
Importancia de la colaboración interdisciplinaria	10/15
Desafíos en la evaluación a largo plazo de la efectividad de los STI	9/15

Un hallazgo notable fue la discrepancia entre la literatura y las opiniones de los expertos respecto a la importancia de la gamificación. Mientras que la gamificación aparecía como una estrategia pedagógica innovadora en el 40% de los artículos analizados, solo 4 de los 15 expertos la consideraron una prioridad para el futuro de los STI.

Los expertos también enfatizaron la necesidad de un mayor enfoque en la formación docente y la integración de los STI en los entornos educativos existentes, un aspecto que estaba relativamente subrepresentado en la literatura analizada.

La triangulación de los datos cuantitativos, el análisis cualitativo de la literatura y las perspectivas de los expertos reveló una convergencia en la importancia de la personalización del aprendizaje y las consideraciones éticas como áreas clave para el futuro desarrollo de los STI. Sin embargo, también se identificaron brechas significativas, particularmente en la investigación sobre la implementación práctica y a gran escala de estos sistemas en diversos contextos educativos.

4. DISCUSIÓN Y CONCLUSIONES

Los resultados de este estudio ofrecen una perspectiva integral sobre las tendencias actuales y los desafíos futuros en la integración de tecnología y pedagogía en los Sistemas de Tutoría Inteligente (STI), revelando un panorama más complejo y matizado que el anticipado inicialmente.

El análisis bibliométrico destaca una concentración geográfica significativa de la investigación en STI, con un claro dominio de países como Estados Unidos, China y Reino Unido. Este hallazgo subraya la urgente necesidad de diversificar la investigación para incorporar perspectivas de diferentes contextos culturales y educativos, especialmente de países en desarrollo. La inclusión de estas diversas perspectivas es crucial para garantizar que los STI sean verdaderamente inclusivos y adaptables a las variadas realidades educativas globales, como lo sugieren Nkambou et al. (2010) en su análisis de los avances en sistemas de tutoría inteligente.

La predominancia del constructivismo y el aprendizaje adaptativo como marcos teóricos principales en la literatura analizada (Jonassen, 1991; Essa, 2016) contrasta notablemente con la emergencia de enfoques híbridos señalada por los expertos entrevistados. Esta discrepancia evidencia una brecha entre la teoría y la práctica en el desarrollo de STI, apuntando hacia la necesidad de investigaciones que exploren la efectividad de estos enfoques híbridos en contextos reales de aprendizaje. La exploración y validación de estos enfoques pedagógicos híbridos, que combinan las fortalezas de diferentes teorías del aprendizaje, se perfila como un área crítica para futuras investigaciones, en línea con las observaciones de Graesser et al. (2018) sobre la evolución de los STI.

Un hallazgo particularmente relevante es la fuerte correlación observada entre las estrategias de personalización del aprendizaje y las consideraciones éticas. Este vínculo, no anticipado en la introducción, subraya la complejidad de implementar sistemas altamente personalizados mientras se salvaguarda la privacidad y se garantiza la equidad. Futuros desarrollos en STI deberán abordar este dilema de manera proactiva, posiblemente a través de enfoques de "ética por diseño" en la arquitectura de estos sistemas. El desarrollo de marcos éticos integrales

que aborden la tensión entre personalización y privacidad se presenta como un desafío crucial para el campo, como lo señalan Arroyo et al. (2014) en su trabajo sobre sistemas multimedia de tutoría inteligente.

La discrepancia entre la prominencia de la gamificación en la literatura y la opinión de los expertos emerge como otro hallazgo significativo. Mientras que la literatura sugiere un entusiasmo por la gamificación como estrategia pedagógica innovadora (Schroeder & Adesope, 2014), los expertos mostraron un escepticismo notable. Esta divergencia apunta a la necesidad de investigaciones más rigurosas sobre la efectividad a largo plazo de las estrategias de gamificación en los STI, más allá de los efectos motivacionales a corto plazo, un aspecto que VanLehn (2011) también destacó en su comparación de diferentes sistemas de tutoría.

Los desafíos de implementación, especialmente en términos de escalabilidad y formación docente, se revelan como preocupaciones centrales tanto en la literatura como en las entrevistas con expertos. Este hallazgo subraya la importancia de considerar no solo los aspectos técnicos y pedagógicos de los STI, sino también los contextos organizacionales y humanos en los que se implementarán. La priorización de la investigación sobre estrategias de implementación efectiva, incluyendo la formación docente y la integración en sistemas educativos existentes, se perfila como un área crucial para el avance del campo, en consonancia con las observaciones de Koedinger y Aleven (2007) sobre el dilema de la asistencia en los tutores cognitivos.

La creciente atención a las consideraciones éticas y sociales, evidenciada tanto en el análisis bibliométrico como en las entrevistas, revela una maduración del campo. Sin embargo, la brecha entre el reconocimiento de estas preocupaciones y la implementación de soluciones concretas sugiere la necesidad de marcos éticos más robustos y aplicables en el desarrollo y despliegue de STI, un aspecto que Essa (2016) anticipó en su visión del futuro de los sistemas de aprendizaje adaptativo.

En conclusión, este estudio revela que la integración efectiva de tecnología y pedagogía en los STI es un desafío multifacético que va más allá de las consideraciones técnicas y pedagógicas inicialmente anticipadas. Los hallazgos subrayan la necesidad de un enfoque holístico que abarque la diversificación de la investigación, la exploración de enfoques pedagógicos híbridos, el desarrollo de marcos éticos integrales, la reevaluación del papel de la gamificación, y la priorización de estrategias de implementación efectiva. El fomento de la colaboración interdisciplinaria emerge como un elemento clave para abordar la complejidad de estos desafíos identificados, como lo sugieren Azevedo y Aleven (2013) en su trabajo sobre metacognición y tecnologías de aprendizaje.

Estas conclusiones representan una contribución significativa al campo, proporcionando una hoja de ruta para futuras investigaciones y desarrollos en STI que vayan más allá de las expectativas iniciales. Al abordar de manera integral los desafíos emergentes en la intersección de la tecnología, la pedagogía y la ética, este estudio sienta las bases para el desarrollo de STI más efectivos, éticos e inclusivos, capaces de responder a las diversas necesidades educativas del siglo XXI, en línea con la visión de futuro propuesta por Magnisalis et al. (2011) para los sistemas adaptativos e inteligentes de apoyo al aprendizaje colaborativo.

5. REFERENCIAS

- Arroyo, I., Bursleson, W., Tai, M., Muldner, K., & Woolf, B. P. (2014). A multimedia intelligent tutoring system for scaffolding self-regulated learning. *International Journal of Artificial Intelligence in Education*, 24(2), 168-207. <https://doi.org/10.1007/s40593-013-0007-9>
- Azevedo, R., & Aleven, V. (2013). Metacognition and Learning Technologies: An Overview of Current Interdisciplinary Research. In *International Handbook of Metacognition and Learning Technologies* (pp. 1-16). Springer. https://doi.org/10.1007/978-1-4419-5546-3_1
- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated Cognition and the Culture of Learning. *Educational Researcher*, 18(1), 32-42. <https://doi.org/10.3102/0013189X018001032>
- Conati, C., Gertner, A., & VanLehn, K. (2002). Using Bayesian Networks to Manage Uncertainty in Student Modeling. *User Modeling and User-Adapted Interaction*, 12(4), 371-417. <https://doi.org/10.1023/A:1021258506583>
- Essa, A. (2016). A Possible Future for Next Generation Adaptive Learning Systems. *International Journal of Artificial Intelligence in Education*, 26(2), 801-819. <https://doi.org/10.1007/s40593-016-0101-y>
- Graesser, A. C., Chipman, P., Haynes, B. C., & Olney, A. (2005). AutoTutor: An Intelligent Tutoring System with Mixed-Initiative Dialogue. *IEEE Transactions on Education*, 48(4), 612-618. <https://doi.org/10.1109/TE.2005.856149>
- Graesser, A. C., Hu, X., & Sottolare, R. (2018). Intelligent Tutoring Systems. In F. Fischer, C. E. Hmelo-Silver, S. R. Goldman, & P. Reimann (Eds.), *International Handbook of the Learning Sciences* (pp. 246-255). Routledge. <https://doi.org/10.4324/9781315617572-24>
- Jonassen, D. H. (1991). Evaluating Constructivistic Learning. *Educational Technology*, 31(9), 28-33.
- Koedinger, K. R., & Aleven, V. (2007). Exploring the Assistance Dilemma in Experiments with Cognitive Tutors. *Educational Psychology Review*, 19(3), 239-264. <https://doi.org/10.1007/s10648-007-9049-0>
- Koedinger, K. R., & Corbett, A. T. (2006). Cognitive Tutors: Technology Bringing Learning Sciences to the Classroom. In K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 61-77). Cambridge University Press. <https://doi.org/10.1017/CBO9780511816833.006>
- Koedinger, K. R., Anderson, J. R., Hadley, W. H., & Mark, M. A. (2003). Toward a Model of Collaboration and Tutoring in Problem-Solving. *Journal of Artificial Intelligence in Education*, 12(1), 47-63.
- Magnisalis, I., Demetriadis, S., & Karakostas, A. (2011). Adaptive and Intelligent Systems for Collaborative Learning Support: A Review of the Field. *IEEE Transactions on Learning Technologies*, 4(1), 5-20. <https://doi.org/10.1109/TLT.2011.2>
- Mayer, R. E. (2004). Should There Be a Three-Strikes Rule Against Pure Discovery Learning? *American Psychologist*, 59(1), 14-19.
- Mayer, R. E. (2009). *Multimedia Learning* (2nd ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511811678>

- Mitrovic, A., & Ohlsson, S. (1999). Evaluation of a Constraint-Based Tutor for a Database Language. *International Journal of Artificial Intelligence in Education*, 10(3-4), 238-256.
- Nkambou, R., Mizoguchi, R., & Bourdeau, J. (2010). *Advances in Intelligent Tutoring Systems*. Springer. <https://doi.org/10.1007/978-3-642-14363-2>
- Piaget, J. (1970). *Science of Education and the Psychology of the Child*. Orion Press.
- Roll, I., Alevan, V., McLaren, B. M., & Koedinger, K. R. (2007). Designing for Metacognition - Applying Cognitive Tutoring Principles to the Tutoring of Help Seeking. *Metacognition and Learning*, 2(2-3), 125-140. <https://doi.org/10.1007/s11409-007-9010-9>
- Schroeder, N. L., & Adesope, O. O. (2014). A Systematic Review of Pedagogical Agents' Persona, Motivation, and Cognitive Load Implications for Learners. *Journal of Research on Technology in Education*, 46(3), 229-251.
- Schwartz, D. L., Chase, C. C., Oppezzo, M. A., & Chin, D. B. (2011). Practicing Versus Inventing with Contrasting Cases: The Effects of Telling First on Learning and Transfer. *Journal of Educational Psychology*, 103(4), 759-775. <https://doi.org/10.1037/a0025140>
- Shute, V. J. (2008). Focus on Formative Feedback. *Review of Educational Research*, 78(1), 153-189. <https://doi.org/10.3102/0034654307313795>
- Soh, L. K., Tan, L. C., Wong, J., Lim, M., & Chen, X. (2017). A Pedagogical Framework for Integrating Effective Tutoring Strategies in Intelligent Tutoring Systems. In *Proceedings of the 25th International Conference on Computers in Education* (pp. 196-205). Asia-Pacific Society for Computers in Education.
- Soller, A., Martínez, A., Jermann, P., & Muehlenbrock, M. (2005). From Mirroring to Guiding: A Review of State of the Art Technology for Supporting Collaborative Learning. *International Journal of Artificial Intelligence in Education*, 15(4), 261-290. <https://dl.acm.org/doi/10.5555/1434935.1434937>
- Tsai, C. C., Shen, P. D., & Lu, Y. J. (2020). Adaptive Computer-Assisted Learning Systems: Behavior, Motivation and Performance. *Educational Technology & Society*, 23(2), 90-106.
- VanLehn, K. (2011). The Relative Effectiveness of Human Tutoring, Intelligent Tutoring Systems, and Other Tutoring Systems. *Educational Psychologist*, 46(4), 197-221. <https://doi.org/10.1080/00461520.2011.611369>


Para citar este artículo:

Ruiz Muñoz, G. F. Integración de la tecnología y la pedagogía en los sistemas de tutoría inteligente. *EduTec, Revista Electrónica de Tecnología Educativa*, (89), 144-155. <https://doi.org/10.21556/edutec.2024.89.3199>



Uso de las TIC por alumnado de aula hospitalaria: Un estudio de caso

Use of ICT by students in a hospital classroom: A case study

 Sara Martínez-Carrera; sara.martinezcarrera@uvigo.es

 Cristina Sánchez-Martínez; c.sanchez@uvigo.es

 Paula Raña-Veloso; paula.r.v.7@gmail.com

Universidade de Vigo (España)

Resumen

El alumnado de las aulas hospitalarias, compuesto por niños y niñas con enfermedades temporales o crónicas, requiere una atención especial tanto física como psicológica. En este contexto, las TIC pueden ser una herramienta muy útil para apoyar su educación y bienestar emocional. La presente investigación establece como objetivo general examinar la percepción del profesorado, alumnado y sus familias acerca del uso de las TIC en el aula hospitalaria, así como hacer una comparativa con el aula ordinaria y el hogar. Para ello se aplicó una metodología cualitativa, concretamente un enfoque etnográfico-narrativo, con un estudio de caso. Participaron en total 11 alumnos y alumnas de aula hospitalaria, 4 padres y madres de jóvenes hospitalizados y 2 profesoras. Los resultados obtenidos tras el análisis de los datos evidencian la satisfacción de los tres colectivos sobre el uso de las TIC, especialmente en el aula hospitalaria. Destacan como ventajas la promoción del aprendizaje activo y metodologías innovadoras, así como el aumento de la motivación del alumnado; aunque en alguna ocasión se puede generar distracción. Sobresale el uso diario del ordenador, la Pizarra Digital Interactiva y el Smartphone, para realizar tareas académicas y recreativas.

Palabras clave: enseñanza, tecnología, aula hospitalaria, escuela, hogar.

Abstract

The student in hospital classrooms, made up of boys and girls with temporary or chronic illnesses, requires special attention, both physical and psychological. In this context, ICT can be a very useful tool to support their education and emotional well-being. The general objective of this research is to examine the perception of teachers, students and their families about the use of ICT in the hospital classroom, as well as to make a comparison with the ordinary classroom and the home. For this purpose, a qualitative methodology was applied, specifically an ethnographic-narrative approach, with a case study. A total of 11 students from the hospital classroom, 4 parents of hospitalized young people and 2 teachers participated. The results obtained after data analysis show the satisfaction of the three groups with the use of ICT, especially in the hospital classroom. The advantages are the promotion of active learning and innovative methodologies, as well as the increase in student motivation; although on some occasions it can be distracting. The daily use of the computer, the Interactive Digital Whiteboard and the Smartphone to carry out academic and recreational tasks stands out.

Keywords: teaching, technology, hospital classroom, school, home.



1. INTRODUCCIÓN

La sociedad se encuentra sumergida en una vorágine de transformaciones que requieren de la población una rápida adecuación y evolución (Siddiq & Scherer, 2019). Esto exige, como señala Kauffman (2015) que los individuos sean vistos como seres integrales, teniendo en cuenta aspectos físicos, biológicos, psíquicos, culturales, sociales e históricos. Frente a esto, el sistema educativo se enfrenta constantemente al desafío de innovar y transformar los procesos pedagógicos para adaptarse a la complejidad de la sociedad (Vico-Bosch & Rebollo-Catalán, 2019). En tal sentido, comprender la complejidad en el contexto educativo es fundamental para abordar los problemas diarios de la enseñanza. La educación debe buscar una conexión entre el conocimiento, el contexto y las cualidades individuales (UNESCO, 2022). De esta forma, el contexto se vuelve un elemento clave en el proceso pedagógico. Uno de los contextos pedagógicos de mayor complejidad hasta el momento ha sido el de las aulas hospitalarias.

En la actualidad, las Tecnologías de la Información y la Comunicación (TIC) forman parte del día a día de la ciudadanía, y participan en los cambios sociales, económicos, culturales y educativos que se producen en la sociedad. En este sentido, las instituciones educativas no pueden mantenerse al margen, puesto que la implantación de las TIC en la educación puede ayudar a enriquecer, transformar y complementar la trayectoria académica y personal del alumnado (Roth-Cohen et al., 2022). En este sentido, autores como Grizzle et al. (2023) ponen de manifiesto la importancia de la educación mediática.

1.1 Aulas hospitalarias

Los niños/as enfermos/as temporalmente o que se enfrentan a enfermedades crónicas, deben continuar aprendiendo, y las aulas hospitalarias son una alternativa para garantizar el derecho a la educación de la infancia (Riera-Negre et al., 2021). Las políticas educativas hospitalarias se desarrollan gracias a la disciplina denominada, Pedagogía Hospitalaria. Según Ocampo y Monsalve (2020), la Pedagogía Hospitalaria es un puente entre la práctica pedagógica y el hospital, y entre la educación y la salud.

La Pedagogía Hospitalaria se recoge en la Carta Europea de los Derechos del Niño Hospitalizado, aprobada por el Parlamento Europeo en 1986 en el que se reconoce el derecho de todo/a niño/a a proseguir su formación escolar durante su permanencia en el hospital; y en la Asociación Europea de Pedagogos Hospitalarios (HOPE). Dicha disciplina acoge las unidades escolares ubicadas en las instituciones hospitalarias denominadas aulas hospitalarias (Asensio-Ramón, 2023).

Las aulas hospitalarias se definen como un recurso del sistema educativo para responder a las necesidades de escolares que dejan de asistir al colegio de manera regular a causa de una enfermedad. Son espacios creados para ofrecer apoyo académico y emocional al alumnado que no puede asistir al colegio por encontrarse hospitalizado (Peña-Hita et al., 2022). Como indican González-González et al. (2021) las aulas hospitalarias están concebidas con la finalidad de continuar el desarrollo educativo, afectivo y social de la infancia que se encuentra hospitalizada, siempre y cuando su estado médico lo permita.

Los/las docentes que desarrollan su trabajo en las aulas hospitalarias presentan ciertos desafíos; deben de tener en cuenta la diversidad de situaciones y condiciones de cada

alumno/a (Salinas, 2020). En este sentido, es imprescindible integrar los aspectos psicosociales del alumnado y de su familia en el proceso de tratamiento de su enfermedad, siendo fundamental la colaboración de todo el equipo que trabaja con los pacientes infantiles: profesorado de las aulas hospitalarias, profesionales de su centro de referencia, compañeros/as, etc. (Fierros et al., 2020). Cabe señalar, como indica Peña-Hita et al. (2022) el desconocimiento existente sobre el servicio de aulas hospitalarias por parte de las familias y del personal médico-sanitario.

1.2. Las TIC en las aulas hospitalarias

Las TIC forman parte de la vida cotidiana, y el proceso de enseñanza-aprendizaje se encuentra mediado por ellas. Como indican González-González et al. (2021), los medios tecnológicos están presentes en los contextos hospitalarios, brindando múltiples posibilidades acerca de la normalización de aquellas personas en situación de enfermedad o convalecencia. Varios autores (Maher, 2020; Ocampo & Lizasoán, 2019) sostienen que la tecnología puede favorecer la integración, inclusión y normalización, creando espacios de atención educativa virtual. Asimismo, la tecnología también permite potenciar la presencialidad con metodologías flexibles, adaptadas al alumnado hospitalizado.

Los/as niños/as y adolescentes hospitalizados son una población en riesgo de exclusión. Debido a que las enfermedades graves y crónicas deben tratarse con hospitalizaciones prolongadas o frecuentes, los y las pacientes son retirados/as de sus entornos sociales: la escuela y la familia (Bastidas-Rivera et al., 2023). Las personas que están en las aulas hospitalarias ven interrumpida su vida de forma inesperada, lo que puede desmotivar y provocar emociones de miedo y tristeza, afectando al rendimiento escolar, entre otros aspectos (De la Mula, 2018). Como indican Mombaers y Donche (2020) la hospitalización regular o a largo plazo puede desempeñar un papel en la motivación académica que, a su vez, afecta al compromiso académico, al aprendizaje y al rendimiento académico. Por este motivo introducir elementos tecnológicos puede favorecer su motivación.

Los/as profesionales de la educación se enfrentan con frecuencia al desafío de incorporar las TIC y el aprendizaje digital en el aula, ya que la demanda de la aplicación de las TIC es inminente desde todos los ámbitos (político, social, educativo...) y contextos. Con todo, muchos/as docentes recalcan la necesidad de adquirir las habilidades TIC necesarias (Falloon, 2020). La competencia en alfabetización digital debe tener un vínculo beneficioso con la implementación de las TIC en el aula, tanto para el crecimiento del profesorado como para el de sus estudiantes (Sánchez-Cruzado et al., 2021).

Varios estudios destacan los beneficios del uso de la tecnología en el aula hospitalaria (Fuentes-Cabrera & Sánchez-Romero, 2020; Serrano & Castañeda., 2016). Mateus et al. (2023), por ejemplo, ponen de manifiesto en su proyecto Aulas Hospitalarias de Fundación Telefónica, los beneficiosos cambios en los paradigmas pedagógicos y estructurales del profesorado, así como en el desarrollo de habilidades para su adaptación a las necesidades individuales del alumnado.

El objetivo general de este estudio es examinar la percepción del profesorado, alumnado y sus familias acerca del uso de las TIC en un aula hospitalaria, así como realizar una comparativa con el aula ordinaria y el hogar. Como objetivos específicos se han determinado:

- Conocer los recursos TIC empleados en el aula hospitalaria, el aula ordinaria y el hogar.
- Detectar la frecuencia de uso de las TIC en el aula hospitalaria, el aula ordinaria y el hogar.
- Identificar las actividades desarrolladas con las TIC en el aula hospitalaria, el aula ordinaria y el hogar.
- Reconocer las ventajas del uso de las TIC en el aula hospitalaria, el aula ordinaria y el hogar.
- Señalar los inconvenientes que presenta la utilización de las TIC en el aula hospitalaria, el aula ordinaria y el hogar.

2. MÉTODO

La presente investigación se enmarca en la metodología cualitativa, concretamente se trata de un estudio de caso. El enfoque es etnográfico-narrativo, Según Creswell (2013) este enfoque implica una perspectiva interpretativa y holística que busca comprender los fenómenos sociales en sus contextos naturales. En concreto, se ha implementado un estudio de caso tal como sugiere Yin (2018). Los estudios de caso permiten una comprensión integral de un fenómeno en específico, contribuyendo al desarrollo de teorías. A través de una exploración profunda de casos individuales, los/as investigadores/as pueden identificar patrones, temas y relaciones que enriquecen o perfeccionan teorías existentes. Este proceso de construcción teórica fortalece la base del conocimiento científico y facilita la toma de decisiones basada en evidencia (George y Bennett, 2005).

2.1. Aproximación contextual y participantes/muestra

El contexto estudiado ha sido seleccionado por su disposición a participar en la investigación. Se trata de un aula hospitalaria situada en el Completo Hospitalario de una ciudad pequeña del interior de la Comunidad Autónoma de Galicia. En el estudio han participado alumnado, profesorado y familias.

En lo referente al alumnado, participaron un total de 11 niños/as que en diferentes momentos se encontraban hospitalizados/as por diversas razones, 7 (63.64%) de los cuales se identifican como género femenino y 4 (36.36%) masculino, y con edades comprendidas entre los 9 y los 14 años, con una media de 12.63. Cursan 1º, 2º y 3º de Educación Secundaria Obligatoria (ESO) y 4º y 5º Educación Primaria y provienen de diferentes centros educativos y localidades de la provincia (Tabla 1).

Tabla 1

Perfil del alumnado

Género	Edad	Curso	Centro educativo	Localidad
Femenino	9	4º EP	CPI IV	Ou
Femenino	14	3º ESO	IES OP	Ou
Femenino	13	2º ESO	IES XFC	AV
Masculino	13	2º ESO	IES XTC	V
Femenino	14	3º ESO	IES XTC	V
Femenino	13	2º ESO	Colegio Plurilingüe DPF	Ou
Masculino	14	2º ESO	IES OP	Ou
Femenino	13	1º ESO	IES 12O	Ou
Masculino	10	5º EP	CEIP Plurilingüe PAN	Ou
Masculino	14	1º ESO	IES LO	OBV
Femenino	12	1º ESO	IES XFC	Ou

Haciendo referencia a las familias de los/as jóvenes hospitalizados/as, participaron 2 padres y 2 madres con edades comprendidas entre los 32 y los 48 años. Sobre el nivel de estudios únicamente 1 persona cuenta con un FP de 2º grado, 1 presenta estudios superiores y las 2 restantes, estudios secundarios. Atendiendo a la situación laboral, 3 se encuentran en activo y 1 está en el paro. Presentan profesiones diversas: confección, operario de fábrica, enfermera y caja-reponedora. Participan en el estudio las 2 profesoras del aula hospitalaria analizada, con 40 y 23 años. 1 de ellas, graduada en Educación infantil y con máster de Educación Secundaria con la especialidad de Orientación; la otra participante cuenta con el grado de Educación Primaria (Tablas 2 y 3).

Tabla 2

Perfil de las familias

Género	Edad	Parentesco	Estudios	Situación laboral	Profesión	Curso de su hijo/a
Masculino	43	Padre	FP 2º grado	En activo	Operario de fábrica	2º ESO
Femenino	48	Madre	Estudios superiores	En activo	Enfermera	3º ESO
Femenino	32	Madre	Estudios secundarios	En paro	Cajera-reponedora	5º EP
Masculino	46	Padre	Estudios secundarios	En activo	Confección	3º ESO

Tabla 3

Perfil del profesorado

Género	Edad	Estudios	Aula hospitalaria	Tipo de centro	Localidad
Femenino	40	Grado de Educación Infantil, Máster de Educación Secundaria por Orientación	Aula hospitalaria de Ou	Centro específico	Ou
Femenino	23	Grado de Educación Primaria	Aula hospitalaria de Ou	Centro específico	Ou

2.2. Técnicas/Instrumentos de recogida de información

Para llevar a cabo la recolección de información en este estudio se aplicaron las técnicas de la encuesta y la entrevista. La primera hace referencia a un proceso de recolección de datos a través de cuestionarios diseñados previamente, en los que se realizan un conjunto de preguntas que responden a los objetivos de la investigación. La segunda, consiste en un método estructurado y sistemático de recopilación de información o datos mediante la interacción directa entre un entrevistador y un entrevistado (Hernández et al., 2016).

Para llevar a cabo las encuestas se elaboraron y aplicaron dos cuestionarios Ad Hoc; uno online a través de Google Forms para el alumnado y otro impreso para las familias. Para su elaboración y validación se ha contado con el apoyo de dos personas expertas en la temática y en el tipo de estudio. Los cuestionarios se formularon bajo la modalidad de preguntas abiertas, las cuales, permiten a las personas encuestadas brindar respuestas con sus propias palabras; permiten conocer sus perspectivas (Martínez, 2007). Los cuestionarios quedaron constituidos, en ambos casos, por una primera parte que recoge los datos demográficos (género, edad, localidad...) y dos bloques de preguntas correspondientes al uso de las TIC (que consta de 4 preguntas) y a sus ventajas y desventajas (que contiene dos preguntas) respectivamente. Cada uno de estos bloques subdivididos a su vez en relación al contexto: aula hospitalaria, aula ordinaria y hogar. En total, cada cuestionario estaba compuesto por 18 preguntas abiertas.

Por otra parte, se aplicó una entrevista semiestructurada que consiste en un conversatorio entre el investigador/a y el/la participante, por medio del cual se pretende hallar las respuestas a un conjunto de interrogantes e inquietudes (Seidman, 2013). Para ello, se sigue un guion o protocolo de entrevista, implicando escuchar al sujeto para comprender sus ideas (Flick, 2015). La entrevista se estructuró, al igual que en el cuestionario, en dos bloques: el uso de las TIC y ventajas y desventajas, además de los datos de perfil. Cada uno de estos bloques con tres apartados correspondiente al contexto: aula hospitalaria, aula ordinaria y hogar. El guion de la entrevista, al igual que los cuestionarios fue elaborado Ad Hoc, con el apoyo de dos personas expertas en la temática y en el tipo de estudio. Del mismo modo, las preguntas eran las mismas.

2.3. Procedimiento de recogida y análisis de datos

Para el desarrollo de este estudio se comenzó por presentar la investigación al personal del aula hospitalaria, solicitando consentimiento informado para garantizar la participación voluntaria y consciente de todas las personas participantes en la investigación. Una vez obtenido el consentimiento, se planteó un cronograma de aplicación que fue aprobado por los

responsables del aula hospitalaria. Se planificaron diversos días, atendiendo a las fechas de hospitalización de los/las estudiantes.

El alumnado participante utilizó un ordenador para responder al cuestionario online. El espacio se adecuaba para que pudiese sentirse cómodo. Se prestaron las orientaciones pertinentes y se resolvieron las dudas. El tiempo de cumplimentación del cuestionario varió entre 20 y 30 minutos.

Para los familiares, el cuestionario se aplicó de forma impresa, facilitando los recursos necesarios. Los padres y madres también contaron con un espacio adecuado para responder de manera tranquila y con la seguridad requerida. Asimismo, se aclararon las dudas pertinentes. El tiempo de cumplimentación del cuestionario varió entre 20 y 30 minutos.

Las entrevistas con el profesorado se llevaron a cabo en el aula hospitalaria. Las preguntas se fueron desarrollando, en un clima agradable y de confianza, lo que favoreció la interacción. Asimismo, se aclararon las dudas pertinentes. Las entrevistas fueron respaldadas en grabaciones que posteriormente fueron transcritas en formato Word para proceder a su análisis.

A los datos obtenidos, tanto los recogidos a través de los cuestionarios online e impresos, como a través de las entrevistas, se le ha aplicado un análisis de contenido. Para delimitar el primer nivel de categorización se partió de los objetivos del estudio y el asesoramiento facilitado por dos expertos. De esta forma, y partiendo de la información bruta se han establecido categorías y subcategorías; triangulando las respuestas obtenidas por los tres colectivos participantes (alumnado, familia y profesorado), organizándolas y cuantificando las frecuencias de acuerdo con la información recopilada: recursos TIC utilizados, frecuencia, actividades realizadas con las TIC, ventajas e inconvenientes de su uso. Cada uno de estos aspectos distribuidos según el contexto: aula hospitalaria, aula ordinaria y hogar. El análisis en profundidad se ha realizado con el programa de Analysis of Qualitative Data (AQUAD), versión 7. A partir de la codificación generada con AQUAD se han elaborado las tablas de resultados, que una vez volcadas en Excel permiten proporcionar una presentación gráfica y sistemática de los mismos. La representación icónica es esencial para ofrecer una interpretación y comprensión efectiva de los resultados.

3. RESULTADOS

Los resultados del estudio se aglutinan en torno a cinco subapartados, dando respuesta a los diferentes objetivos específicos de la investigación.

3.1. Recursos TIC utilizados

El alumnado del aula hospitalaria sostiene que los recursos TIC que más utiliza son el ordenador ($f = 10/11$; 90.9%) y la Pizarra Digital Interactiva (PDI) ($f = 1/10$; 0.9%). No obstante, el profesorado añade la tableta, la impresora y la Nintendo ($f = 2/2$; 100% en todas ellas). Por su parte, la familia considera que el recurso TIC más utilizado en el aula hospitalaria es el ordenador ($f = 3/4$; 75%) (Figura 1).

A continuación, se presentan algunos extractos ilustrativos:

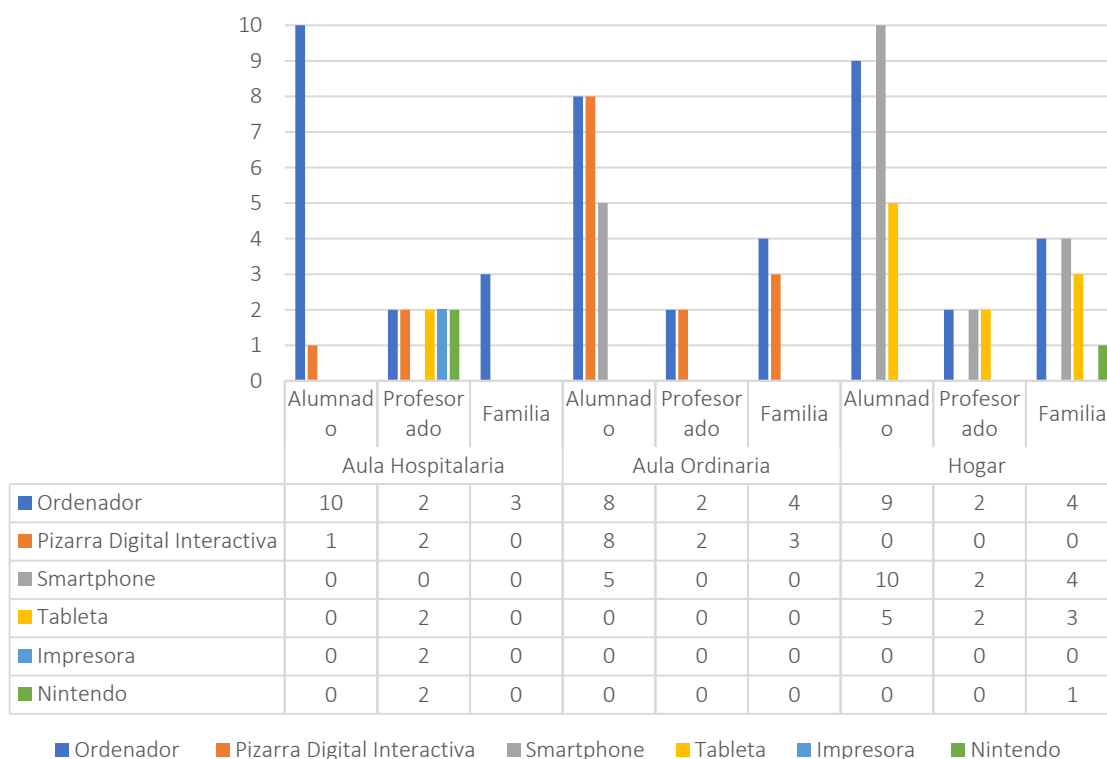
“En el aula hospitalaria utilizamos principalmente el ordenador y la PDI” (Cuestionario abierto 1, alumna, 9 años).

“Sin duda, en el aula hospitalaria utilizamos el ordenador y la PDI, aunque también la tableta y la impresora” (Entrevista 1, profesora, 40 años).

“Yo creo que en el aula hospitalaria utilizan mucho el ordenador” (Cuestionario abierto 4, padre, 46 años).

Figura 1

Recursos tecnológicos empleados por el alumnado del aula hospitalaria



El alumnado considera que, en el aula ordinaria, los recursos tecnológicos más utilizados son el ordenador y la PDI (f = 8/11; 72.72% cada uno); también incluyen el Smartphone (f = 5/11; 45.45%), en menor frecuencia. Para el profesorado y la familia, los recursos más utilizados en este espacio son el ordenador (profesorado: f = 2/2; 100%/ familia f = 4/4; 100%) y la PDI (profesorado: f = 2/2; 100%/ familia f = 3/4; 75%).

Haciendo referencia a los recursos TIC utilizados en el hogar, el alumnado destaca el uso del Smartphone (f = 10/11; 90.90%), el ordenador (f = 9/11; 81%) y la tableta (f = 5/11; 45.45%). El profesorado, coincide plenamente (f = 2/2; 100% cada uno); mientras que la familia añade la Nintendo (f = 1/4; 25%).

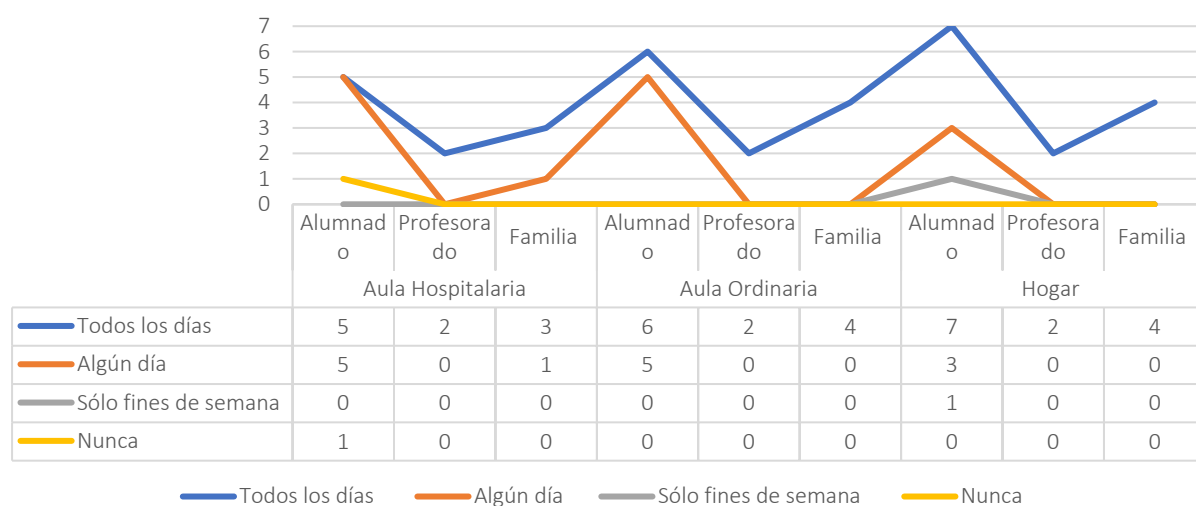
3.2. Frecuencia de uso de las TIC

Los y las estudiantes ponen de manifiesto que utilizan los recursos TIC todos los días, o por lo menos algún día de la semana (f = 5/11; 45.45% cada uno); excepcionalmente un estudiante

indica que nunca los utiliza (f = 1/11; 9.09%). La totalidad del profesorado sostiene que en el aula hospitalaria se utilizan los recursos digitales todos los días (f = 2/2; 100%). Por su parte, la familia considera que el alumnado utiliza los recursos todos los días (f = 3/4; 75 %) o al menos algún día de la semana (f = 1/4; 25%) (Figura 2).

Figura 2

Temporalización de uso de los recursos TIC por el alumnado del aula hospitalaria



Atendiendo al aula ordinaria, el propio alumnado confirma que utiliza los recursos digitales todos los días (f = 6/11; 54.54%) o al menos algún día de la semana (f = 5/11; 45.45%). Para el profesorado y la familia, estos recursos son utilizados todos los días en el aula ordinaria (profesorado f = 2/2; 100% /familia f = 4/4; 100%).

A continuación, se presentan algún extracto ejemplificador:

“Depende de su centro educativo. Algunos/as alumnos/as solo emplean el ordenador en alguna materia en concreto, como por ejemplo en tecnología. Otros en cambio lo emplean todos los días porque usan E-Dixgal y ya lo necesitan para entrar en su aula virtual. La Pizarra Digital Interactiva, en general la usan a diario” (Entrevista 1, profesora, 40 años).

En el hogar, más de la mitad del alumnado sostiene que utilizan las TIC todos los días (f = 7/11; 63.63%), alguno, por otra parte, afirma que lo hacen algún día (f = 3/11; 27.27%) y únicamente un alumno indica que el uso de dispositivos digitales en el hogar se supedita exclusivamente a los fines de semana (f = 1/11; 9.09%). La totalidad del profesorado y de las familias asegura que el alumnado utiliza las TIC en el hogar todos los días (f = 2/2; 100%) (f = 4/4; 100%).

3.3. Actividades realizadas con las TIC

En el aula hospitalaria los recursos tecnológicos son utilizados con mayor frecuencia, según el alumnado, para la realización de actividades académicas (f = 11/11; 100%). No obstante, también se emplean para el desarrollo de actividades recreativas y de ocio (f=4/11; 36.36%) y, en menor medida, de actividades comunicativas (f = 2/11; 18.18%). De acuerdo con el

profesorado, en el aula hospitalaria los/las niños/as utilizan las TIC tanto para la realización de actividades académicas (f = 2/2; 100%), como recreativas (f = 2/2; 100%) y comunicativas (f = 2/2; 100%). Por su parte, la familia considera que el alumnado en este contexto utiliza las TIC para la realización de actividades académicas (f = 4/4; 100%) y comunicativas (f = 4/4; 100%), aunque en un caso se considera que pueden tener un uso recreativo/de ocio (f = 1/4; 25%) (Figura 3).

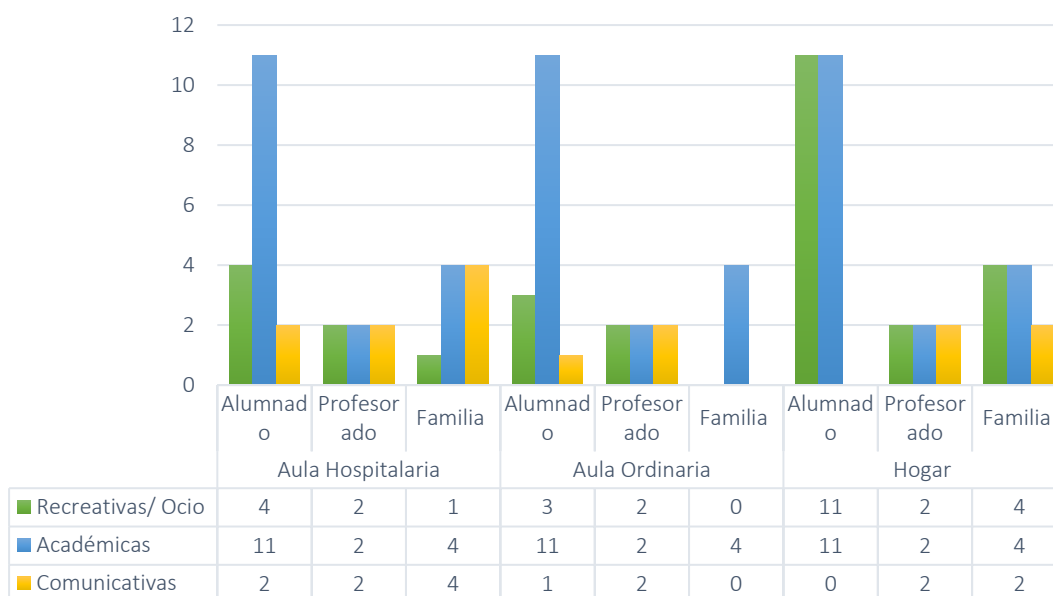
Algunos extractos ilustrativos son:

“Las TIC se utilizan mucho para contactar con el colegio/instituto de los niños/as mediante correo electrónico y que puedan continuar realizando los ejercicios que le manden los profesores, así como los exámenes en caso de que pasen mucho tiempo hospitalizados” (Entrevista 1, Profesora, 40 años).

“En el aula hospitalaria yo creo que utilizan las TIC para contactar con los/las profesores/as, para utilizar el aula virtual (deberes), para buscar información, para realizar los trabajos y deberes...” (Cuestionario abierto, madre, 48 años).

Figura 3

Prácticas educativas con las TIC llevada a cabo por el alumnado de aula hospitalaria



La totalidad del alumnado considera que en el aula ordinaria, utilizan las TIC principalmente para la realización de actividades académicas (f = 11/11; 100%). No obstante, en alguna ocasión indican que desarrollan actividades recreativas (f = 3/11; 27.27%) y comunicativas (f = 1/11; 9.9%). Por su parte, el profesorado manifiesta que sus estudiantes, en el aula ordinaria, utilizan los recursos digitales tanto para el desarrollo de actividades académicas (f = 2/2; 100%), como comunicativas (f = 2/2; 100%) y recreativas (f = 2/2; 100%). De acuerdo con las familias, los recursos TIC en el centro educativo, se emplean únicamente para realizar actividades académicas (f = 4/4; 100%).

A continuación, un extracto ilustrativo:

“En el aula ordinaria el profesor utiliza la PDI para explicar los temas, nos enseña ejemplos, vídeos explicativos...Nosotros utilizamos el ordenador para buscar información y realizar trabajos en grupo. Cuando nos dejan utilizar el Smartphone, lo hago para buscar palabras en el diccionario online de gallego. Utilizamos mucho los

dispositivos para hacer actividades académicas. Por otra parte, también utilizamos la aplicación Zoom y hacemos videollamadas con otras clases, es muy divertido comunicarnos así. Cuando vamos bien de tiempo, el profesor nos deja repasar contenido a través de juegos online ¡nos encanta!” (Cuestionario abierto 6, alumna, 13 años).

Con respecto al contexto del hogar, la totalidad del alumnado sostiene que utiliza los recursos TIC para realizar actividades académicas ($f = 11/11$; 100%) y recreativas ($f = 11/11$; 100%). El profesorado por su parte hace referencia al uso de dispositivos digitales para el desarrollo de actividades académicas, recreativas y comunicativas ($f = 2/2$; 100% cada una). La familia considera que sus hijos/as utilizan las TIC en el hogar principalmente para llevar a cabo praxis académicas y recreativas ($f = 4/4$; 100% cada una) y con menor frecuencia comunicativas ($f = 2/4$; 50%).

3.4. Ventajas del uso de las TIC

La totalidad del alumnado coincide en que una de las ventajas del uso de las TIC en el aula hospitalaria consiste en la promoción del aprendizaje activo ($f = 11/11$; 100%). En menor medida, los/las niños/as destacan como aspecto positivo la motivación que provoca en ellos/as el uso de dispositivos digitales ($f = 3/11$; 27.27%), y la generación de prácticas innovadoras ($f = 3/11$; 27.27%). Asimismo, consideran que facilitan la comunicación y las relaciones personales ($f = 2/11$; 18.18%). Por su parte, las docentes destacan en igual medida las siguientes ventajas: promoción del aprendizaje activo, incremento de la motivación, impulso de prácticas innovadoras, promoción del aprendizaje lúdico y de la comunicación ($f = 2/2$; 100%). La familia resalta, como ventaja del uso de las TIC en el aula hospitalaria el aprendizaje activo que se genera ($f = 3/4$; 75%), así como el aumento de la motivación ($f = 1/4$; 25%) y las comunicaciones ($f = 1/4$; 25%) (Figura 4).

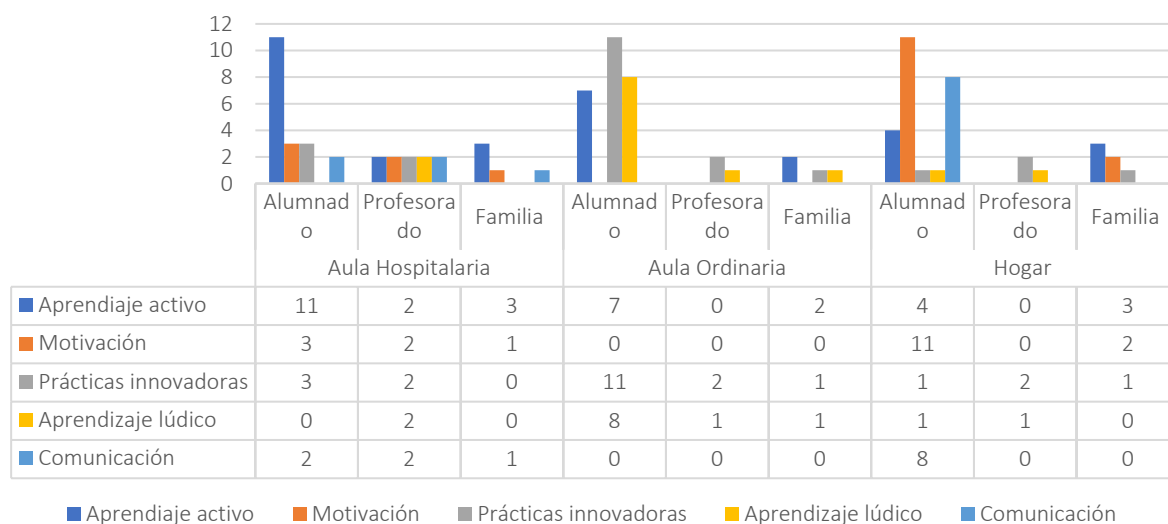
A continuación, se muestran algunos ejemplos que ayudan a ilustrar lo mencionado anteriormente:

“Gracias a las TIC puedo mantener el contacto fluido con el centro educativo de los/as niños/as hospitalizados/as y que así puedan seguir realizando las actividades escolares, incluso los exámenes, si es necesario. Las TIC les permiten seguir con su formación durante su permanencia en el hospital de una forma activa y cuando vuelvan a su centro educativo irán prácticamente al ritmo de sus compañeros/as. También les permiten divertirse jugando y olvidarse, por un momento, de su situación” (Entrevista 1, profesora, 40 años).

“Gracias a las TIC los/las chicos/as no quedan atrasados en los estudios, pueden buscar información y aprender de forma activa, lo que aumenta su motivación. Además, les permite la comunicación con el colegio y los/las compañeros/as” (Cuestionario abierto 2, madre, 48 años).

Figura 4

Ventajas del empleo de las TIC por el alumnado de aula hospitalaria



En el contexto de aula ordinaria, la principal ventaja del uso de las TIC para el alumnado es el desarrollo de prácticas innovadoras ($f = 11/11$; 100%). Asimismo, destaca como aspecto positivo la promoción del aprendizaje lúdico ($f = 8/11$; 72.72%) y activo ($f = 7/11$; 63.63%). El profesorado, por su parte, señala que la mayor ventaja en este espacio es la realización de prácticas innovadoras ($f = 2/2$; 100%), así como la promoción del aprendizaje lúdico ($f = 1/2$; 50%). En tanto, la familia resalta positivamente el aprendizaje activo que se promueve ($f = 2/4$; 50%), así como el desarrollo de prácticas innovadoras ($f = 1/4$; 25%) y del aprendizaje lúdico ($f = 1/4$; 25%).

En el hogar, el alumnado considera que el uso de recursos tecnológicos incrementa su motivación ($f = 11/11$; 100%). Asimismo, le permite establecer comunicaciones y relaciones con otras personas ($f = 8/11$; 72.72%) y promueve un aprendizaje activo ($f = 4/11$; 36.36%). Las docentes, destacan la posibilidad de realizar prácticas innovadoras ($f = 2/2$; 100%). La familia, por su parte señala como aspecto positivo el aprendizaje activo ($f = 3/4$; 75%), y el incremento de la motivación ($f = 2/4$; 50%).

1.2 Inconvenientes del uso de las TIC

La mayor parte del alumnado considera, que el uso de las TIC en el aula hospitalaria no presenta ninguna desventaja o inconveniente ($f = 7/11$; 63.63%). No obstante, algún alumno/a señala que los dispositivos digitales pueden presentar problemas técnicos ($f = 2/11$; 18.18%), así como producir distracción ($f = 1/11$; 9.09%) y pérdida de la privacidad ($f = 1/11$; 9.09%). Para las

docentes y la familia, en el contexto de aula hospitalaria las TIC no tienen ninguna desventaja (profesorado $f = 2/2$; 100% / familia $f = 3/4$; 74%) (Figura 5).

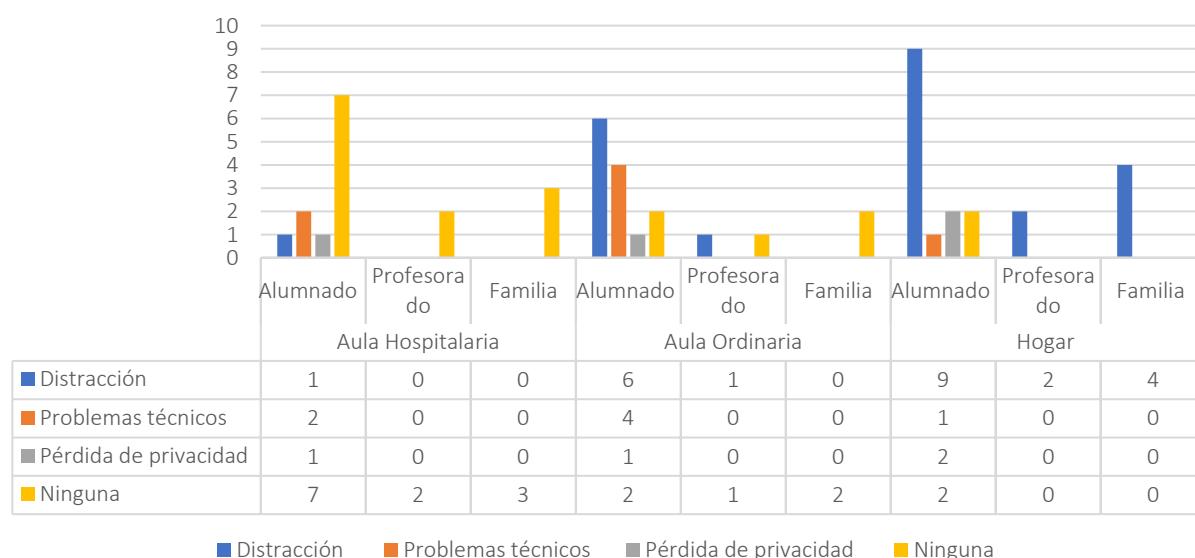
Se presentan algunos extractos ilustrativos:

“Puede haber algún problema técnico. En algún momento no puedes acceder a alguna página web o aplicación a la hora que quieras y/o pueden estar restringidas” (Cuestionario abierto 2, alumna, 14 años).

“En principio, ninguna desventaja que resaltar” (Cuestionario abierto 2, madre, 48 años).

Figura 5

Desventajas del empleo de las TIC por el alumnado de aula hospitalaria



En el aula ordinaria, la mayoría de los y las estudiantes resaltan como principal desventaja la distracción que les puede causar ($f = 6/11$; 54.54%), así como el enfrentarse a problemas técnicos ($f = 4/11$; 36.36%). Para el profesorado, una docente cree que puede generar distracción, mientras que la otra no considera que exista ningún inconveniente. La familia coincide con esta última ($f = 2/4$; 50%).

En cuanto al uso de las TIC en el hogar, los niños y niñas consideran la generación de distracción una desventaja ($f = 9/11$; 81.81%). En algún caso, señalan como inconveniente la pérdida de privacidad ($f = 2/11$; 18.18%), así como la existencia de problemas técnicos ($f = 2/11$; 18.18%). Para las profesoras, la mayor desventaja del uso de los recursos digitales en el hogar se basa en la distracción ($f = 2/2$; 100%), coincidiendo con la familia ($f = 4/4$; 100%).

4. DISCUSIÓN Y CONCLUSIONES

El objetivo de esta investigación fue descubrir la percepción del profesorado, alumnado y sus familias acerca del uso de las TIC en un aula hospitalaria, así como realizar una comparativa con el aula ordinaria y el hogar.

De este estudio se desprende que, en el aula hospitalaria y el aula ordinaria, el recurso TIC más utilizado es el ordenador seguido de la PDI; mientras que en el hogar son el Smartphone y el ordenador los recursos más empleados. Estas herramientas digitales se utilizan diariamente en los tres contextos, formando parte del día a día de los/las niños/as. Varios estudios ponen de manifiesto la prevalencia del ordenador (Bjursten et al., 2023; Lee & Kim, 2023), la PDI (Deveci-Topal et al., 2023; Saed et al., 2024) y el Smartphone (Hartley & Andújar, 2022; Wang et al., 2023) en contextos educativos. En este sentido, Peña-Hita et al., (2022) en su estudio señalan que las aulas hospitalarias disponen de suficientes recursos materiales, adaptados al nivel educativo de los/as pacientes.

De esta investigación se extrae que las actividades más desarrolladas con las TIC en los espacios analizados son las de índole académica; seguidas de las recreativas y de ocio. Varios estudios (Morales, 2019 & Trost, 2019) ponen de manifiesto la existencia del amplio abanico de herramientas y aplicaciones informáticas que pueden ser usadas en el contexto hospitalario: herramientas para el desarrollo de actividades escolares, como por ejemplo de trabajo colaborativo, o redes sociales y videojuegos, que no se limitan al uso plenamente didáctico sino que amplían el uso social y lúdico-educativo de las tecnologías.

El uso de la tecnología demuestra beneficios positivos de su utilización, así como una buena aceptación por parte del alumnado, de sus familias y los/las profesionales de la salud y educadores/as (Dawe et al., 2019). En este sentido, entre las ventajas de uso de las TIC, en el aula hospitalaria, destaca la promoción del aprendizaje activo, al permitir al alumnado participar en su propio aprendizaje. Fuentes-Cabrera y Sánchez-Romero (2020) y Mateus et al. (2023), muestran que la tecnología puede utilizarse para mejorar el proceso de enseñanza y aprendizaje de niños/as enfermos/as hospitalizados/as. Se resalta la importancia de la inclusión de tecnologías para la mejora del bienestar del niño/a enfermo/a, que a su vez, consolida la continuidad del proceso de enseñanza-aprendizaje y posibilita la colaboración y comunicación con sus iguales. En el aula ordinaria, su principal ventaja es el uso de metodologías innovadoras. Por su parte, en el hogar estos recursos incrementan la motivación. En esta línea, muchos estudios analizan los efectos positivos del uso de las TIC en la enseñanza (Garzón et al., 2020; Wang et al., 2021).

En general no se han detectado grandes desventajas respecto al uso de las TIC en el aula hospitalaria. Sin embargo, se puede señalar la distracción que pueden generar los dispositivos en el hogar y en el aula ordinaria. Vázquez-Cano et al. (2022) añaden otros aspectos como los problemas de privacidad, el discernimiento de información confiable y relevante, así como la infoxicación.

Como conclusiones cabe señalar que los colectivos implicados en la investigación muestran su satisfacción con el uso de las TIC en los tres contextos analizados, haciendo especial énfasis en los beneficios que presenta su utilización en el aula hospitalaria. En la actualidad, las TIC son una herramienta de gran utilidad en las aulas hospitalarias, contexto educativo cuyo alumnado presenta unas características especiales al estar enfermo. En este sentido, la tecnología se convierte en una herramienta fundamental para que estos/as niños/as no pierdan el ritmo escolar mientras se encuentren hospitalizados/as. Además, les permite comunicarse con sus profesores/as, compañeros/as y familiares; y entretenerse, lo que contribuye positivamente a su estado emocional.

Las principales limitaciones de este estudio hacen referencia a la muestra del mismo. En un contexto educativo tan especial como el aula hospitalaria, y tratándose de un estudio de caso cualitativo, el número de participantes es limitado. Además, varios/as niños/as no asistían al aula hospitalaria regularmente, lo que dificultó la recogida de datos. Las familias también han supuesto una dificultad en la investigación al no participar muchas de ellas. En cuanto al profesorado, se contaba exclusivamente con dos maestras para toda el aula. Con todo, la investigación supone una contribución a la comunidad científica y educativa, al permitir conocer el uso que hace de las TIC el alumnado de aula hospitalaria, tanto en ésta, como en el aula ordinaria y el hogar.

En futuras investigaciones se contempla hacer un estudio comparativo, atendiendo a distintas aulas hospitalarias de varias ciudades y/o Comunidades Autónomas. Asimismo, sería interesante, implementar un programa para mejorar la competencia digital del alumnado de aula hospitalaria, que se podría extender al profesorado y a las familias, con la finalidad de propiciar espacios de colaboración, e integrar propuestas para hacer más inclusiva la atención educativa hospitalaria.

5. AGRADECIMIENTOS

Nos gustaría agradecer la implicación y colaboración de los niños/as hospitalizados, así como de sus familias y profesoras. En un momento complicado agradecemos especialmente su disposición.

6. REFERENCIAS

- Asensio-Ramón, P. (2023). El papel del Tercer Sector: el voluntariado y las ONGs en las aulas hospitalarias en España. *Revista Española de Educación Comparada*, 42, 378-398. <https://revistas.uned.es/index.php/REEC/article/view/31502/27227>
- Bastidas-Rivera, S., Sánchez, J.E., Sierra-Gil, Y.L. y González-Cabal, L.F. (2023). The Hospital Classroom: An Opportunity for Educational Inclusion. *Human Arenas* <https://doi.org/10.1007/s42087-023-00362-6>
- Bjursten, E.L., Nilsson, T. y Gumaelius, L. (2023). Computer programming in primary schools: Swedish Technology Teachers' pedagogical strategies. *International Journal of Technology and Design Education*, 33, 1345–1368. <https://doi.org/10.1007/s10798-022-09786-7>
- Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches* (3rd ed.). Sage Publications.
- Dawe, J., Sutherland, C., Barco, A. y Broadbent, E. (2019). Can social robots help children in healthcare contexts? A scoping review. *BMJ paediatrics open*, 3(1), 1-16. <https://doi.org/10.1136/bmjpo-2018-000371>

- De la Mula-Fuentes, B., Quintana, M., Rimbau, J., Martínez-Mejías, A., Socorro, M., Rivera-Pérez, C. y Garolera, M. (2018). Ansiedad, miedos hospitalarios y alteraciones conductuales en la hospitalización infantil. *Actas Españolas de Psiquiatría*, 46(2), 42-50.
- Deveci-Topal, A., Kolburan-Geçer, A. y Çoban-Budak, E. (2023). An analysis of the utility of digital materials for high school students with intellectual disability and their effects on academic success. *Universal Access in the Information Society* 22, 95–110. <https://doi.org/10.1007/s10209-021-00840-0>
- Falloon, G. (2020). From digital literacy to digital competence: the teacher digital competency (TDC) framework. *Educational Technology Research and Development*, 68, 2449–2472. <https://doi.org/10.1007/s11423-020-09767-4>
- Fierros, M., Bengoechea, C., Yáñez, S., Martínez, C. y López-Ibor, B. (2020). Un colegio en una unidad de oncología pediátrica: El papel crucial de la educación en el contexto de una enfermedad grave. *Revista Electrónica Educare*, 24(1), 1-16. <https://doi.org/10.15359/ree.24-1.5>
- Flick, U. (2015). *El diseño de la investigación cualitativa*. Morata
- Fuentes-Cabrera, M.J. y Sánchez-Romero, C. (2021). Analysis of the social and educational advantages of ICT for the sick child. *Revista de Educación Inclusiva*, 13, 59-75. <https://revistaeducacioninclusiva.es/index.php/REI/article/view/505>
- Garzón, J., Baldiris, S., Gutiérrez, J. y Pavón, J. (2020). How do pedagogical approaches affect the impact of augmented reality on education? A meta-analysis and research synthesis. *Educational Research Review*, 31, 100334. <https://doi.org/10.1016/j.edurev.2020.100334>
- George, A. L. y Bennett, A. (2005). *Case studies and theory development in the social sciences*. Cambridge University Press.
- González-González, C., Violant Holz, V., Infante Moro, A., Cáceres García, L. y Guzmán Franco, M.D. (2021). Educational robotics in inclusive contexts: The case of the hospital classrooms. *Educación XX1*, 24(1), 375-403, <http://doi.org/10.5944/educXX1.27047>
- Grizzle, A., Wilson, C., Tuazon, R., Cheung, C. K., Lau, J., Fischer, R., ... & Gulston, C. (2023). *Pensar críticamente, hacer clic sabiamente. Currículum de Alfabetización Mediática e Informativa para educadores y estudiantes*. UNESCO.
- Hartley, K. y Andújar, A. (2022). Smartphones and Learning: An Extension of M-Learning or a Distinct Area of Inquiry. *Education Sciences* 12(1), 50. <https://doi.org/10.3390/educsci12010050>
- Hernández, R., Fernández, C. y Baptista, P. (2016). *Metodología de la Investigación* (6ta Edición). Mac Graw Hill Education.
- Kauffman, S. (2015). *La humanidad en un mundo complejo*. Tusquets Editores.
- Lee, H. y Kim, Y. (2023). Exploring the Effects of Computer and Smart Device-Assisted Learning on Students' Achievements: Empirical Evidence from Korea. *Sustainability*. 15(17):13241. <https://doi.org/10.3390/su151713241>

- Maher, D. (2020). Pre-Service Teachers' Digital Competencies to Support School Students' Digital Literacies. In Handbook of Research on Literacy and Digital Technology Integration in Teacher Education (pp. 29-46). IGI Global. <https://doi.org/10.4018/978-1-7998-1461-0.ch002>
- Martínez, R. A. (2007). *La investigación en la práctica educativa: Guía metodológica de investigación para el diagnóstico y evaluación en los centros docentes*. Fareso, S. A.
- Mateus, J.C., Sotelo, J.M. y Fernandez, A. (2023). ICT for diversity: Analysis of two experiences of inclusive and hospital education in Peru. *Revista de Educación Inclusiva* 16(1), 8-23. <https://revistaeducacioninclusiva.es/index.php/REI/article/view/793/702>
- Mombaers, T. y Donche, V. (2020). Hospital School Students' Academic Motivation and Support Needs: A Self-Determination Perspective. *Frontiers in Education*, 5. <https://doi.org/10.3389/feduc.2020.00106>
- Morales, L.O. (2019). La pedagogía hospitalaria, un reto gigante para la virtualidad. *Revista Reflexiones y Saberes*, 10, 17-24. <https://revistavirtual.ucn.edu.co/index.php/RevistaRyS/article/view/1067>
- Ocampo, A. G., y Lizasoain, O. R. (2019). Pedagogía hospitalaria: trayectorias de desarrollo intelectual, conquistas profesionales y desafíos de futuro. *Revista Intersaberes*, 13(29), 415-423.
- Ocampo, A. y Monsalve, C. (2020). Epistemología de la pedagogía hospitalaria. *Revista Educación las Américas*, 10(1), 118–128. <https://doi.org/10.35811/rea.v10i0.93>
- Peña-Hita, MA., Pegalajar-Palomino, MC. y Carpio-Fernández, M.V (2022). La educación desde las aulas hospitalarias: percepciones de la familia y del personal médico-sanitario. *Educar* 58(2), 517-531. <https://doi.org/10.5565/rev/educar.1521>
- Riera-Negre, L., Paz-Lourido, B., Negre, F., Rosselló, M.R. y Verger, S. (2021). Self-Perception of Quality of Life and Emotional Well-Being among Students Attending Hospital Classrooms during COVID-19 Pandemic. *Healthcare*, 9(8), 943. <https://doi.org/10.3390/healthcare9080943>
- Roth-Cohen, O., Rosenberg, H. y Lissitsa, S. (2022). Are you talking to me? Generation X, Y, Z responses to mobile advertising. *Convergence*, 28(3), 761-780. <https://doi.org/10.1177/13548565211047342>
- Saed, H., Al-Refai, A.M., Aladwan, D., Abu-Rahme, M., Alnajjar, E., Masalha, F. y Basheti, I. (2024). Factors Influencing Academic Staff Attitudes Towards Smart Boards Use in Education Process in Applied Science University. *Pegem Journal of Education and Instruction*, 14(1), 20-28. <https://doi.org/10.47750/pegegog.14.01.03>
- Salinas, S.C. y Ormeño, F.R. (2020). Concepciones sobre afectividad en docentes que trabajan en aulas hospitalarias en Chile. *Educación*, 29, 27–47. <https://doi.org/10.18800/educacion.202001.002>
- Sánchez-Cruzado, C., Santiago Campión, R. y Sánchez-Compañía, M. (2021). Teacher digital literacy: The indisputable challenge after COVID-19. *Sustainability* 13, 1858. <https://doi.org/10.3390/su13041858>

- Seidman, I. E. (2013). *Interviewing as qualitative research: A guide for researchers in education and the social sciences*. Teachers College Press.
- Serrano, J. L. y Castañeda, L. (2016). Proyecto EDUMOBSPITALARIOS: desarrollo profesional docente e innovación con m-learning en aulas hospitalarias. *EduTec. Revista Electrónica De Tecnología Educativa*, 55, a327. <https://doi.org/10.21556/edutec.2016.55.641>
- Siddiq, F., y Scherer, R. (2019). Is there a gender gap? A meta-analysis of the gender differences in students' ICT literacy. *Educational Research Review*, 27, 205-217. <https://doi.org/10.1016/j.edurev.2019.03.007>
- Trost, M.J., Ford, A.R., Kysh, L., Gold, J.I. y Matari, M. (2019). Socially Assistive Robots for Helping Pediatric Distress and Pain. *The Clinical Journal of Pain*, 35(5), 451-458. <https://doi.org/10.1097/AJP.0000000000000688>
- UNESCO (2022). Reimaginar juntos nuestros futuros: un nuevo contrato social para la educación. UNESCO y SM. <https://unesdoc.unesco.org/ark:/48223/pf0000381560>
- Vázquez-Cano, E., Parra-González, M.ª E., Segura-Robles, A. y López-Meneses, E. (2022). The negative effects of technology on education: a bibliometric and topic modeling mapping analysis (2008-2019). *International Journal of Instruction*, 15(2), 37-60. <https://doi.org/10.29333/iji.2022.1523a>
- Vico-Bosch, A. y Rebollo-Catalán, Á. (2019). Women's learning on internet and social networks: validation and general results of a scale. *Educación XX1*, 22(1), 375-400. <https://doi.org/10.5944/educxx1.21469>
- Wang, J., Tigelaar, D. E. y Admiraal, W. (2021). Rural teachers' sharing of digital educational resources: From motivation to behavior. *Computers & Education*, 161, 104055. <https://doi.org/10.1016/j.compedu.2020.104055>
- Wang, J.C., Hsieh, C.Y. y Kung, S.H. (2023). The impact of smartphone use on learning effectiveness: A case study of primary school students. *Education and Information Technologies* 28, 6287–6320. <https://doi.org/10.1007/s10639-022-11430-9>
- Yin, R. K. (2018). *Case study research and applications: Design and methods*. Sage Publications.

Para citar este artículo:

Martínez-Carrera, S., Sánchez-Martínez, C. y Raña-Veloso, P. Uso de las TIC por alumnado de aula hospitalaria: Un estudio de caso. *EduTec, Revista Electrónica de Tecnología Educativa*, 89, 156-173. <https://doi.org/10.21556/edutec.2024.89.3281>




¿Cómo influye el nivel socioeconómico en la autopercepción de la competencia digital del alumnado de educación secundaria en España?

How does socioeconomic level influence the self-perception of digital competence of secondary school students in Spain?

Luz Mayra Niño-Cortés; luzmayra.nino@estudiants.urv.cat; Universitat Rovira i Virgili (España)

 Carme Grimalt-Álvaro; carme.grimalt@uab.cat; Universitat Autònoma de Barcelona (España)

 Irma Zoraida Sanabria Cárdenas; isanabri@unet.edu.ve; Universidad Nacional Experimental del Táchira (Venezuela)

 Mireia Usart Rodríguez; mireia.usart@urv.cat; Universitat Rovira i Virgili (España)

Resumen

En un mundo cada vez más digitalizado, resulta crucial comprender la brecha digital en educación. Esta brecha, derivada de desigualdades sociales para adquirir habilidades digitales, impulsa a investigar el desarrollo de la Competencia Digital (CD) y sus factores determinantes. Este estudio indaga la influencia de la variable socioeconómica en la autopercepción de la CD del alumnado de la Educación Secundaria en España. Se utilizó un enfoque cuantitativo, con un diseño de encuesta y una muestra de 2411 alumnos/as, empleando el cuestionario Digitalis-ESO. A través del análisis descriptivo y comparativo, se identificaron disparidades en CD entre niveles socioeconómicos, así como una relación significativa entre el nivel socioeconómico, CD y ciclo de estudio. Los resultados muestran que, a medida que el nivel socioeconómico y el ciclo de estudio aumentan, también lo hace la CD autopercebida, principalmente en alumnado de nivel socioeconómico bajo y medio durante la Educación Secundaria Obligatoria (ESO). Este hallazgo evidencia la vulnerabilidad del alumnado de entornos socioeconómicos bajos, en particular, en la primera etapa de la ESO y en mayor medida en los aspectos académicos de la CD. Así, se evidencia una brecha digital en la educación secundaria, que coloca al alumnado socioeconómicamente desfavorecido en desventaja para adquirir CD.

Palabras clave: Competencia Digital autopercebida, nivel socioeconómico, estudiantes, educación secundaria.

Abstract

In an increasingly digitalized world, it is crucial to understand the digital divide in education. This divide, derived from social inequalities in acquiring digital skills, prompts research into the development of Digital Competence (DC) and its determinants. This study investigates the influence of the socioeconomic variable on students' self-perception of DC in Secondary Education in Spain. A quantitative approach was used, with a survey design and a sample of 2411 students, using the Digitalis-ESO questionnaire. Through descriptive and comparative analysis, disparities in DC between socioeconomic levels were identified, as well as a significant relationship between socioeconomic level, DC and study cycle. The results show that, as socioeconomic status and cycle study increase, so does self-perceived DC, mainly in students of low and middle socio-economic status during lower-secondary school (ESO). This finding shows the vulnerability of students from low socio-economic backgrounds, particularly in the first stage of lower-secondary and to a greater extent in the academic aspects of DC. Thus, there is evidence of a digital divide in secondary education, which places socio-economically disadvantaged students at a disadvantage in acquiring DC.

Keywords: Self-perceived digital competence, socioeconomic level, students, secondary education.



1. INTRODUCCIÓN

La brecha digital entendida como la desigualdad de oportunidades, en lo social y familiar, para el desarrollo de CD (Unesco, 2022) representa un desafío que requiere atención exhaustiva y sólida de los factores que la determinan. Desde el contexto de la educación secundaria y, en un mundo cada vez más digitalizado, es esencial investigar y comprender la Competencia Digital (CD) como un constructo que dota al alumnado de habilidades para participar de manera activa en la sociedad, adquiriendo destrezas para el uso crítico y responsable de las tecnologías digitales (TD).

En este escenario, el factor socioeconómico emerge como un elemento clave que puede incidir en la brecha digital en la educación y predecir la CD. Por ello, la presente investigación tiene como propósito examinar la influencia del nivel socioeconómico y ciclo de estudio en el desarrollo de la CD del alumnado de la Educación Secundaria Obligatoria (ESO) y, bachillerato. Fundamentado en el currículo oficial de España, que reconoce esta etapa educativa por su importancia en la adquisición y desarrollo de competencias clave en el alumnado (LOE, 3/2020).

El entorno global del alumnado, junto con el acceso o no a las TD, influye en su CD. Investigaciones revelan que el alumnado de secundaria tiene niveles medios de CD autopercebida (Colás-Bravo et al., 2017), con diferencias según su entorno socioeconómico (Soria-Pérez et al., 2022). A medida que avanzan en sus estudios, tienden a sentirse más competentes digitalmente (Niño-Cortés et al., 2023) especialmente aquellos de niveles socioeconómicos más altos (Scherer & Siddiq, 2019). Así que algunos estudios sugieren que el nivel socioeconómico puede impactar sobre la CD (Matamala, 2015) pero implementar estrategias educativas reforzadas podría mitigar esta influencia (Hatlevik et al., 2015).

Esto hace fundamental examinar el impacto socioeconómico en el desarrollo de la CD del alumnado, especialmente en la transición de la ESO hacia la etapa postobligatoria. Aunque este fenómeno de la brecha digital ha sido discutido fuera del ámbito académico, aún faltan estudios integrales, entre lo social, lo cognitivo y tecnológico, para abordar estas desigualdades y promover una educación equitativa, alineada con la Agenda 2030 y el Plan Nacional de competencias digitales – España digital 2025 (Comisión Europea, 2020).

1.1. La Competencia Digital (CD) del alumnado de educación secundaria

La CD del alumnado es un constructo multidimensional, explicado desde distintas acepciones y contextos de investigación (Comisión Europea, 2017; Gisbert et al., 2016). Larraz (2013) destaca la importancia de la CD al capacitar al alumnado para gestionar información y contenido digital de manera eficaz y crítica; habilidades esenciales para participar en la sociedad desde una identidad digital. Por tanto, de acuerdo con Larraz, este estudio reconoce la CD en términos de uso y manejo de la tecnología digital en ámbitos informacionales, tecnológicos, multimedia y comunicativos.

En el ámbito de las ciencias sociales, y en concreto en la educación secundaria, ha surgido la necesidad de plantear estrategias que permitan medir la CD del alumnado, y así, entender su desarrollo y el impacto en la mejora educativa y social en su conjunto. Para ello, existen distintas formas e instrumentos que pretenden medir esta competencia (González-Rodríguez & Urbina-Ramírez, 2020), por una parte, pruebas objetivas de desempeño (Baeza-González et al., 2022; Fernández Abuín, 2016), y por otra, cuestionarios que miden la autopercepción del alumnado sobre su CD (Colás-Bravo et al., 2017).

Las pruebas de diagnóstico de autopercepción de la CD son las más utilizadas, mientras que las pruebas objetivas de desempeño se realizan a menor escala (González-Rodríguez & Urbina-Ramírez, 2020). Esta diversidad en el método para medir la CD asevera las distintas acepciones y marcos de referencia que intentan delimitar este constructo y, las dimensiones que la componen (Colás-Bravo et al., 2017). Respecto a las pruebas de desempeño de CD, pocos estudios existentes muestran que el alumnado de secundaria no tiene suficiente nivel de CD o se encuentra en estado inicial; en países bajos (Van Deursen & Van Diepen, 2013) y en España (Fernández-Abuín, 2016). Mientras que, en Latinoamérica, no se encuentran resultados sobre este análisis de la CD.

Los reportes de la CD autopercebida, en España, sugieren que el alumnado de secundaria se percibe con un nivel de CD media, con mayor habilidad para buscar información en internet, y en menor medida, para crear contenido y gestionar la seguridad en red (Colás Bravo et al., 2017; Vila-Couñago, 2020). Asimismo, estudios recientes, como el de Verdú-Pina (2024) indican que el alumnado de secundaria se percibe con un nivel de CD entre medio y alto, con diferencias significativas en función de otras variables, como el curso de estudio. Además, investigaciones como la de Matamala (2015) señala que los bajos niveles de CD podrían estar relacionados con el entorno socioeconómico del alumnado.

En síntesis, el ámbito de las ciencias de la educación ha asumido el reto de valorar y comprender la CD, enfocado principalmente en los factores individuales que influyen en el desarrollo de esta competencia (Fernández-Mellizo & Manzano, 2018) y en menor proporción sobre los factores contextuales que podrían incidir en su desarrollo. Así, en el presente estudio se busca profundizar en el nivel de CD autopercebida del alumnado más allá de sus propias características individuales, y sobre la relación de esta competencia con la brecha digital en la educación.

1.2. La brecha digital en educación. El nivel socioeconómico asociado a la CD del alumnado

Atender la brecha digital en educación se ha convertido en una herramienta para promover la inclusión digital del alumnado (Casillas et al., 2022). Una vez superadas las disparidades de cobertura y acceso a la red, la dificultad para integrarse a la sociedad digital se centra en el desarrollo de la CD (Márquez-Fernández, 2023). Esto conlleva en asumir la importancia de la brecha digital desde la posibilidad de desarrollar capacidades y habilidades que permitan a los individuos aprovechar las ventajas de los entornos tecnológicos (la formación virtual, teletrabajo, desarrollo de negocios o entretenimiento) (González-Benito, 2023) y, la desigualdad de oportunidades, en lo social y familiar (OCDE, 2001; UNESCO 2022).

La literatura señala que, el alumnado de familias con bajo nivel socioeconómico carece de apoyo social para adquirir conocimientos sobre tecnología lo que resulta en una menor confianza en habilidades tecnológicas (Scherer & Siddiq, 2019). Un alto capital cultural en las familias se asocia positivamente con la CD, refleja interés por el conocimiento y el aprendizaje (Hatlevik et al., 2015; Ren et al., 2022) e inculca altas expectativas sobre la educación (Ren et al., 2022). Mientras que, estudios recientes señalan que un alto nivel cultural y económico puede no ser significativo para mejorar la CD (Casillas-Martín et al., 2021).

Las alfabetizaciones que componen la CD y el entorno socioeconómico también muestran desequilibrios, y señalan la falta de sinergia entre el centro de estudio y la familia (Casillas-Martín et al., 2022). Fraga-Varela et al. (2020) indican que la alfabetización informacional podría desarrollarse de igual manera indistintamente del contexto socioeconómico (ya que el centro juega un papel importante), mientras que, la alfabetización comunicativa es promovida, en mayor medida, por el entorno familiar (modelo a seguir de algún integrante de la familia) o por el mundo de los videojuegos. Así, la falta de acceso a la tecnología, asociada al nivel socioeconómico familiar, puede afectar el desarrollo de CD, lo que podría afectar la capacidad futura del alumnado para competir en la economía digital (Zubillaga & Gortazar, 2020).

Aunado al análisis de la brecha digital asociada al nivel socioeconómico, la literatura sugiere que existe una relación positiva entre el nivel educativo y la CD del alumnado (Siddiq & Scherer, 2019), y que la CD aumenta con la edad (Zhang & Zhu, 2016). No obstante, es escasa la literatura que triangula lo sociodemográfico, académico y tecnológico, por tanto, es en este contexto donde los resultados del presente estudio constituyen uno de los principales aportes.

Se podría decir, la literatura existente es aún limitada para comprender la complejidad entre el nivel socioeconómico y el desarrollo de la CD (Fernández-Mellizo & Manzano, 2018). Sobre esto, las reformas en el Programme for International Student Assessment (PISA), reconocen la importancia de factores contextuales y la igualdad en la educación, tanto en el rendimiento del alumnado como en el acceso y uso de las TIC (OCDE, 2019). Aunque PISA no evalúa la CD directamente, ha considerado medir habilidades digitales en las ediciones 2018 y 2022, como la competencia lectora en entornos digitales y la “alfabetización digital” de los jóvenes en términos de acceso, uso y actitud hacia las tecnologías (OCDE, 2019). Esto resalta la importancia de investigar cómo las desigualdades socioeconómicas pueden afectar en el desarrollo de la CD del alumnado y, por ende, el acceso equitativo a la educación. Datos esenciales para orientar políticas y estrategias educativas en pro de la reducción de la brecha digital en educación.

1.3. Preguntas de investigación:

Se pretende estudiar la influencia del nivel socioeconómico y ciclo de estudio en el desarrollo de la CD del alumnado de educación secundaria de España. En concreto, responder a:

P1. ¿Qué relaciones se pueden identificar entre el nivel de CD autopercebida y el nivel socioeconómico del alumnado de educación secundaria, para cada ciclo de estudio?

P2. ¿Qué relaciones se pueden identificar entre cada uno de los indicadores de la CD autopercebida y el nivel socioeconómico del alumnado de educación secundaria, para cada ciclo de estudio?

P3. ¿Qué variaciones se pueden identificar en el nivel de CD autopercibida del alumnado del mismo nivel socioeconómico, a lo largo de la educación secundaria?

2. MÉTODO

El estudio se realizó bajo una metodología cuantitativa y diseño de encuesta, con la finalidad de identificar relaciones entre la variabilidad de la autopercepción de la CD del alumnado según el nivel socioeconómico y los ciclos de estudio de la Educación Secundaria Obligatoria (ESO) y bachillerato, en España. Se utilizó el cuestionario Digitalis-ESO (Niño-Cortés et al., 2023), versión actualizada del instrumento INCOTIC-ESO de González Martínez et al. (2012). El cuestionario consta de 19 indicadores que miden la CD autopercibida, operacionalizando la definición propuesta por Larraz (2013), mediante una escala Likert de 5 puntos (1: no sé hacerlo –5: sé hacerlo sin dudar) con un coeficiente de fiabilidad de $\alpha=0,908$.

Para conocer el nivel socioeconómico, el cuestionario pregunta sobre el tipo de trabajo que realiza el padre o la madre del alumnado. Esto permitió operacionalizar los datos ocupacionales de la familia y estimar el cálculo del nivel socioeconómico. Datos que se tabularon según la estructura de la Clasificación Internacional Uniforme de Ocupaciones (CIUO-08) y el Índice Socio Económico Internacional (ISEI) (Ganzeboom & Treiman, 2010). Se codificó, en conjunto, con microdatos aportados por la encuesta de ocupación del INE (2022), y se categorizó la información en tres niveles socioeconómicos: nivel bajo ($10 \leq x \leq 29$), nivel medio ($30 \leq x < 49$) y nivel alto (≥ 50). De manera complementaria, se recogieron datos demográficos, como la edad, nivel educativo, repitencia de curso, municipio y la comunidad autónoma del centro educativo.

La recogida de datos se realizó durante el último semestre del curso 2019-2020 y el primer semestre del curso 2020-2021. Los datos se alojaron en el servidor de la universidad, se anonimizaron y volcaron en una hoja de cálculo para crear la base de datos. El proceso de selección de la muestra, elaboración y aplicación de los instrumentos, así como el tratamiento de datos, siguió estrictamente los principios éticos de anonimidad y conformidad de cesión de datos marcados por la British Educational Research Association (BERA) (2018) y contó con la aprobación del comité de ética de la investigación de la universidad (Ref. CEIPSA-2021-PR-0046).

2.1. Muestra

El estudio utilizó un muestreo no probabilístico por conveniencia para seleccionar 130 centros de educación secundaria, de las diversas comunidades autónomas de España. El muestreo fue considerado por las restricciones de la COVID-19 y la conveniencia de acceder y contar con la disponibilidad de participantes en un tiempo determinado. Esto permitió asegurar la viabilidad y eficiencia en la recolección de los datos (Creswell, 2008).

La muestra final estuvo conformada por 4116 alumnos/as de los ciclos de estudio de la ESO y bachillerato. La muestra, está distribuida de acuerdo con el currículo oficial de España (LOE, 3/2020) y corresponde a alumnos/as de: 1º ciclo de la ESO, que cursan 1º, 2º y 3º año escolar, con una edad entre 12 y 15 años; 2º ciclo de la ESO, aquellos que cursan 4º año escolar, con

edad entre 15 y 16 años, y, Bachillerato, 5º y 6º año escolar de la educación secundaria no obligatoria, con una edad entre 16 y 18 años (Ver Tabla 1). De este total de alumnos/as se descartaron 567 datos perdidos, individuos que no se identificaron con ningún nivel socioeconómico o prefirieron no responder. Adicionalmente se eligió de manera aleatoria una submuestra, para asegurar la representatividad en cada nivel socioeconómico y ciclo de estudio, de acuerdo con los datos de la población de España (INE, 2022). Lo que resultó en una muestra final de 2411 estudiantes.

Tabla 1

Distribución de la muestra por ciclo de estudio y nivel socioeconómico.

Ciclo de estudio	Nivel Socioeconómico		
	Bajo (n=375)	Medio (n=1223)	Alto (n=813)
1º ciclo ESO (n=1312)	14%	50%	36%
2º ciclo ESO (n=874)	17%	52%	31%
Bachillerato (n=225)	18%	51%	31%
Muestra total (n=2411)	15%	51%	34%

2.2. Análisis

Se examinaron las distribuciones de la mediana de CD autopercebida global y en detalle para cada indicador que mide la CD, en cada ciclo de estudio (1º y 2º ciclo de la ESO, bachillerato) y niveles socioeconómicos (bajo, medio y alto). Posteriormente se realizaron pruebas estadísticas para analizar las diferencias significativas en la CD autopercebida entre los grupos de nivel socioeconómico para cada ciclo de estudio. Se utilizó la prueba de comparaciones múltiples de Kruskal-Wallis con un nivel de significancia de $p < 0,05$, debido a la naturaleza ordinal de las variables (McDonald, 2014). Procedimiento que se llevó a cabo mediante el programa IBM SPSS Statistics v28 para Windows y Excel.

3. RESULTADOS

3.1. P1. ¿Qué relaciones se pueden identificar entre el nivel global de CD autopercebida y el nivel socioeconómico del alumnado de educación secundaria de España, para cada ciclo de estudio?

El alumnado se percibe como competente digital, en todos los ciclos de la educación secundaria, independientemente de su nivel socioeconómico, con una mediana superior a 3,55 (Ver Tabla 2). No obstante, en detalle, la percepción acerca de su CD varía en función del nivel socioeconómico y el ciclo de estudio.

En términos generales, se observa que a medida que aumenta el nivel socioeconómico, la autopercepción de la CD también aumenta. Esta tendencia se muestra tanto en 1º y 2º ciclo de la ESO donde el alumnado de nivel socioeconómico alto se percibe como los más competentes

digitales, en comparación con sus compañeros/as de nivel socioeconómico medio y bajo, y, el alumnado de nivel socioeconómico bajo se percibe como el menos competente. Mientras que, en bachillerato, la variabilidad de la CD autopercibida es un tanto homogénea entre los diversos niveles socioeconómicos, donde el alumnado de nivel socioeconómico medio destaca como el más competente digital.

Tabla 2

Distribución del promedio de CD para cada ciclo de estudio y nivel socioeconómico. Valores sobre 5.

Ciclo de estudio	Nivel socioeconómico	Mdna	DT
1º ciclo de ESO (n=1312)	Bajo	3,55	0,61
	Medio	3,55	0,62
	Alto	3,70	0,59
CD global 1º ciclo de la ESO:		3,60	0,61
2º ciclo de ESO (n=874)	Bajo	3,60	0,59
	Medio	3,75	0,61
	Alto	3,75	0,59
CD global 2º ciclo de la ESO:		3,73	0,60
Bachillerato (n=225)	Bajo	3,70	0,49
	Medio	3,75	0,56
	Alto	3,73	0,50
CD global bachillerato:		3,75	0,53

Nota: M: promedio de CD autopercibida; DT: desviación típica

El aumento gradual en la puntuación de CD autopercibida entre los niveles socioeconómicos se muestra con mayor fuerza en el 1º ciclo de la ESO, entre 3,55 y 3,70. Mientras que, en 2º ciclo de la ESO la puntuación de la CD autopercibida se mantiene entre el alumnado de nivel socioeconómico medio y alto. No obstante, entre estos ciclos de estudio siempre es el alumnado de nivel socioeconómico alto los que tienen mejor autopercepción de CD. En el caso de bachillerato, son los alumnos/as de nivel socioeconómico medio, los que se perciben como los más competentes digitales, con una puntuación de 3,75, valor que se equipara con la puntuación de CD del alumnado de 2º ciclo de la ESO. En el extremo opuesto, el alumnado de nivel socioeconómico bajo son los que siempre se perciben como los menos competentes digitales, tanto en la etapa de la ESO como en bachillerato.

Para medir estadísticamente las diferencias observadas en la CD autopercibida entre los grupos de niveles socioeconómicos, se realizó la prueba de Kruskal-Wallis (ver Tabla 3), considerando todos los ciclos de estudio en su conjunto (n=2411; $p < .05$). Mientras que, para el 1º ciclo de la ESO, la distribución de la puntuación de CD global es significativamente diferente considerando las tres categorías de nivel socioeconómico (n=1312; $p < .05$), en el 2º ciclo de la ESO ($p=0.116$) y bachillerato ($p=.480$) las diferencias observadas no alcanzan una significancia estadística. Así pues, el nivel global de CD autopercibida en el alumnado de los diversos niveles socioeconómicos es estadísticamente similar en el 2º ciclo de la ESO y bachillerato.

Tabla 3

Análisis del promedio de CD entre parejas de niveles socioeconómicos en cada ciclo de estudio. Resultados de prueba de Kruskal-Wallis.

Ciclo de estudio	Nivel socioeconómico	Estadístico de prueba	Error estándar	Sig.	Sig. Total
1º ciclo de ESO (n=1312)	Bajo-Medio	-10,771	31,307	0,731	<,001
	Bajo-Alto	-92,602	32,584	0,004	
	Medio-Alto	-81,831	22,885	0,000	
2º ciclo de ESO (n=874)	Bajo-Medio	-27,085	23,977	0,259	.166
	Bajo-Alto	-48,697	25,940	0,060	
	Medio-Alto	-21,613	19,377	0,265	
Bachillerato (n=225)	Bajo-Medio	-3,098	11,943	0,795	0,480
	Bajo-Alto	-13,393	12,896	0,299	
	Medio-Alto	-10,295	9,863	0,297	
Toda la muestra (n=2411)	Bajo-Medio	-37,450	41,078	0,362	<,001
	Bajo-Alto	-144,448	43,441	0,001	
	Medio-Alto	-106,998	31,491	0,001	

Nota: Cada fila prueba la hipótesis nula que las distribuciones de la Muestra 1 y la Muestra 2 son iguales. 1º: primer ciclo de la ESO; 2º: segundo ciclo de la ESO; M: promedio de CD; DT: desviación típica.

La comparativa entre parejas de categorías de nivel socioeconómico para cada ciclo de estudio solo arroja resultados significativos en 1º ciclo de la ESO. En este ciclo de estudio hay diferencias significativas entre el alumnado de nivel socioeconómico bajo y alto ($p=.0004$) y entre el nivel socioeconómico medio y alto ($p=0,000$), mientras que la comparativa entre nivel socioeconómico bajo y medio no muestra significancia ($p=0.731$). En resumen, estos resultados sugieren que existe una relación significativa entre el nivel socioeconómico, la CD autopercebida y el ciclo de estudio del alumnado de la ESO, mientras que en bachillerato pareciera no ser significativa.

3.2. P2. ¿Qué relaciones se identifican en cada uno de los indicadores de la CD autopercebida según el nivel socioeconómico en cada etapa de la educación secundaria?

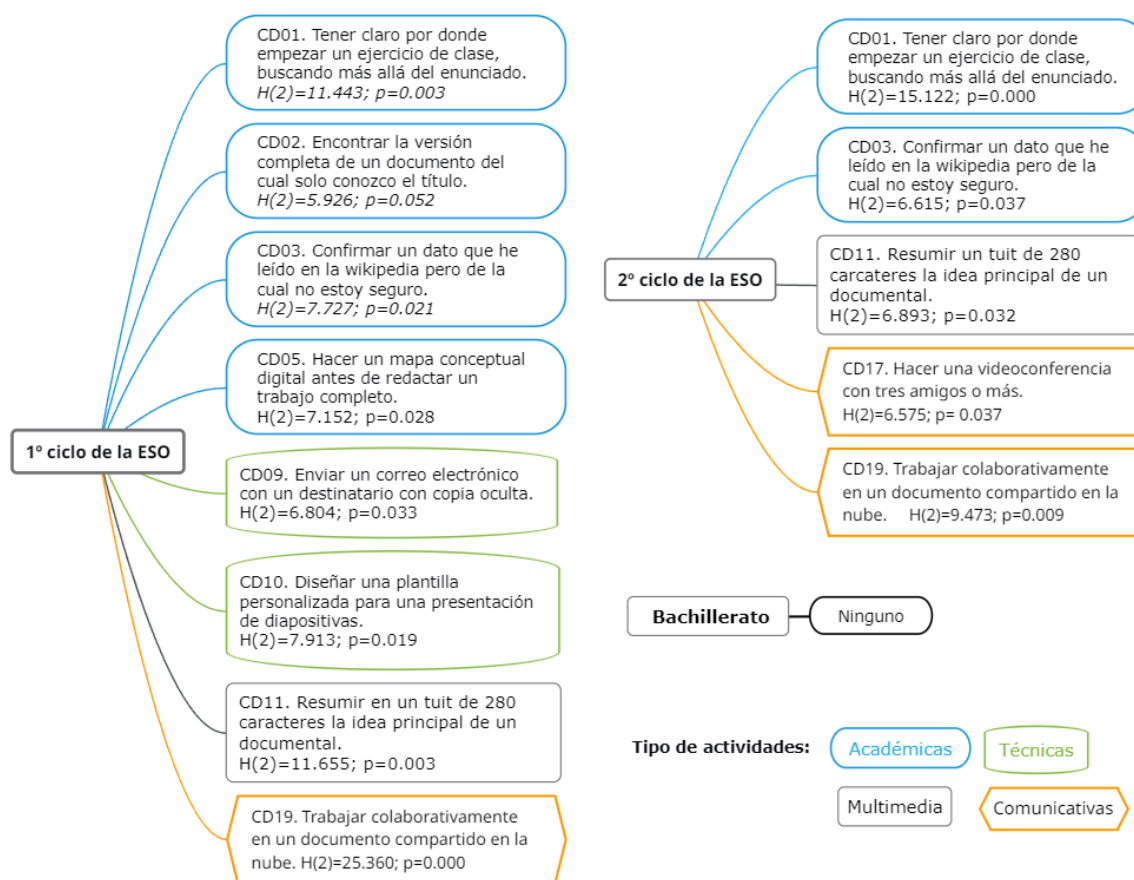
Además de la evaluación de la autopercepción global, se realizó un análisis detallado de cada uno de los indicadores que miden la CD. Los resultados revelan que, en la muestra total ($n=2411$), se encontraron diferencias significativas en 9 de 19 indicadores para los tres ciclos de la educación secundaria en relación con los diferentes niveles socioeconómicos ($p<.05$) (ver Figura 1).

De acuerdo con los resultados obtenidos, en la muestra total ($n=2411$), se encuentran diferencias estadísticamente significativas en el 47% de los indicadores de la CD según el nivel socioeconómico. Concretamente, en estos indicadores el alumnado que se ubica en un nivel

socioeconómico bajo tiende a percibirse como menos competente digital que sus compañeros/as de nivel socioeconómico alto, con un valor de significancia de $p < .05$. Los indicadores en los que se encuentran estas diferencias se pueden relacionar con cuatro tipologías de actividades. En detalle, 4 de 5 indicadores asociados con actividades académicas (CD01, CD02, CD03, CD05), 1 de 5 indicadores de actividades técnicas (CD09, CD10), 2 de 4 indicadores de actividades multimedia (CD11 y CD12), y 2 de 5, de actividades comunicativas (CD17 y CD19). Así pues, los resultados indican que el alumnado de nivel socioeconómico bajo tiene mayor dificultad para desarrollar el aspecto académico de la CD (por ejemplo: búsqueda de información relevante y fiable en la web, conocer herramientas y aplicaciones que les ayuden a analizar y sintetizar información, resumir un tuit de 280 caracteres, trabajar colaborativamente en un documento, entre otros).

Figura 1

Diferencias en los indicadores de la CD según el nivel socioeconómico para cada ciclo de la educación secundaria.



En cuanto al análisis de los indicadores de la CD para cada ciclo de estudio, se observa que es en el 1º ciclo de la ESO donde se encuentra la mayor cantidad de diferencias significativas entre los niveles socioeconómicos. Estas diferencias se corresponden con 8 de 19 indicadores, en concreto, para actividades académicas (CD01, CD02, CD03, CD05), técnicas (CD09, CD10), multimedia (CD11), y comunicativas (CD19). A medida que el alumnado avanza en sus estudios, se evidencia una disminución de las diferencias significativas en los indicadores de la CD

autopercebida según el nivel socioeconómico. En el 2º ciclo de la ESO, estas diferencias son significativas solo en cinco indicadores (CD01, CD03, CD11, CD017, CD19), mientras que, para bachillerato, el análisis estadístico no muestra significancia en ningún indicador. Los resultados se corresponden con lo observado en la CD global y, en consecuencia, ratifican la influencia del nivel socioeconómico en la autopercepción de esta competencia, con mayor fuerza en el primer ciclo de la ESO, y para realizar actividades académicas.

3.3. P3. ¿Qué variaciones se pueden identificar en el nivel de CD autopercebida para el alumnado de un mismo nivel socioeconómico a lo largo de la educación secundaria?

Los hallazgos previos sugieren que el alumnado mejora su percepción de CD a medida que adquiere más conocimientos y habilidades a lo largo de la educación secundaria. Los datos presentados en la Tabla 2, muestran que las puntuaciones de la CD autopercebida aumentan a medida que el alumnado de un mismo nivel socioeconómico avanza en la ESO y bachillerato, equiparándose entre niveles socioeconómicos. Así pues, por ejemplo, para el nivel socioeconómico bajo, la puntuación de CD autopercebida varía entre 3,55 (1º ciclo de la ESO) y 3,70 en bachillerato, mientras que, para el nivel socioeconómico alto, el promedio de CD oscila entre 3,70 (1º ciclo de la ESO) y 3,73 en bachillerato (ver Tabla 2). Es decir, el estudiantado de un mismo nivel socioeconómico mejora su autopercepción de CD a medida que avanza en los estudios indistintamente del nivel socioeconómico, especialmente en el estudiantado de nivel socioeconómico bajo. Estas tendencias son consistentes con los resultados presentados en los apartados anteriores.

Al comparar estadísticamente la puntuación de CD autopercebida entre los ciclos de estudio para un mismo nivel socioeconómico, se confirman algunas de estas diferencias observadas. La puntuación de CD autopercebida no es la misma entre todos los ciclos de estudio y categorías de nivel socioeconómico, para la muestra total de 2411 alumnos/as, con un valor de $p < .05$ (ver Tabla 4). Sin embargo, las diferencias en la puntuación de CD autopercebida en cada categoría de nivel socioeconómico solo son significativas para el nivel socioeconómico medio ($n=1223$), con un valor de $p < 0.00$. Además, en este nivel socioeconómico, la comparación múltiple por parejas de los diferentes ciclos de estudio indica que hay una significancia entre el 1º y 2º ciclo de la ESO ($p < 0.00$) y entre 1º ciclo de la ESO y bachillerato ($p < 0.00$). Las diferencias entre la puntuación de CD para los niveles socioeconómico bajo ($n=375$) y alto ($n=813$), y entre las parejas en los diferentes ciclos de estudio, no son estadísticamente significativas.

Tabla 4

Análisis del promedio de CD entre parejas de ciclos de estudio en cada nivel socioeconómico. Resultados de prueba de Kruskal-Wallis.

Nivel socio económico	Ciclo de estudio	Estadístico de prueba	Error estándar	Sig.	Sig. Total
Bajo (n=375)	1º - 2º	-18,206	11,938	0,127	0,158
	1º - Bachillerato	-29,133	18,858	0,122	
	2º - Bachillerato	-10,927	19,337	0,572	
CD global nivel socioeconómico bajo. Mdna: 3,60. DT: 0,59					
Medio (n=1223)	1º - 2º	-83,703	21,533	0,000	0,000
	1º - Bachillerato	-92,355	35,723	0,010	
	2º - Bachillerato	-8,652	36,819	0,814	
CD global nivel socioeconómico medio. Mdna: 3,65. DT: 0,62					
Alto (n=813)	1º - 2º	-25,137	17,920	0,161	0,173
	1º - Bachillerato	-45,507	30,058	0,130	
	2º - Bachillerato	-20,370	31,497	0,518	
CD global nivel socioeconómico alto. Mdna: 3,70. DT: 0,59					
Toda la muestra (n=2411)	1º - 2º	-121,795	30,385	0,000	0,000
	1º - Bachillerato	-159,993	50,215	0,001	
	2º - Bachillerato	-38,198	52,024	0,463	

Nota: Cada fila prueba la hipótesis nula que las distribuciones de la Muestra 1 y la Muestra 2 son iguales. 1º: primer ciclo de la ESO; 2º: segundo ciclo de la ESO; M: promedio de CD; DT: desviación típica.

4. DISCUSIÓN Y CONCLUSIONES

La presente investigación exploró la influencia del nivel socioeconómico en la CD autopercibida del alumnado de Educación Secundaria en España. Se encontraron diferencias significativas en la autopercepción de CD relacionadas con los niveles socioeconómicos y ciclos de estudio, lo que contribuye a comprender cómo afectan estas variables en el incremento de la brecha digital en la educación.

A lo largo de toda la educación secundaria, se observó una relación significativa entre el nivel socioeconómico, la CD y el ciclo de estudio del alumnado. Aunque las diferencias en CD autopercibida entre los tres niveles socioeconómicos son pequeñas, presentan una tendencia muy marcada, principalmente en los ciclos de la ESO, donde el alumnado de nivel socioeconómico alto tiende a percibir su CD de manera más positiva en comparación con los estudiantes de niveles socioeconómicos más bajos, y, además, a mayor nivel de estudio y nivel socioeconómico, mayor percepción de CD. Mientras que, en bachillerato las puntuaciones de CD de nivel socioeconómico medio y alto se equiparan con el alumnado de 2º ciclo de la ESO.

Por tanto, a lo largo de toda la educación secundaria es el alumnado de nivel socioeconómico bajo quien siempre tiene peor percepción de su CD. Estos resultados concuerdan con estudios previos que indican que el alumnado se percibe con mayor nivel de CD cuando está matriculado en cursos superiores e interactúan en niveles socioeconómicos más altos (Fernández-Mellizo & Manzano, 2018). Además, esta afirmación es consistentemente significativa al incluir el nivel socioeconómico alto en las parejas comparativas de estos, por ejemplo, al comparar nivel socioeconómico bajo-alto, nivel socioeconómico medio-alto, siempre hay diferencias significativas, mientras que, entre nivel bajo y medio, no hay significancia.

Aunado a la relación intrínseca entre las tres variables en estudio, se encontró que la mayor brecha digital en términos socioeconómicos está en los primeros años de la educación secundaria (1º ciclo de la ESO). Siendo escasa la literatura sobre este análisis en el contexto español, hace que estos hallazgos encontrados en el presente estudio sean significativos y destaquen la vulnerabilidad del alumnado de nivel socioeconómico bajo, en particular, en la primera etapa de la ESO. Esto muestra relación con estudios que vinculan la CD del alumnado con el nivel de estudio y la edad (Siddiq & Scherer, 2019; Zhang & Zhu, 2016), lo que sugiere que los alumnos/as en sus etapas iniciales educativas carecen de formación en herramientas digitales y se encuentra en los primeros contactos con la tecnología en la escuela, en consecuencia, se espera que se perciban con menor CD.

El análisis detallado de los indicadores de la CD autopercebida confirma las diferencias en esta competencia y precisa los aspectos con mayor debilidad o fortaleza entre los diferentes niveles socioeconómicos. En concreto, el nivel socioeconómico tiene mayor incidencia en los aspectos académicos de la CD, que, para los técnicos, multimedia y comunicativos, por tanto, el alumnado de nivel socioeconómico bajo presenta mayor dificultad para realizar actividades académicas en línea. Estos resultados difieren de algunas investigaciones previas (Fraga-Varela et al., 2020) que sugieren que las actividades académicas en línea se benefician de la interacción en la escuela, mientras que las actividades comunicativas y multimedia están más relacionadas con el entorno familiar y socioeconómico. No obstante, el presente estudio pone de manifiesto que el nivel socioeconómico ejerce mayor incidencia en el alumnado de 1º ciclo de la ESO, en mayor medida para realizar actividades digitales de índole académica. Este hallazgo sugiere una posible asociación entre la formación recibida por los alumnos/as y el respaldo de recursos tecnológicos del centro educativo. Además, se infiere que estos resultados guardan una relación con el progreso de la CD a medida que el alumnado avanza en su trayectoria académica, y que esto sucede indistintamente del nivel socioeconómico.

Los resultados de las diferencias en la autopercepción de la CD dentro de un mismo nivel socioeconómico, a lo largo de la educación secundaria, indican que dicha percepción aumenta a medida que el alumnado avanza en sus estudios, y esta tendencia se muestra con mayor fuerza para el nivel socioeconómico bajo. Estos resultados evidencian la significativa relación entre el nivel socioeconómico y CD autopercebida, y destacan que el alumnado de entornos socioeconómicos desfavorecidos se enfrenta a desventajas debido a la carencia de respaldo social para adquirir habilidades tecnológicas, tal como lo señalan Ren et al. (2022) y Casillas-Martín et al. (2021). Además, en este contexto, se observa la importancia de la formación a lo largo de la educación secundaria (ciclo de estudio) que favorece y mejora la CD autopercebida, en todos los niveles socioeconómicos.

En conclusión, se observa una brecha digital en la educación secundaria que muestra implicaciones para el desarrollo de la CD del alumnado. Esta brecha señala al entorno socioeconómico en el que interactúa el alumnado como un factor que restringe su progreso en condiciones equitativas y limita la adquisición de habilidades necesarias para su integración en la sociedad digital. Por tanto, es esencial implementar políticas educativas y prácticas docentes que enfoquen esfuerzos en la reducción de la brecha digital, especialmente en los primeros años de la ESO y en centros educativos que atienden a entornos desfavorecidos. Esto incluye la implementación de programas de alfabetización digital adaptados, la mejora de la infraestructura tecnológica en las escuelas y la capacitación continua del profesorado en el uso de estrategias didácticas digitales.

Además, la interrelación entre el contexto social y la CD sugiere que las intervenciones deben ser proactivas y sostenidas, enfocadas en la formación docente y la provisión de recursos a los estudiantes en riesgo. Por tanto, como línea de investigación futura, a nivel de centro de estudio, se sugiere fortalecer e implementar políticas formativas institucionales y planes de acción centrados en cultura y educación digital, en concreto, dirigido al alumnado de niveles socioeconómicos más bajos y en los primeros años de la ESO, con el objetivo de priorizar y abordar la brecha digital en la educación. Finalmente, es importante mencionar que una de las principales limitantes del estudio radica en la muestra reducida de estudiantes en el segundo ciclo de la ESO y en bachillerato.

5. AGRADECIMIENTOS

Este estudio se realizó en el marco del proyecto MindGAP, financiado por Fundación Bancaria “La Caixa” (“LCF/PR/SR19/52540001”). Además, la autora Luz Mayra Niño-Cortés es becaria de Fundación Carolina.

6. REFERENCIAS

- Baeza-González, A., Lázaro-Cantabrana, J. L., & Sanromà-Giménez, M. (2022). Evaluación de la competencia digital del alumnado de ciclo superior de primaria en Cataluña: [Assessment of primary education students' digital competence in Catalonia]. *Pixel-Bit: Revista de Medios y Educación*, 64, 265-298. <https://doi.org/10.12795/pixelbit.93927>
- British Educational Research Association [BERA] (2018). Ethical Guidelines for Educational Research, 4th edition, London. <https://www.bera.ac.uk/researchers-resources/publications/ethicalguidelines-for-educational-research-2018>
- Casillas-Martín, S., Cabezas-González, M., García-Valcárcel Muñoz-Repiso, A., & Basilotta-Gómez-Pablos, V. (2021). *Modelos de mediación sociofamiliares en el desarrollo de la competencia digital*. *Revista electrónica de investigación educativa*, 23, e26, 1-18. <https://doi.org/10.24320/redie.2021.23.e26.3839>
- Casillas-Martín, S., Cabezas-González, M., & García-Valcárcel Muñoz-Repiso, A. (2022). Influencia de variables sociofamiliares en la competencia digital en comunicación y colaboración: [Influence of socio-familial variables on digital competence in

- communication and collaboration]. *Pixel-Bit. Revista De Medios Y Educación*, 63, 7–33. <https://doi.org/10.12795/pixelbit.84595>
- Colás Bravo, M.P., Conde Jiménez, J. y Reyes de Cózar, S. (2017). Competencias digitales del alumnado no universitario. *RELATEC. Revista Latinoamericana de Tecnología Educativa*, 16 (1). <http://dx.medra.org/10.17398/1695-288X.16.1.7>
- Comisión Europea. (2020). Plan de acción de educación digital 2021-2027. <https://education.ec.europa.eu/es/focus-topics/digital-education/action-plan>
- Comisión Europea (Centro Común de Investigación). Brande, L., Carretero, S., Vuorikari, R. (2017). DigComp 2.0: el marco de competencias digitales para los ciudadanos, Oficina de Publicaciones. <https://data.europa.eu/doi/10.2791/11517>
- Creswell, J. W., & Guetterman, T.C. (2013). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Pearson.
- Fernández-Mellizo, M. & Manzano, D. (2018). Anàlisi de les diferències en la competència digital dels alumnes espanyols. *Papers: revista de sociologia*, 103 (2), 175-198. <https://raco.cat/index.php/Papers/article/view/336759>
- Fraga-Varela, F., Vila-Couñago, E., & Pernas-Morado, E. (2019). Aprendizajes ausentes en la Competencia Digital de preadolescentes: un estudio de casos pertenecientes a contextos socioculturales desfavorables. *Revista de Educación a Distancia (RED)*, 19(61). <https://doi.org/10.6018/red/61/04>
- Fernández-Abuín, J. P. (2016). La adquisición y desarrollo de la competencia digital en alumnos de educación secundaria. Estudio de caso. *Cuadernos de Investigación Educativa*, 7(2), 83–98. <https://doi.org/10.18861/cied.2016.7.2.2612>
- Ganzeboom, H. B., & Treiman, D. J. (2010). Occupational status measures for the new International Standard Classification of Occupations ISCO-08; with a discussion of the new classification. In Annual Conference of International Social Survey Programme, Lisbon.
- Gisbert, M., González, J. y Esteve, F. (2016). Competencia digital y competencia digital docente: una panorámica sobre el estado de la cuestión. *RIITE. Revista Interuniversitaria de Investigación en Tecnología Educativa*, 0, 74-83. <https://doi.org/10.6018/riite2016/257631>
- González-Benito, A., Gutiérrez-de-Rozas, B., & Otero-Mayer, A. (2023). La brecha digital como factor de exclusión social: situación actual en España. *Cuestiones Pedagógicas. Revista De Ciencias De La Educación*, 2(31), 103–128. <https://doi.org/10.12795/CP.2022.i31.v2.06>
- González-Rodríguez, C., & Urbina-Ramírez, S. (2020). Análisis de instrumentos para el diagnóstico de la competencia digital. *RiiTE Revista Interuniversitaria de Investigación en Tecnología Educativa*, 9. <https://doi.org/10.6018/riite.411101>

- González-Martínez, J., Espuny Vidal, C., de Cid Ibeas, M. J., & Gisbert Cervera, M. (2012). INCOTIC-ESO. Cómo autoevaluar y diagnosticar la competencia digital en la escuela 2.0. *Revista de Investigación Educativa*, 30(2), 287–302. <https://doi.org/10.6018/rie.30.2.117941>
- Hatlevik, O. E., Guðmundsdóttir, G. B., Loi, M. (2015). Examining factors predicting students' digital competence. *Journal of Information Technology Education: Research*, 14, 123-137. <http://www.jite.org/documents/Vol14/JITEV14ResearchP123-137Hatlevik0873.pdf>
- INE (2022). Encuesta de condiciones de vida. Instituto Nacional de Estadística. https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176807&menu=resultados&idp=1254735976608#!tabs-1254736194793
- Larraz, V. (2013). *La competencia digital a la Universitat* [Tesis doctoral, Universitat d'Andorra]. TDX. <https://cutt.ly/l88xp8p>
- Ley Orgánica de Educación (LOE) 3/2020, de 29 de diciembre, por la que se modifica la Ley Orgánica 2/2006, de 3 de mayo. BOE, 340, de 30 de diciembre de 2020. <https://www.boe.es/eli/es/lo/2020/12/29/3>
- Matamala Riquelme, C. (2015). Factores predictivos de las competencias TIC en alumnos chilenos de secundaria. *Revista Iberoamericana De Educación*, 67(1), 121-136. <https://doi.org/10.35362/rie671267>
- Márquez Fernández, Juan Miguel (2023). *Monográficos España Digital: Competencias Digitales*. Observatorio Nacional de Tecnología y Sociedad (ONTSI). Ministerio de Asuntos Económicos y Transformación Digital 02/2023. <https://biblio.ontsi.red.es/cgi-bin/koha/opac-detail.pl?biblionumber=7326>
- Niño-Cortés, L.M., Grimalt-Álvaro, C., Lores-Gómez, B., & Usart, M. (2023). Brecha digital de género en secundaria: diferencias en competencia autopercebida y actitud hacia la tecnología. [The digital gender gap in secondary school: differences in self-perceived competence and attitude towards technology]. *Educación XX1*, 26(2), 299-322. <https://doi.org/10.5944/educxx1.34587>
- OECD (2001), "Understanding the Digital Divide", *OECD Digital Economy Papers*, No. 49, OECD Publishing, Paris, <https://doi.org/10.1787/236405667766>.
- OECD (2019), *PISA 2018 Results (Volume I): What Students Know and Can Do*. PISA, OECD Publishing, Paris, <https://doi.org/10.1787/5f07c754-en>
- Ren, W., Zhu, X., & Yang, J. (2022). The SES-based difference of adolescents' digital skills and usages: An explanation from family cultural capital. *Computers & Education*, 177, 104382. <https://doi.org/10.1016/j.compedu.2021.104382>
- Scherer, R., & Siddiq, F. (2019). The relation between students' socioeconomic status and ICT literacy: Findings from a meta-analysis. *Computers & Education*, 138, 13-32. <https://doi.org/10.1016/j.compedu.2019.04.011>

- Soria Pérez, Y., Sebastiani Elías, Y., & Lujano Ortega, Y. (2022). Análisis comparativo de la competencia digital en estudiantes de educación secundaria. En *Desafíos y perspectivas de la educación* (61-71). Editorial Idicap Pacífico. <https://doi.org/10.53595/eip.006.2022.ch.4>
- Unesco (18 de julio de 2022). *El derecho a la educación. Reducir la brecha digital y garantizar la protección en el ciberespacio*. <https://www.unesco.org/es/right-education/digitalization>
- Van Deursen, A. J., & Van Diepen, S. (2013). Information and strategic Internet skills of secondary students: A performance test. *Computers & Education*, 63, 218-226. <https://doi.org/10.1016/j.compedu.2012.12.007>
- Vila-Couñago, E., Rodríguez-Groba, A., & Martínez-Piñeiro, E. (2020). La competencia digital de los preadolescentes gallegos/as antes de la pandemia: ¿y ahora qué?. *Revista Latinoamericana de Tecnología Educativa - RELATEC*, 19(2), 9-27. <https://doi.org/10.17398/1695-288X.19.2.9>
- Zhang, H. y Zhu, C. (2016). A Study of Digital Media Literacy of the 5th and 6th grade primary students in Beijing. *The Asia-Pacific Education Researcher*, 25(4), 579-592. <http://dx.doi.org/10.1007/s40299-016-0285-2>
- Zubillaga y Gortazar (2020). COVID-19 Y EDUCACIÓN: problemas, respuestas y escenarios. *Fundación Cotec para la Innovación*. <http://www.cop.es/uploads/PDF/COVID19-EDUCACION-PROBLEMAS-RESPUESTAS.pdf>.


Para citar este artículo:

Niño-Cortés, L. M., Grimalt-Álvaro, C., Sanabria, I. Z. y Usart, M. ¿Cómo influye el nivel socioeconómico en la autopercepción de la competencia digital del alumnado de educación secundaria en España? *EduTec, Revista Electrónica de Tecnología Educativa*, 89, 174-189. <https://doi.org/10.21556/edutec.2024.89.3277>



Redes sociales y aprendizaje móvil en Educación Física: Creación colaborativa de principios de diseño

Social networks and mobile learning in Physical Education: Creating design principles collaboratively

-  Francisco José Montiel-Ruiz, franciscojose.montiel2@um.es Universidad de Murcia (España)
-  Antonio Calderón, antonio.calderon@ul.ie University of Limerick (Irlanda)
-  María del Mar Sánchez-Vera, mmarsanchez@um.es, Universidad de Murcia (España)
-  Isabel María Solano-Fernández, imsolano@um.es, Universidad de Murcia (España)

Resumen

La Educación Física debe centrarse en el desarrollo de los dominios físico, cognitivo, social y afectivo del estudiante, sirviéndose de las potencialidades de las metodologías activas y el uso de recursos tecnológicos. Para ello, deben existir principios de diseño que guíen las estrategias pedagógico-digitales a implementar por parte de los docentes del área. El objetivo de este estudio es la creación de estos principios mediante la colaboración de profesorado y alumnado a través de un proyecto de intervención colaborativo que tuvo lugar en un contexto real de aula. El método de investigación se corresponde con el de Investigación Basada en Diseño en un proceso iterativo de tres años de duración para el diseño, implementación y evaluación de dicha intervención educativa. Participan diez profesores que imparten docencia a 17 grupos de Educación Secundaria de los niveles de 4º ESO y 1º de Bachillerato (un total de 372 estudiantes, siendo 189 chicas y 184 chicos). Los resultados muestran una serie de principios de diseño para implementar estrategias pedagógico-digitales. Destaca el uso de entornos virtuales de enseñanza-aprendizaje específicos para Educación Física que combinen sistemas de gestión de contenidos, redes sociales y monitorización de actividad física, así como la inclusión de narrativas de gamificación.

Palabras clave: Investigación Basada en Diseño, Educación Física, aplicaciones móviles, redes sociales.

Abstract

Physical Education should focus on the development of the student's physical, cognitive, social, and affective domains, using, among others, the potential of active methodologies and the use of technological resources. To this end, there must be design principles that guide the pedagogical-digital strategies to be implemented by teachers in the area. The aim of this study is the creation of these principles through the collaboration of teachers and students through a collaborative intervention project that took place in a real classroom context. The research method corresponds to Design-Based Research, in a three-year iterative process for the design, implementation, and evaluation of this educational intervention. The participating teachers teach 17 groups of Secondary Education at the 4th ESO and 1st Baccalaureate levels (a total of 372 students, 189 girls and 184 boys). The results of this study show a series of design principles for implementing pedagogical-digital strategies. Gamification narratives to guide teaching programming stand out, as well as the use of virtual teaching-learning environments specific to Physical Education that combine content management systems, social networks and physical activity monitoring.

Keywords: Design Based Research, Physical Education, mobile applications, social networks.



1. INTRODUCCIÓN

Actualmente la Educación Física (en adelante, EF) parece estar reducida a la perspectiva patogénica en la que el objetivo del área es la realización de actividad física para la prevención del sedentarismo y la obesidad (Quennerstedt, 2019). Sin embargo, la EF debe ser una pieza clave para el desarrollo de los aspectos físico, cognitivo, social y afectivo (Bailey, 2006). En esta línea, surge la visión salutogénica (Super et al., 2021) en la que se busca promover experiencias significativas en EF a través de la interacción social del alumnado, el uso de desafíos y retos, el desarrollo de aprendizajes significativos y la necesidad de diversión (Quennerstedt et al., 2024).

Una visión que entiende la realización de actividad física desde la motivación intrínseca derivada del significado personal, satisfacción y alegría, en mayor medida que la motivación extrínseca que antepone los posibles beneficios de pérdida de peso o prevención de enfermedades (Beni et al. 2019). En esa búsqueda por el desarrollo de los dominios que engloban el aprendizaje de la EF (físico, cognitivo, social y afectivo) se apuesta por la implementación de metodologías activas relacionadas con la gamificación, el Aprendizaje Basado en Proyectos, el aprendizaje cooperativo y el aula invertida, apoyadas en el uso de distintas tecnologías digitales (en adelante, TD) (Cañizares y Carbonero, 2018; Casey y Fernández-Río, 2019; Contreras y Gutiérrez, 2017; Fernández-Río et al., 2020; Østerlie y Kjelaas, 2019; Sargent y Casey, 2020).

El uso de los medios tecnológicos debe enfocarse en la consecución de los objetivos de aprendizaje a través de estrategias metodológicas pedagógico-digitales bajo el paraguas de la Tecnología Educativa (en adelante, TE) como disciplina (Prendes et al., 2018; Sánchez-Vera, 2023). La necesidad de volver a poner el énfasis en esta disciplina ha sido señalada por autores como Cabero (2016), Martínez (2016), Sancho et al. (2018) y Castañeda et al. (2020). Las últimas investigaciones resaltan la importancia de investigar en TE, a partir del método de Investigación Basada en el Diseño por su capacidad de producir principios de diseño que retroalimentan y mejoran el proceso de intervención (de-Benito y Salinas, 2016; Reeves, 2006; van den Berg y Bozalek, 2024).

La investigación en TE debe vertebrarse a través de teorías y modelos que le ofrezcan solidez (Hew et al., 2019), como la teoría de la autorregulación del aprendizaje (Self-Regulated Learning o SRL) (Zimmerman y Schunk, 2011); el modelo TPACK, el cual integra el conocimiento del contenido de la asignatura, el conocimiento de la tecnología que se utiliza y el conocimiento pedagógico de las estrategias didácticas adecuadas (Mishra y Koehler, 2006); o, el modelo SAMR que clasifica la integración tecnológica en educación en cuatro niveles: sustitución, aumento, modificación y redefinición. Discurrir por estos cuatro niveles implica evolucionar de una situación en la que únicamente se sustituya una herramienta tradicional por una tecnología hasta la redefinición y creación de nuevas tareas gracias al uso de la tecnología (Puentedura, 2020).

En los procesos de aprendizaje mediados por TD en EF se han utilizado diferentes herramientas y entornos virtuales de aprendizaje tratando de partir de los principios del trabajo colaborativo. Principalmente, se utilizan redes sociales para formar una comunidad colaborativa de aprendizaje, cuya aplicabilidad específica para el área de EF se ha analizado en diversas investigaciones (Carpenter y Linton 2018; Casey et al., 2017; Goodyear et al, 2019).

Los diseños de intervención mediados por TD en EF deben plantearse desde una premisa en la que la tecnología no sea una mera herramienta de gestión del conocimiento o comunicación virtual sino que también pueda ser un aliciente para el movimiento y la práctica de actividad física del alumnado (Monguillot et al., 2015). Esto deriva en el aprovechamiento de las características de algunas aplicaciones móviles como pueden ser los sensores de movimiento, las cámaras de fotos/vídeo o el sistema GPS, utilizados para el desarrollo de contenidos específicos de EF (Díaz-Barahona, 2020; Moreno et al., 2015; Papastergiou et al., 2021; Pulido et al., 2016; Victoria, 2020).

En definitiva, este estudio tiene como objetivo la creación de principios básicos para el diseño de estrategias pedagógico-digitales en el ámbito de la EF a partir de la colaboración con los docentes participantes, atendiendo al diseño participativo en TE de Gros y Durall (2020), teniendo en cuenta el cumplimiento de la normativa curricular y conduciendo a la inclusión de las TD como facilitadoras de la actividad física del alumnado.

2. MÉTODO

Este estudio se enmarca en la Investigación Basada en el Diseño (en adelante IBD), atendiendo a Reeves (2006). Nace de un enfoque que procura abordar problemas educativos complejos en contextos reales en colaboración con otros docentes e integrar principios de diseño con avances tecnológicos para dar soluciones plausibles. Todo ello mediante una investigación rigurosa y reflexiva para probar y refinar entornos de aprendizaje innovadores y definir nuevas estrategias pedagógico-digitales (Wang y Hannafin, 2005; Reeves, 2006; de-Benito y Salinas, 2016; Valverde-Berrocoso, 2016; van den Berg y Bozalek, 2024;).

2.1. Intervención educativa

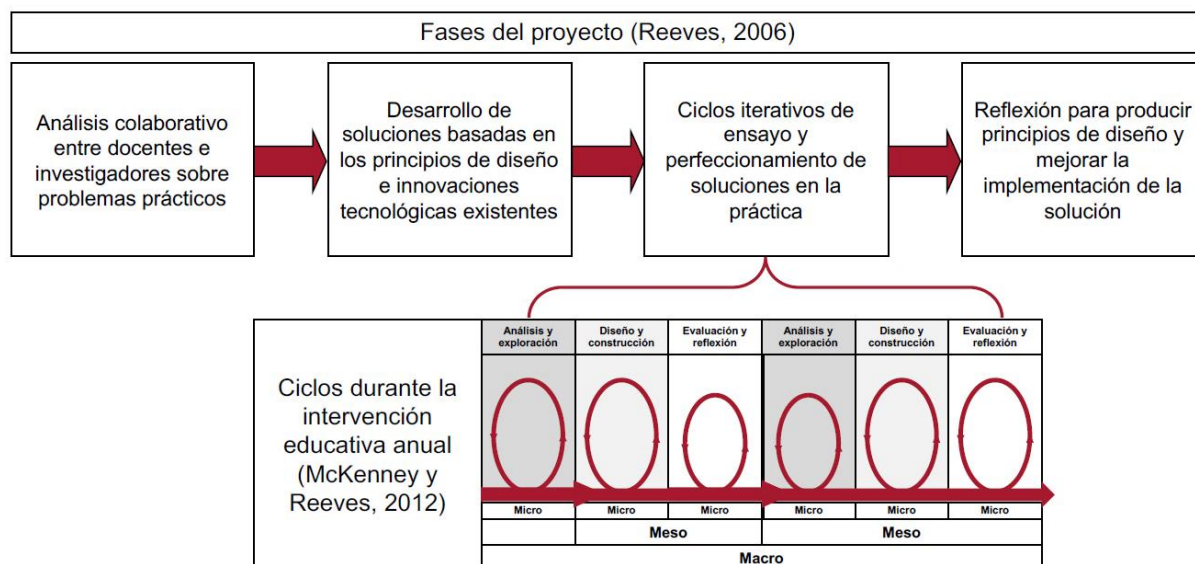
La finalidad de nuestra investigación se construye sobre la colaboración y retroalimentación de docentes y alumnado. Esa colaboración se lleva a cabo en la IBD a través de cuatro fases interrelacionadas dadas a conocer por Reeves (2006). Para el desarrollo de este estudio, durante la tercera fase en la que se llevan a cabo diferentes ciclos de implementación, se ha complementado con el modelo de ciclos integrados de McKenney y Reeves (2012) como se observa en la Figura 1.

La intervención educativa llevada a cabo para este estudio está ligada a la colaboración con docentes en ejercicio, cuya participación en las distintas fases conllevó un proceso de tres años basado en el diseño y rediseño de una programación didáctica, su implementación y la posterior evaluación y definición de principios de diseño (Montiel-Ruiz, 2020). Las fases de investigación supusieron multitud de iteraciones de análisis, cocreación y evaluación por lo que a continuación se sintetizan las acciones fundamentales llevadas a cabo durante todo el proceso.

La primera fase se llevó a cabo durante el primer año de trabajo de este estudio, en el que se desarrolló el análisis colaborativo entre docentes e investigadores sobre problemas prácticos en EF en la etapa de secundaria, así como los desafíos y retos que supone la inclusión de las TD en el área.

Figura 1

Integración de ciclos de intervención de McKenney y Reeves (2012) en las fases de IBD de Reeves (2006)



En la segunda fase los investigadores diseñaron una propuesta preliminar de programación didáctica en la que las unidades formativas estaban integradas por diferentes retos educativos teniendo en cuenta la normativa curricular del área. Posteriormente, los docentes participantes colaboraron en el rediseño y adaptación de dichos retos. La propuesta final cocreada con el profesorado se presentó al alumnado con una narrativa de gamificación compuesta por diferentes actividades y retos, estructurada en etapas, como se observa en la Figura 2.

Figura 2

Estructuración lineal de los retos en la narrativa de gamificación para el alumnado

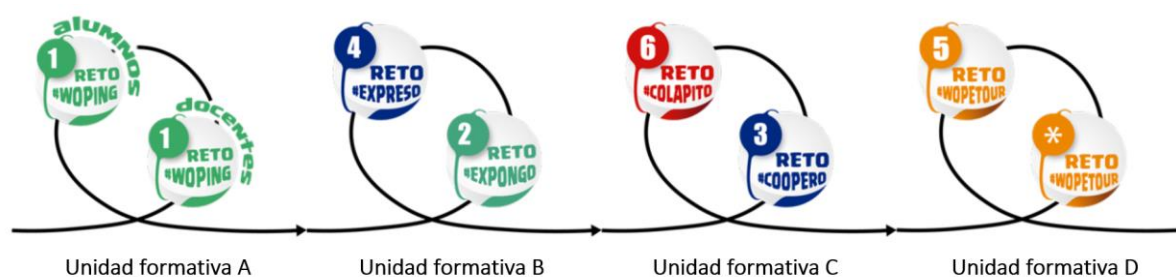


Las tres etapas divididas en colores se corresponden con cada uno de los trimestres que conforman el curso escolar, vertebradas por el hilo conductor de la narrativa de gamificación que se creó. Un primer trimestre de “entrenamiento”, un segundo trimestre de “juegos de invierno” y un tercero de “juegos de la juventud”. En definitiva, una estructuración lineal de los retos, presentada a los alumnos conforme se implementaron durante el curso escolar.

Sin embargo, esto no se corresponde con la estructuración de ciclos iterativos de ensayo de la IBD, que se observa en la tercera fase propuesta para este estudio. Para ello, los retos se agruparon en diferentes unidades formativas en función de los estándares de aprendizaje abordados. En la Figura 3 se observa dicha estructura.

Figura 3

Estructuración cíclica de los retos en el proceso de mejora en cada unidad formativa



La red social educativa Edmodo fue la elegida para la intervención debido a que es una plataforma en la que destaca la capacidad de generar un entorno seguro para docentes y estudiantes. Pese a que ya no se encuentra activa, se recomienda la utilización de aplicaciones con características similares, como puede ser Google Classroom (Dimo et al., 2024; Lobo, 2023) y que además cuenta en la actualidad con las licencias educativas adquiridas por las consejerías de educación de las provincias en las que se llevó a cabo este estudio. Permitió durante los años de intervención mantener una estricta privacidad de los usuarios creando perfiles anonimizados sin estar asociados a correos electrónicos ni ningún tipo de dato personal del alumno e imposibilitando las conversaciones privadas entre alumnos (Montiel-Ruiz, 2020; Montiel-Ruiz, 2022).

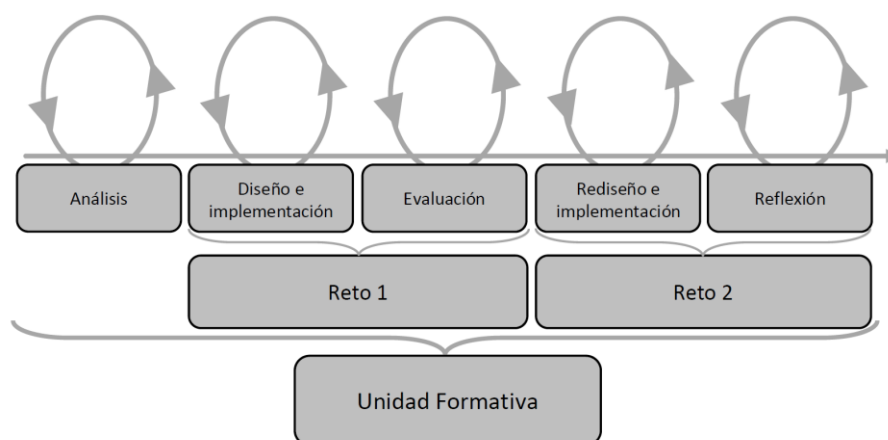
Así, se lanzó el reto “woping” de iniciación al uso de las TD que se utilizaron durante el curso, como la propia red social educativa. Fue una primera conexión y formación, tanto para docentes como para alumnos, relacionado con los estándares de aprendizaje transversales vinculados a la competencia digital. El reto “woping para docentes” se realizó en el curso anterior a la intervención, durante el proceso de diseño y cocreación de dicha programación anual. Así, los microciclos de análisis, retroalimentación y mejora para el reto “woping de los alumnos” se realizaron de un curso para otro. En el resto de las unidades formativas estos microciclos ocurrieron durante el propio curso académico entre trimestres. Así, se observa como el segundo reto denominado “expongo” y llevado a cabo en el primer trimestre se relaciona con el cuarto reto denominado “expreso” e implementado en el segundo trimestre. De igual manera, el tercer reto lineal denominado “coopero” e implementado en el segundo trimestre tenía relación con el sexto reto lineal “colapito”, llevado a cabo en el tercer trimestre. Por último, el reto “wopetour” destinado a la realización de actividad física en el tiempo libre

del alumnado se llevó a cabo durante todo el segundo y tercer trimestre, permitiendo los microciclos de mejora durante el periodo de vacaciones.

Todo esto facilitaba que existiera un proceso de análisis dentro de cada unidad formativa (Figura 4) para evaluar las actividades abordadas en el primer reto y poder rediseñar e introducir mejoras en las actividades pertenecientes al segundo reto de la misma unidad formativa. Tomando como ejemplo la Unidad Formativa C, en primer lugar se llevaron a cabo juegos cooperativos durante el segundo trimestre (reto 3: “coopero”) que se analizaron y evaluaron permitiendo tomar decisiones de mejora para llevar a cabo los juegos de colaboración y competición (reto 6: “colapito, juegos en los que colaboro y compito”). De igual manera ocurrió en la Unidad Formativa B al desarrollar contenidos de expresión oral y corporal.

Figura 4

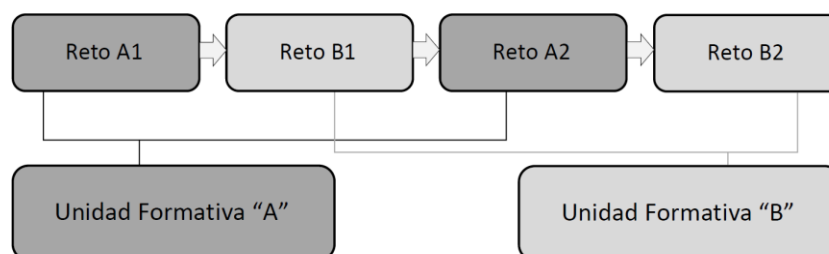
Ciclos correspondientes a cada unidad formativa



Así, en la temporalización lineal se intercalaban actividades y retos de diferentes unidades formativas, como se observa en la Figura 5, para que pudiera existir el tiempo necesario en el que llevar a cabo el proceso de retroalimentación y mejora de las actividades siguientes.

Figura 5

Retos escalonados entre unidades formativas para el proceso de mejora



Una vez finalizada la intervención educativa, comenzó la cuarta fase. La reflexión y colaboración con los docentes y estudiantes para producir principios de diseño destinados al desarrollo de estrategias pedagógico-digitales en EF.

2.2. Muestra

Se ha trabajado con un muestreo no probabilístico por conveniencia. La muestra participante ha sido de 17 grupos de Educación Secundaria de los niveles de 4º ESO y 1º de Bachillerato (un total de 373 estudiantes, siendo 189 chicas y 184 chicos) y 10 docentes de EF procedentes de ocho centros educativos, siete de la Región de Murcia y uno Tenerife, que se ofrecieron voluntarios a participar tras conocer la invitación formal. Dicha invitación se compartió a través de la etiqueta “#edufis” en redes sociales, que es la más popular y seguida en el área. Además, dichos centros voluntarios debían cumplir con los criterios de selección (1) centro de Educación Secundaria; y (2) docentes de EF interesados en participar en la investigación. Asimismo, el alumnado debía cumplimentar una declaración de consentimiento informado, proceso que se detalla posteriormente.

2.3. Instrumentos

En esta investigación se han utilizado diferentes instrumentos para la recogida de datos:

- Grupo focal realizado con el profesorado participante de la experiencia. Para Trigueros et al. (2018), el objetivo central de un grupo focal es “recopilar información sobre un tópico específico, pero se hace en forma grupal debido a que se entiende que la participación individual se puede enriquecer a través de la interacción entre los miembros de dicho grupo” (p.110). Se utiliza un guion con diversas cuestiones para moderar este grupo focal con el objetivo de centrar la temática en la búsqueda de resultados finales. Las preguntas utilizadas fueron revisadas y validadas por un grupo de tres expertos, dos de ellos en TE y uno de ellos en EF.
- Registro de los datos recogidos en un grupo de WhatsApp docente activo durante toda la fase de implementación. WhatsApp puede ser una herramienta muy potente para tener en cuenta en la investigación, entre otras cosas, porque agiliza la comunicación y la colaboración (Jailojev et al., 2021). Nos ha dado la posibilidad de generar un espacio virtual de trabajo entre los docentes participantes como medio de interacción y retroalimentación para el rediseño permanente del entorno de aprendizaje.
- Registro de impresiones de los docentes en formato carta en el que expresaban sus vivencias en el desarrollo del proyecto con el objetivo de exponer los puntos fuertes e inconvenientes de este. Este instrumento se encuadra dentro de los documentos personales de carácter narrativo, es decir, se trata de testimonios en los que los participantes narran los hechos guiados por sus sensaciones, juicios y propia personalidad (Trigueros et al., 2018).
- Entrevistas grupales con el alumnado. Las preguntas de la entrevista fueron validadas mediante la técnica de juicio de expertos, que es una estrategia con grandes potencialidades para ser aplicada en el contexto de estudios sobre tecnología en la educación (Cabero y Llorente, 2013). Se revisaron y validaron por un grupo de tres expertos, dos de ellos en TE y uno de ellos en EF, que tras un proceso de análisis, comprensión y adecuación a partir de algunos cambios dieron lugar a las entrevistas finales. Se pretendía tener en cuenta la perspectiva del alumnado a la hora de refinar los principios de diseño y conocer sus opiniones respecto a los efectos del programa.

2.4. Procedimiento de investigación

La intervención educativa recibió informe favorable por unanimidad desde la Comisión de Ética de Investigación de la Universidad de Murcia. Se entregó al alumnado una “hoja de información a la persona participante” con los detalles necesarios para cumplimentar la “declaración de consentimiento informado”. La recogida de datos se llevó a cabo de múltiples formas. Durante todo el proceso de implementación del proyecto estuvo en funcionamiento el “Grupo de docentes en WhatsApp” que se exportó al finalizar resultando en documento textual (.txt) con 2796 mensajes de chat.

Además, una vez completada la intervención educativa, se realizó un grupo focal virtual con los docentes resultando en un archivo de vídeo (.mp4) de una hora y diecinueve minutos. Los profesores también cumplimentaron una carta narrativa de su experiencia derivando en 9 documentos de texto con una media de 200 palabras. Por último, la entrevista grupal al alumnado se aplicó a una muestra incidental de estudiantes procedentes de todos los grupos, realizadas en la jornada final del proyecto posterior a la finalización del curso escolar, en un clima de aula agradable, donde la honestidad de las respuestas ya no sería influenciado por posibles repercusiones en la calificación de la asignatura. Se obtuvieron dos archivos de audio (.mp3) con una duración total de veintidós minutos.

2.5. Análisis de datos

El análisis de datos ha sido de carácter cualitativo y se ha desarrollado a partir del software NVivo 12. La estrategia utilizada ha sido el análisis de contenido a partir de la información recogida en los instrumentos descritos anteriormente, y siguiendo el procedimiento de categorización y codificación establecido por autores como Miles y Huberman (1984). El uso de estos instrumentos ha permitido la triangulación de datos como estrategia para la investigación educativa. Como afirman Aguilar y Barroso (2015), “la triangulación es un procedimiento que nos va a permitir obtener un mayor control de calidad en el proceso de investigación y garantía de validez, credibilidad y rigor en los resultados alcanzados” (p.73).

A la hora de realizar el análisis, se ha trabajado en la conjunción de diferentes métodos de análisis de documentos cuya prioridad ha sido el sentido de los textos y cuyos análisis está integrado en un proceso cíclico o circular de refinamiento continuado frente a la linealidad del análisis de datos cuantitativos. Es decir, “la intención no es transformar la palabra, el sentimiento, la creencia o las teorías de nuestros participantes en dato numérico. Muy al contrario, será la interpretación y la búsqueda de los porqués lo que marcará la hoja de ruta” (Trigueros et al., 2018, p. 38).

En este proceso se han seguido tres pasos fundamentales: la reducción de datos para la selección y condensación mediante codificaciones y categorías. Se realizó a partir de las recomendaciones de Braun y Clarke, (2019) por lo que esta investigación se circunscribe a un enfoque interpretativo basado en una secuencia inductiva, mediada por el investigador.

3. RESULTADOS

La conjunción de todos los datos codificados recogidos a través de los diferentes instrumentos permite interrelacionar la información para la creación de principios de diseño. Estos principios deben definir estrategias útiles para docentes que pretendan integrar el uso de redes sociales y aprendizaje móvil en el área de EF en la etapa de Secundaria, con objeto de aumentar la actividad física del alumnado.

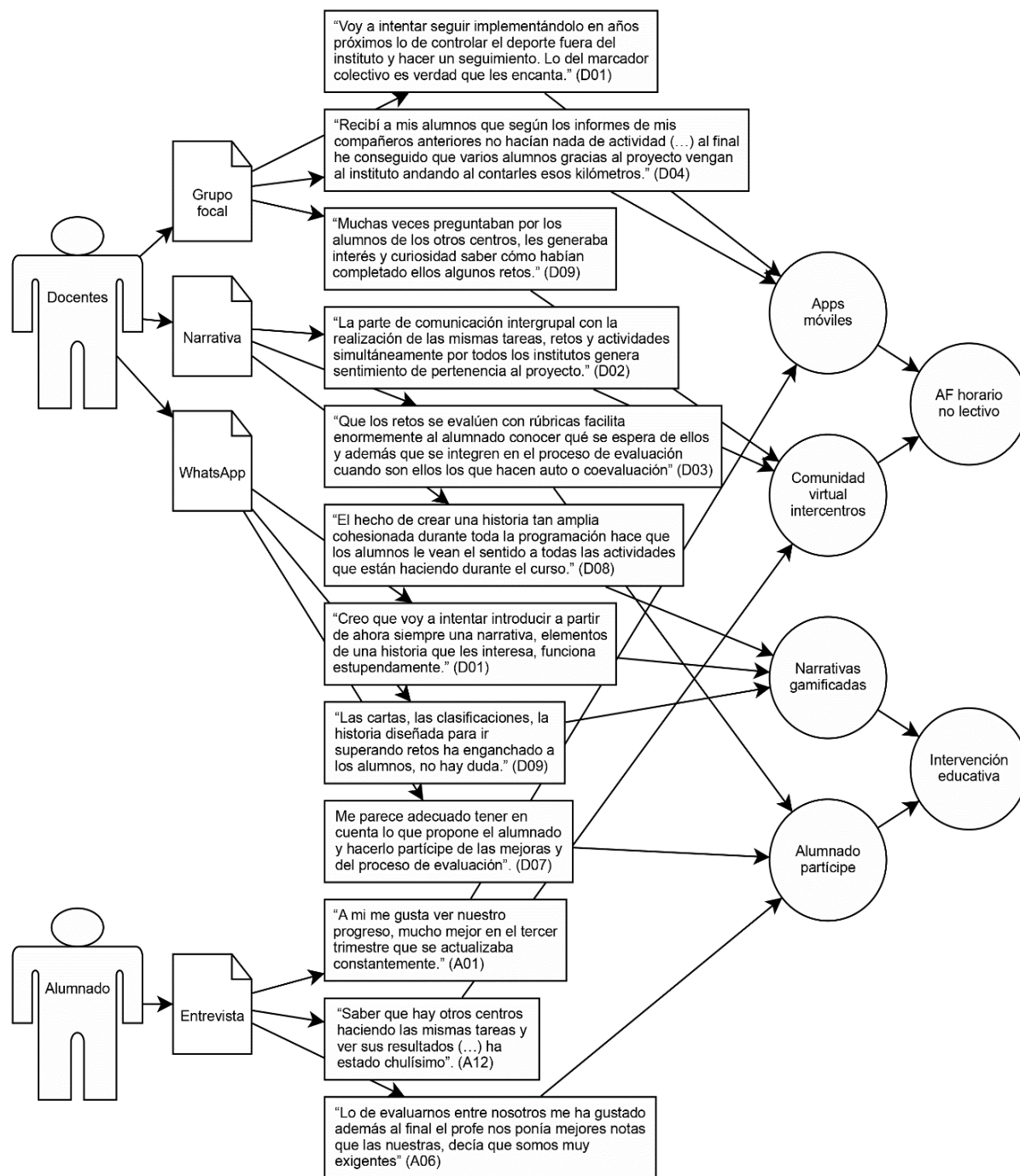
Analizando los principios de diseño en detalle, en la Figura 6 se puede observar gráficamente dicho proceso en el que, según la percepción de los docentes, se extrae como estrategia pedagógica-digital prioritaria la necesidad de utilizar aplicaciones móviles para monitorizar la realización de actividad física en el horario no lectivo del alumnado. Para ello, debe poseer la característica de realizar un seguimiento y conteo de la actividad física realizada, principalmente en forma de marcador colectivo, para colaborar entre compañeros de clase o entre diferentes grupos. En esta línea, se evidencia también la necesidad de, si existe la oportunidad, proyectar este tipo de intervenciones sobre la colaboración y participación de diferentes centros educativos para generar una comunidad de alumnos y docentes amplia alrededor del proyecto. El uso de redes sociales es necesario para la adecuada realización de esta labor.

En relación con el análisis de la intervención educativa se define como estrategia pedagógico-digital prioritaria la implementación de narrativas de gamificación. Esto se evidencia de nuevo en la Figura 6. Si se decide gamificar el proceso de enseñanza-aprendizaje, la gamificación del programa de intervención debe integrar un seguimiento del progreso del alumnado a través de insignias, rankings, niveles y recompensas a través de medios digitales que faciliten su uso. Debe ser construido sobre una narrativa flexible y adaptable a las circunstancias y preferencias del alumnado y ser una narrativa o historia lineal que cohesione las unidades formativas implementadas con un objetivo final.

Además, docentes y alumnos muestran satisfacción ante la participación de estos últimos en el proceso de evaluación como en el diseño de la narrativa. Esto pone de manifiesto que el alumnado debe integrarse en un proceso de evaluación formativa y compartida pudiendo implementar rúbricas de autoevaluación y coevaluación. Debe ser partícipe, además, del propio rediseño continuado de la intervención educativa mediante el diálogo y la negociación con el docente y sus compañeros.

Figura 6

Análisis de contenido en NVivo12. Categorías emergentes de “actividad física en horario no lectivo” e “intervención educativa”



Respecto al análisis realizado en torno a la categoría que emerge sobre el uso de recursos digitales se extrae como estrategia pedagógico-digital prioritaria la utilización de vídeos e infografías para la explicación de tareas o exposición de contenidos, además de para promover la curiosidad previa al inicio de unidades formativas. Ejemplo de ello son las aportaciones de los participantes, tanto docentes como alumnado, que se recogen en la Figura 7. Esta categoría se caracteriza por el debate que genera el uso del móvil y el consenso de los participantes para implementarlo como recurso educativo, formando en su uso crítico desde un enfoque educativo, aprovechando todas sus potencialidades con fines didácticos.

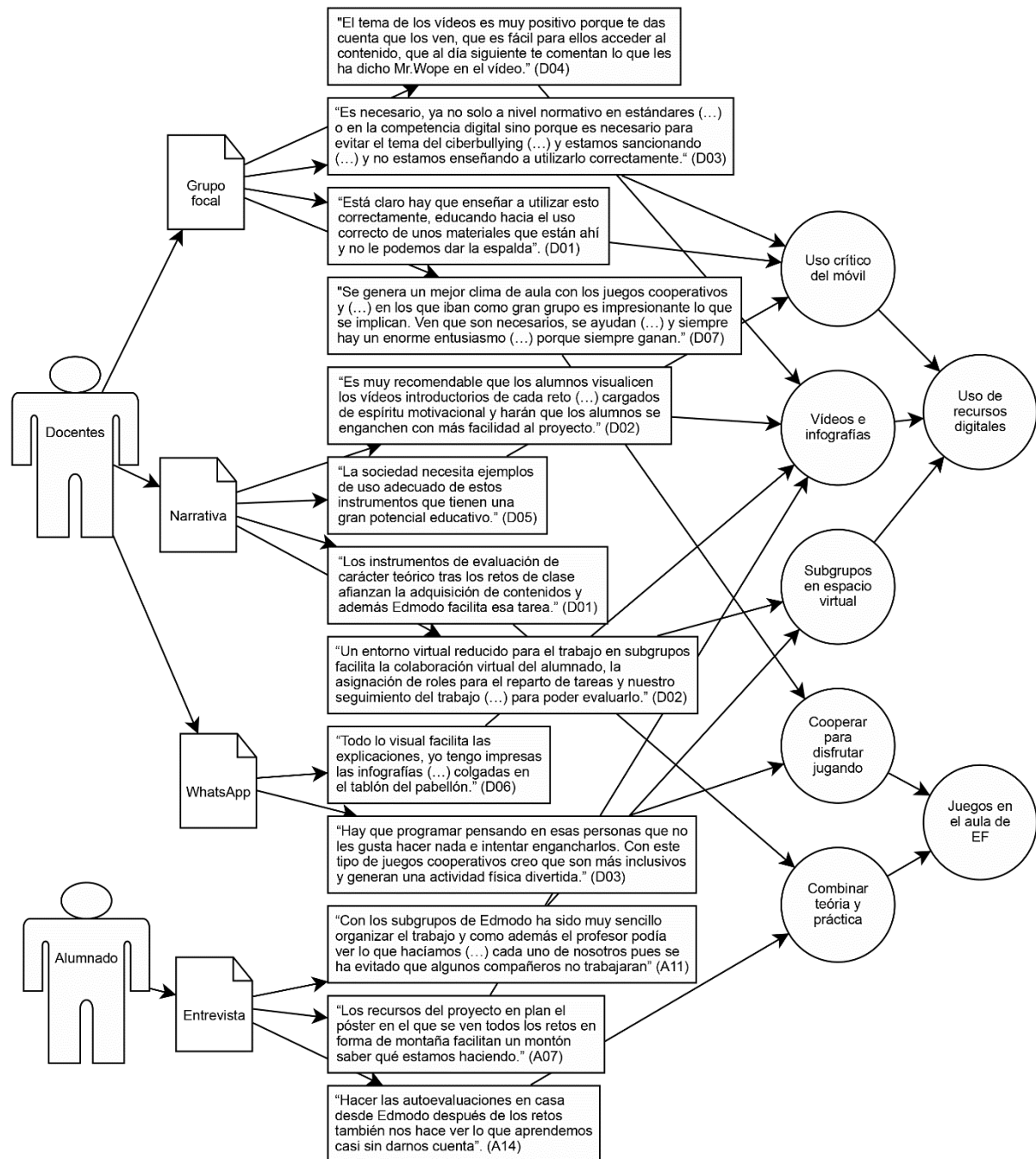
Por otra parte, tanto docentes como alumnado muestran su satisfacción al utilizar subgrupos de trabajo en espacios virtuales para facilitar el aprendizaje cooperativo como se ejemplifica en la Figura 7. Esto pone de manifiesto que la colaboración virtual se potencia a partir de subgrupos de trabajo con la creación de un espacio específico dentro del aula virtual. Espacios que simplifican la comunicación entre compañeros, para compartir archivos o fuentes de información y facilitando el seguimiento y evaluación del proceso de aprendizaje del alumno por parte del docente al quedar constancia en la red el trabajo realizado por cada uno de ellos. Además, existen multitud de potencialidades dentro de los subgrupos virtuales como que cada compañero puede tener un rol diferenciado y una tarea concreta.

En cuanto a la categoría relacionada con los juegos y deportes de equipo en el aula de EF, la perspectiva los participantes pone de manifiesto una estrategia clave: cooperar para disfrutar jugando. Así, se debe dar prioridad a los juegos basados en la cooperación en gran grupo ya que facilitan la inclusión activa de todos con un objetivo común. Además, el docente puede implementar estrategias competitivas proponiendo diferentes objetivos a superar -metas o competir con otro centro educativo con la cooperación de todo el grupo- sin necesidad de crear una competición presencial entre compañeros de clase.

En esta misma categoría, se hace imprescindible como estrategia pedagógico-digital la combinación de teoría y práctica mediante el uso de las TD. Así, tras el desarrollo de contenidos de carácter práctico se debe aprovechar el entorno virtual de aprendizaje para implementar trabajos o pruebas teóricas que sirvan para afianzar y relacionar el contenido teórico-práctico además de para evidenciar su desarrollo. Algunas afirmaciones de docentes y alumnos que desarrollan esta idea se recogen en la Figura 7.

Figura 7

Análisis de contenido en NVivo12. Categorías emergentes de “uso de recursos digitales” y “juegos en el aula de EF”

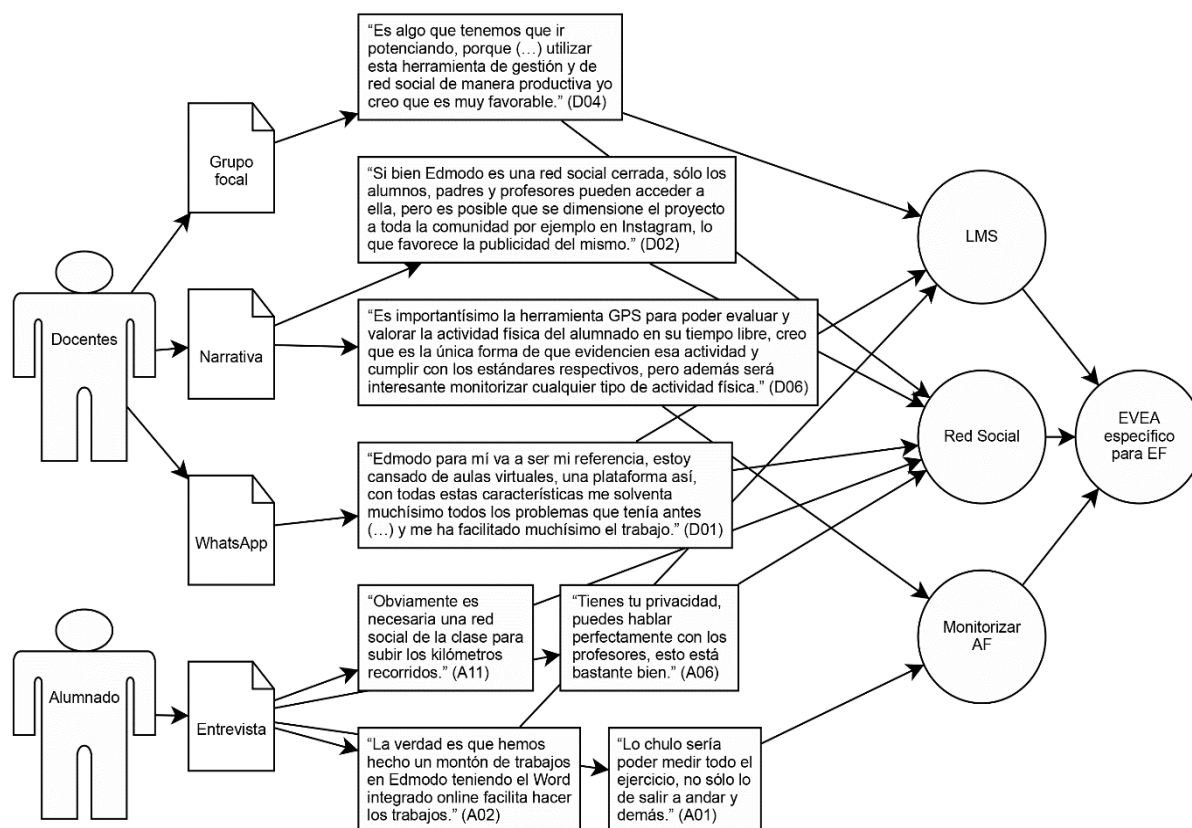


Por último, como estrategia pedagógico-digital resultante del análisis de contenido del espacio virtual utilizado (Figura 8), se observa que el entorno virtual de enseñanza-aprendizaje(EVEA) ideal para EF debe estar integrado por tres elementos clave:

- Un sistema de gestión del aprendizaje para administrar recursos, materiales, actividades de formación y hacer seguimiento del proceso de aprendizaje del alumnado (D01, D04, A02).
- Una red social que permita la colaboración virtual entre docentes, entre alumnos y entre ambos a partir de espacios de aula privados y en espacios de colaboración abiertos a toda la comunidad interesada (D01, D02, D04, A06, A11).
- Una herramienta de monitorización de actividad física del alumnado para su utilización en horario no lectivo (D06, A01).

Figura 8

Análisis de contenido en NVivo12. Categoría emergente de "EVEA específico para EF"



4. DISCUSIÓN Y CONCLUSIONES

La finalidad del presente estudio ha sido la creación de principios básicos para el diseño de estrategias pedagógico-digitales en el ámbito de la EF a partir de la colaboración con docentes y la percepción de los estudiantes. Los principios que emergen de la investigación pueden estructurarse en cinco grandes categorías, que sintetizamos a continuación:

- La propuesta al alumnado de realización de actividad física en horario no lectivo debe venir acompañada de un plan que integre la utilización de aplicaciones móviles y redes sociales y buscar la posibilidad de colaborar con otros centros educativos de manera virtual.

- Tal intervención educativa debe construirse sobre una narrativa gamificada que tenga en cuenta el contexto del alumnado, así como la participación de estos en el rediseño de dicha intervención.
- El uso de recursos digitales debe construirse sobre el análisis crítico de los dispositivos móviles por parte del alumnado, las posibilidades de diseño de contenidos por parte del docente desde vídeos a infografías y las ventajas de gestionar subgrupos de trabajo en entornos virtuales.
- Entre los tipos de juego en el aula de EF debería priorizarse la elección de juegos cooperativos, tanto en gran grupo como subgrupo, donde se compite y trabaja por alcanzar objetivos comunes. Además, el desarrollo práctico de los contenidos en esas sesiones de juego debería reforzarse de manera teórica posteriormente a la sesión, preferiblemente a través de autoevaluaciones en el espacio virtual del aula.
- El entorno virtual ideal para el proceso de enseñanza-aprendizaje en EF de Secundaria debería aunar las posibilidades incluidas en un sistema de gestión de contenidos, en una red social educativa y en una aplicación móvil de monitorización de la actividad física.

Es necesario indicar que estas estrategias pedagógico-digitales son transferibles a otros contextos educativos pero los efectos derivados de esta investigación no son generalizables. Hay que tener en cuenta que se ha construido sobre las propuestas de diseño participativo y creación colaborativa en TE basadas en la IBD en un proyecto en el que ha predominado la colaboración entre los docentes participantes, el cumplimiento de la normativa curricular actual y la inclusión de TD guiadas al desarrollo del objetivo principal de etapa de la EF, que no es otro que el movimiento y el aumento de la actividad física del alumnado.

La sucesión e interrelación de las estrategias pedagógico-digitales generadas se ponen de manifiesto en otras investigaciones o experiencias de diversas maneras. Se observa la importancia del uso de aplicaciones móviles para monitorizar actividad física y compartirla con los compañeros en un espacio virtual común en el que se producen procesos de gamificación con rankings y recompensas. Es el caso de Moreno et al. (2015) y Pulido et al. (2016) quienes se sirvieron de las aplicaciones Endomondo y Runtastic, respectivamente, para registrar la actividad física de estudiantes de Secundaria ordenándola en diferentes clasificaciones, consiguiendo mejorar la participación e implicación de los estudiantes. Tal uso de aplicaciones móviles para la mejora de la enseñanza en EF va en la línea de Díaz-Barahona (2020) quien concluyó que estas aplicaciones y el uso de características móviles como la geolocalización, vídeos o redes sociales enriquecen el proceso de aprendizaje.

Respecto a la importancia de las redes sociales para la implicación del alumnado en EF, Monguillot et al. (2015) determinaron que se convierten en una potente herramienta para fomentar hábitos de vida saludables en los adolescentes tras haber diseñado e implementado diferentes situaciones de aprendizaje mediadas por TD con alumnos de Secundaria. Goodyear et al. (2019) confirmaron que las redes sociales ofrecen un alto grado de identificación individual dentro de la comunidad, algo a lo que ya hacían referencia en investigaciones anteriores los propios Goodyear y Casey (2013). Tal como se evidencia en este estudio, una plataforma como Edmodo permite generar los habituales procesos de comunicación de las redes sociales, así como integrar las características de gestión del aprendizaje de un sistema de gestión de contenidos virtual y todo ello siendo un entorno útil y motivador para establecer

relaciones interpersonales del alumnado que comparte actividad física. Sobejano et al. (2016) destacaron la importancia de la relación entre actividad física y redes sociales tras una experiencia con alumnos de Secundaria en la que debían compartir con sus compañeros de clase la actividad física realizada en su tiempo libre mediante la plataforma de Edmodo.

El uso de estrategias gamificadoras con narrativa lineal y un objetivo final, utilizando estrategias cooperativas como el marcador colectivo, facilita el fomento de la actividad física del alumnado. Resultados similares reúne Victoria (2020) en una revisión bibliográfica que relaciona la gamificación y el uso de la tecnología en EF destacando elementos como sistema de niveles, puntos, clasificaciones o insignias y el uso de aplicaciones móviles por su efecto motivador en el alumnado hacia la práctica regular de actividad física y deportiva. Además, se destaca que estos procesos de gamificación deben ser flexibles para coordinar diferentes actividades de largo recorrido en el tiempo siendo similar al estudio realizado por Fernández-Río et al. (2020) quienes concluyeron que la gamificación implementada a largo plazo aumenta la motivación del alumnado más allá del efecto novedad inicial.

La intervención educativa debe incluir el uso de recursos digitales como los vídeos e infografías que facilitan el proceso de enseñanza-aprendizaje motivando al aprendizaje, ambientando la narrativa de gamificación y ganando en tiempo de compromiso motor del alumnado. Østerlie y Kjelaas (2019) reportaron resultados similares tras un estudio de aula invertida en EF de Secundaria otorgando al uso de los vídeos un elevado poder de motivación entre el alumnado. Papastergiou et al. (2021) demostraron que con el uso de tabletas y aplicaciones móviles especializadas se ganaba en tiempo de compromiso motor, permitiendo al docente liberarse del deber de demostrar repetidamente diferentes ejercicios y pudiendo proporcionar retroalimentación individualizada a sus estudiantes.

Otra de las estrategias priorizadas por la perspectiva de los participantes es la importancia del aprendizaje cooperativo, tanto en los espacios virtuales como en las sesiones de EF a través de la estructuración en subgrupos. Casey y Fernández-Río (2019) exponen la importancia de utilizar técnicas cooperativas para el aprendizaje en EF y para el desarrollo del ámbito afectivo. Rivera-Pérez et al. (2020) también corroboran la asociación positiva entre el aprendizaje cooperativo y el desarrollo del dominio afectivo de los estudiantes, por lo que promulgan la necesidad de incrementar la implementación de los principios de trabajo cooperativo en el área de EF.

De igual forma destaca el papel activo de los alumnos como participantes en el rediseño de los retos y como actores principales del proceso de evaluación. Así, la evaluación se convierte en una forma de diálogo y de toma de decisiones conjuntas entre alumnos y docente. Han primado los procesos de autoevaluación y coevaluación cuyas mejoras van en la línea de López-Pastor, et al. (2005) respecto a poder formar al alumnado en los contenidos trabajados desde la retroalimentación, siendo coherente con las metodologías activas y de fácil asimilación a cualquier contenido de EF por su viabilidad y adaptabilidad a diferentes unidades formativas. Supone también la mejora en las variables de aprendizaje, implicación y rendimiento académico del alumnado.

Los resultados extraídos no son estancos, sino que fluyen entre ellos y están intrínsecamente interrelacionados. Con objeto de sintetizar las estrategias pedagógico-digitales resultantes y

para dar respuesta a qué estrategias deberían priorizarse en una intervención en EF de Secundaria basado en el uso de aplicaciones móviles y redes sociales se concluye que:

Se debería implementar un entorno virtual de Enseñanza-Aprendizaje específico para EF que integre las características de tres herramientas principales: las posibilidades de un sistema virtual de gestión de contenidos; las características de una red social educativa con espacios abiertos y privados; y, las potencialidades de monitorización de la actividad física. Asimismo, integrar una narrativa de gamificación se convierte en una pieza clave, incluyendo los elementos de cualquier proceso gamificado, teniendo en cuenta la participación de los estudiantes para el diseño de esa narrativa y para los procesos de evaluación formativa y compartida. No debemos perder de vista que el docente debería servirse del diseño de vídeos e infografías para explicar el contenido o para establecer la narrativa de la gamificación.

Otras cuestiones importantes son tanto que los contenidos prácticos deben reforzarse a través de actividades teóricas mediadas por el uso de las TD, como que los juegos cooperativos deben seleccionarse como prioritarios, ya que facilitan la integración de todos los alumnos en el grupo y generan dinámicas positivas. Por último, la estrategia didáctica de formar subgrupos en el entorno virtual debe tenerse en cuenta para facilitar la colaboración de los estudiantes. Asimismo, el programa de intervención diseñado mejora si existe colaboración entre diferentes centros educativo.

Una limitación del presente estudio es la contextualización específica de la que surgen los principios de diseño, por lo que se hace necesario plantear como línea de investigación futura replicar en otros contextos el programa de intervención codiseñado introduciendo los principios de diseño generados. En esta línea de acciones futuras, sería razonable sustituir el uso de los teléfonos móviles por pulseras de salud y actividad para la monitorización de actividad física. Se limitaría el uso del smartphone, siempre polémico, y se ganaría en capacidad de monitorización de multitud de variables gracias a los sensores de estos dispositivos. Se podrían trabajar contenidos como las zonas de frecuencia cardiaca saludable, calorías, índice de masa corporal o estado del sueño. Contenidos teóricos, a veces abstractos, que el alumnado pasaría a conocer y medir en su propio cuerpo y actividad física.

5. REFERENCIAS

- Aguilar, S. y Barroso, J. (2015). La triangulación de datos como estrategia en investigación educativa. *Pixel-Bit. Revista de medios y educación*, (47), 73-88. <https://doi.org/10.12795/pixelbit.2015.i47.05>
- Bailey, R. (2006). Physical education and sport in schools: A review of benefits and outcomes. *Journal of school health*, 76(8), 397-401. <https://doi.org/10.1111/j.1746-1561.2006.00132.x>
- Beni, S., Fletcher, T. y Chróinín, D. N. (2019). Using features of meaningful experiences to guide primary physical education practice. *European Physical Education Review*, 25(3), 599-615. <https://doi.org/10.1177%2F1356336X18755050>

- Braun, V. y 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>
- Cabero, J. y Llorente, M.C. (2013). La aplicación del juicio de experto como técnica de evaluación de las Tecnologías de la Información y la Comunicación (TIC), *Revista de Tecnología de Información y Comunicación en Educación*, 7(2), 11-22. <https://revistaeduweb.org/index.php/eduweb/article/view/206>
- Cabero, J. (2016). ¿Qué debemos aprender de las pasadas investigaciones en Tecnología Educativa? *RIITE Revista Interuniversitaria de Investigación en Tecnología Educativa*, 0, 23-33. <https://doi.org/10.6018/riite/2016/256741>
- Cañizares, J.M. y Carbonero, C. (2018). *Las TIC en la escuela actual: Nuevas metodologías didácticas en Educación Física*. Wanceulen.
- Carpenter, J. P. y Linton, J. N. (2018). Educators' perspectives on the impact of Edcamp unconference professional learning. *Teaching and Teacher Education*, 73, 56-69. <https://doi.org/10.1016/j.tate.2018.03.014>
- Casey, A. y Fernández-Río, J. (2019). Cooperative learning and the affective domain. *Journal of Physical Education, Recreation y Dance*, 90(3), 12-17. <https://doi.org/10.1080/07303084.2019.1559671>
- Casey, A., Goodyear, V.A. y Armour, K. M. (2017). Rethinking the relationship between pedagogy, technology and learning in health and physical education. *Sport, education and society*, 22(2), 288-304. <https://doi.org/10.1080/13573322.2016.1226792>
- Castañeda, L., Salinas, J. y Adell, J. (2020). Hacia una visión contemporánea de la Tecnología Educativa. *Digital Education Review*, 37, 240-268. <https://raco.cat/index.php/DER/article/view/375035>
- Contreras, O. y Gutiérrez, D. (2017). *El Aprendizaje Basado en Proyectos en Educación Física*. INDE.
- de-Benito, B. y Salinas, J. M. (2016). La Investigación Basada en Diseño en Tecnología Educativa. *RIITE Revista Interuniversitaria de Investigación en Tecnología Educativa*, 0, 44-59. <https://doi.org/10.6018/riite2016/260631>
- Díaz-Barahona, J. (2020). Retos y oportunidades de la tecnología móvil en la educación física. *Retos*, (37), 763-773. <https://doi.org/10.47197/retos.v37i37.68851>
- Dimo, S., Bandojo, H. I., Estilo, K., Abalayan, J., Achas, M. y Lobo, J. (2024). Utilizing TAM Framework in Students' Behavioral Intention and Use towards Google Classroom for Physical Education. *Journal of Educators Online*, 21(2), 1-13. <https://doi.org/10.9743/JEO.2024.21.2.11>
- Fernández-Río, J., de las Heras, E., González, T., Trillo, V. y Palomares, J. (2020). Gamification and physical education. Viability and preliminary views from students and teachers. *Physical Education and Sport Pedagogy*, 25(5), 509-524. <https://doi.org/10.1080/17408989.2020.1743253>

- Goodyear, V. A., Parker, M. y Casey, A. (2019). Social media and teacher professional learning communities. *Physical Education and Sport Pedagogy*, 24(5), 421-433. <https://doi.org/10.1080/17408989.2019.1617263>
- Gros, B. y Durall, E. (2020). Retos y oportunidades del diseño participativo en tecnología educativa. *Revista electrónica de Tecnología Educativa*, 74, 12-24. <http://dx.doi.org/10.21556/edutec.2020.74.1761>
- Hew, K.F., Min, L., Tang, Y., Jia, C. y Lo, C.K. (2019). Where is the “theory” within the field of educational technology research? *British Journal of Educational Technology*, 50(3), 956-971. <http://dx.doi.org/10.1111/bjet.12770>
- Jailobaev, T., Jailobaeva, K., Baialieva, M., Baialieva, G. y Asilbekova, G. (2021). WhatsApp Groups in Social Research: New Opportunities for Fieldwork Communication and Management. *Bulletin de Methodologie Sociologique*, 49, 60-82. <https://doi.org/10.1177/0759106320978337>
- Lobo, J. (2023). Students’ acceptance of google classroom as an effective pedagogical tool for Physical Education. *IJERI: International Journal of Educational Research and Innovation*, (20), 1–15. <https://doi.org/10.46661/ijeri.7535>
- López-Pastor, V. M., González, M. y Barba, J.J. (2005). La participación del alumnado en la evaluación: la autoevaluación, la coevaluación y la evaluación compartida. *Tándem Didáctica de la Educación Física*, (17), 21-37. <http://hdl.handle.net/11162/21846>
- Martínez, F. (2016). Sentado en el andén. *RIITE Revista Interuniversitaria de Investigación en Tecnología Educativa*, 0, 17-22. <https://doi.org/10.6018/riite/2016/258131>
- McKenney, S. y Reeves, T.C. (2012). *Conducting Educational Design Research*. Routledge. <https://doi.org/10.4324/9780203818183>
- Miles, M. y Huberman, A.M. (1984). *Qualitative data analysis. A source book of new methods*. Sage.
- Mishra, P. y Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, 108(6), 1017-1054. <https://doi.org/10.1111/j.1467-9620.2006.00684.x>
- Monguillot, M., González, C. y Guitert, M. (2015). Diseño de situaciones de aprendizaje mediadas por TIC en Educación Física. *Revista Iberoamericana de Educación*, 68(2), 63-82. <https://doi.org/10.35362/rie682185>
- Montiel-Ruiz, F. J. (2020). *Aplicaciones móviles y redes sociales para la enseñanza de Educación Física en Secundaria*. Tesis Doctoral. Universidad de Murcia.
- Montiel-Ruiz, F. J. (2022). Aprendizaje cooperativo y evaluación formativa de la expresión corporal en entornos virtuales. *EmásF: revista digital de educación física*, (75), 179-192.
- Moreno, H., Cuevas, R., Pulido, J.J. y García, T. (2015). Influencia de una aplicación de telefonía móvil de entrenamiento sobre las necesidades psicológicas y la motivación autodeterminada en escolares. *Cuadernos de Psicología del Deporte*, 15(2), 71-78. <https://revistas.um.es/cpd/article/view/233831>

- Østerlie, O. y Kjelaas, I. (2019). The perception of adolescents' encounter with a flipped learning intervention in Norwegian physical education. *Frontiers in Education*, 4,(114). <https://doi.org/10.3389/educ.2019.00114>
- Papastergiou, M., Natsis, P., Vernadakis, N. y Antoniou, P. (2021). Introducing tablets and a mobile fitness application into primary school physical education. *Education and Information Technologies*, 26(1), 799-816. <https://doi.org/10.1007/s10639-020-10289-y>
- Prendes, M. P., Solano, I. M., Serrano, J. L., González, V. y Román, M. M. (2018). Entornos Personales de Aprendizaje para la comprensión y desarrollo de la Competencia Digital: análisis de los estudiantes universitarios en España. *Educatio Siglo XXI*, 36(2), 115-134. <http://dx.doi.org/10.6018/j/333081>
- Puentedura, R.R. (2020, January 25). SAMR - A research perspective. http://hippasus.com/rrpweblog/archives/2020/01/SAMR_AResearchPerspective.pdf
- Pulido, J. J., Sánchez, D., Sánchez, P. A., González, I. y García, T. (2016). Proyecto MOVIL-ÍZATE: Fomento de la actividad física en escolares mediante las Apps móviles. *Retos. Nuevas tendencias en Educación Física, Deporte y Recreación*, (30), 3-8. <https://recyt.fecyt.es/index.php/retos/article/view/34258>
- Quennerstedt, M. (2019). Healthy physical education-on the possibility of learning health. *Physical Education and Sport Pedagogy*, 24(1), 1-15. <https://doi.org/10.1080/17408989.2018.1539705>
- Quennerstedt, M., Backman, E. y Mikael, J. (2024). Returning to the river: the salutogenic model as a theory to explore the relation between outdoor activities and health. *Journal of Adventure Education and Outdoor Learning*, 1-16. <https://doi.org/10.1080/14729679.2024.2342305>
- Reeves, T. C. (2006). Design research from the technology perspective. En J. V. Akker, K. Gravemeijer, S. McKenney y N. Nieveen (Eds.), *Educational design research* (pp. 86-109). Routledge. <https://doi.org/10.4324/9780203088364>
- Rivera-Pérez, S., León-del-Barco, B., Fernández-Río, J., González-Bernal, J. J. y Iglesias, D. (2020). Linking Cooperative Learning and Emotional Intelligence in Physical Education: Transition across School Stages. *International Journal of Environmental Research and Public Health*, 17(14), 5090. <https://doi.org/10.3390/ijerph17145090>
- Sánchez-Vera, M.M. (2023). Los desafíos de la Tecnología Educativa. *RiiTE. Revista interuniversitaria de investigación en Tecnología Educativa*, 14, 1-5. <https://doi.org/10.6018/rriite.572131>
- Super, S., Verkooijen, K. y Koelen, M. (2021). A salutogenic perspective on sport-for-development research. *Social Science y Medicine*, 268, 113376. <https://doi.org/10.1016/j.socscimed.2020.113376>
- Sancho, J.M., Alonso, C. y Sánchez, J. A. (2018). Miradas retro-prospectivas sobre las Tecnologías Educativas. *Educatio Siglo XXI*, 36 (2), 209-228. <http://dx.doi.org/10.6018/j/333051>
- Sargent, J. y Casey, A. (2020). Flipped learning, pedagogy and digital technology: Establishing consistent practice to optimise lesson time. *European Physical Education Review*, 26(1), 70-84. <https://doi.org/10.1177%2F1356336X19826603>

- Sobejano, M., Hortigüela, D., Pérez-Pueyo, A. y Hernando, A. (2016). La implicación del alumnado de educación física a través del uso de las redes sociales en el aula. *Revista Arbitrada del Centro de Investigación y Estudios Gerenciales* (24), 73-81. <https://bit.ly/3IGNQ9s>
- Trigueros, C., Rivera, E. y Rivera, I. (2018). *Técnicas conversacionales y narrativas. Investigación cualitativa con software NVivo*. Universidad de Granada.
- Valverde-Berrocoso, J. (2016). La investigación en Tecnología Educativa y las nuevas ecologías del aprendizaje: Design-Based Research (DBR) como enfoque metodológico. *RIITE Revista Interuniversitaria de Investigación en Tecnología Educativa*, 0, 60-73. <https://doi.org/10.6018/riite/2016/257931>
- van den Berg, C., Bozalek, V. (2024). Design principles to develop digital innovation skills: a design-based research approach. *Education Tech Research Dev.* 72, 947–972 <https://doi.org/10.1007/s11423-023-10308-y>
- Victoria, C. (2020). Herramientas TIC para la gamificación en Educación Física. *EduTec. Revista Electrónica de Tecnología Educativa*, (71), 67-83. <https://doi.org/10.21556/edutec.2020.71.1453>
- Wang, F. y Hannafin, M.J. (2005) Design-based research and technology-enhanced learning environments. *Educational Technology Research y Development*, 53(5), 5-23. <http://dx.doi.org/10.1007/BF02504682>
- Zimmerman, B. J. y Schunk, D. H. (2011). *Handbook of Self-Regulation of Learning and Performance*. Routledge.

Para citar este artículo:

Montiel-Ruiz, F. J., Calderón, A., Sánchez-Vera, M. M., Solano-Fernández, I. M. (2024). Redes sociales y aprendizaje móvil en Educación Física: Creación colaborativa de principios de diseño. *EduTec, Revista Electrónica de Tecnología Educativa*, 89, 190-209. <https://doi.org/10.21556/edutec.2024.89.3191>