

Rosana Satorre Cuerda (Ed.)

Docencia universitaria en transformación: escenarios emergentes

Horizontes Universidad

Octaedro

Rosana Satorre Cuerda (Ed.)

Docencia universitaria en transformación: escenarios emergentes

Octaedro
Editorial

Docencia universitaria en transformación: escenarios emergentes

Edición: Rosana Satorre Cuerda

Revisión y maquetación: ICE de la Universidad de Alicante

Primera edición: noviembre de 2026

© De la edición: Rosana Satorre Cuerda

© Del texto: Las autoras y autores

© De esta edición:

Ediciones OCTAEDRO, S.L.
C/ Bailén, 5 – 08010 Barcelona
Tel.: 93 246 40 02 – Fax: 93 231 18 68
www.octaedro.com – octaedro@octaedro.com

ISBN: 978-84-1079-508-2

Producción: Ediciones Octaedro

La revisión de los trabajos se ha realizado de forma rigurosa, siguiendo el protocolo de revisión por pares.

Cualquier forma de reproducción, distribución, comunicación pública o transformación de esta obra solo puede ser realizada con la autorización de sus titulares, salvo excepción prevista por la ley. Diríjase a CEDRO (Centro Español de Derechos Reprográficos, www.cedro.org) si necesita fotocopiar o escanear algún fragmento de esta obra.

NOTA EDITORIAL: Las opiniones y contenidos de los textos publicados en esta obra son de responsabilidad exclusiva de los autores.

COMITÉ TÉCNICO:
Neus Pellin Buades, Universidad de Alicante

COMITÉ CIENTÍFICO INTERNACIONAL:
Aires, Luísa. Universidade Aberta de Portugal
Antón Ruiz, Josefa A. Universidad de Alicante
Aparicio Flores, Pilar. Universidad de Alicante
Aznar Crespo, Pablo. Universidad de Alicante
Baez Carratalá, Alejandro. Universidad de Alicante
Balteiro Fernández, María Isabel. Universidad de Alicante
Berná Martínez, José Vicente. Universidad de Alicante
Bressole Bolaños, Doris Verónica.Universidad Central de Ecuador
Buquet Corleto, Ana Gabriela. Universidad Nacional Autónoma de México (UNAM)
Carrasco Rodríguez, Antonio. Universidad de Alicante
Castro López, María Ángeles. Universidad de Alicante
De Juana Espinosa, Susana. Universidad de Alicante
Escabias Lloret, Pilar. Universidad de Alicante
Fernández Pascual, María Dolores. Universidad de Alicante
Formigós Bolea, Juan Antonio. Universidad de Alicante
García Jaén, Miguel. Universidad de Alicante
Gonzálvez Maciá, Carolina. Universidad de Alicante
Hernández Amorós, María José. Universidad de Alicante
Ibáñez Hernández, Ana Isabel. Universidad de Alicante
Jiménez Olmedo, José Manuel. Universidad de Alicante
Marcillas Piquer, Isabel. Universidad de Alicante
Meza Bolaños, Doris Verónica. Universidad Central de Ecuador
Molina Jordá, José Miguel. Universidad de Alicante
Rovira-Collado, José. Universidad de Alicante
Saiz Noeda, Maximiliano. Universidad de Alicante
Sanmartín López, Ricardo. Universidad de Alicante
Santos Ruiz, Ana María. Universidad de Alicante
Sentana Gadea, Irene. Universidad de Alicante
Sepulcre Martínez, Juan Matías. Universidad de Alicante
Villalba Clemente, Francisco Gabriel. Universidad de Alicante

Índice

Presentación

1. *Innovación docente en Fisiología Animal: adaptación de una práctica de laboratorio al aprendizaje autónomo del estudiante aplicando el principio de las tres R*
   Armando Alberola Die, Isabel Ivorra Pastor, Andrés Morales Calderón 3
2. *Diagnosticar para innovar: perfil basal del engagement, pensamiento crítico y cohesión grupal en una muestra de estudiantes universitarios*
   J. A. Antón-Ruiz, V. Clement Carbonell; L. J. Viñuelas Barberá, N. Albaladejo Blázquez, N. Ruiz Robledillo, I. Portilla Tamarit 14
3. *Aprendizaje adaptativo con inteligencia artificial para reforzar la enseñanza de Anatomía Humana en Ciencias de la Salud*
   Eva Ausó, Gema Esquiva Sobrino, Ana Cifuentes, Damián R. Martínez St. John, J. Víctor García Velasco, Javier Sanz Reig, María de las Nieves Pérez Climent, Erika Meléndez-Oliva* 23
4. *De la rigurosidad escrita a la síntesis visual: el papel del informe y el póster en ingeniería ante la IA generativa*
   Fidel Aznar Gregori, Ramón Rizo Aldeguer, Pilar Arques Corrales, Mar Pujol López, Patricia Compañ Rosique, Javier Botana Gómez, Miguel Angel Lozano Ortega, Francisco José Mora Lizán, Juan Antonio Puchol García, Mª José Pujol López 31
5. *Del lenguaje especializado al lenguaje claro: innovación docente y alfabetización en salud mediante el etiquetado alimentario*
   Isabel Balteiro 40
6. *From detection to traceability: A segmented dossier architecture for assessment under generative AI in a History of America course*
   Antonio Carrasco-Rodríguez 54
7. *El Seminario de lecturas como escenario de formación teórico-práctica y reflexividad crítica en Trabajo Social*
   Mercedes Cuenca Silvestre, Javier Ferrer Aracil, Belén Parra Ramajo 66
8. *Visible Only to Those Who Know Where to Look: Multimodal Literacy, Visual Grammar, and Critical Reading in the L2 Classroom*
   Pilar Escabias Lloret 77
9. *Aprender italiano con la Inteligencia Artificial Generativa: autoaprendizaje y autocorrección*
   Maria Angelica Giordano Paredes, María Soledad Villarrubia Zúñiga 91
10. *Innovación docente a través de la historia familiar como parte del aprendizaje significativo en el desarrollo personal y la elección profesional*
    Asunción Lillo-Beneyto, José Emiliano Ramírez-García, ... 102

11. *Visualizar la comprensión traductora: externalización gráfica del razonamiento en la formación inicial de traductores*
Victoria I. Miralles Wieczorek 113

12. *La inteligencia artificial generativa en la enseñanza del Derecho: una estrategia para desarrollar las habilidades de pensamiento crítico que sustentan el razonamiento jurídico*
Lucía Molina Martínez, María Cristina Berenguer Albaladejo, Andrea García Jurado, Amanda Judith Gil Marín, María Remedios Guilabert Vidal, Lucía López Gutiérrez, Isabela Erika Ungureanu Orosz 123

13. *Aplicación de metodologías activas en la universidad y su relevancia en el contexto educativo superior*
Ignasi Navarro Soria, Inés Rubio Sánchez Molina, Alba Gutiérrez Martínez, Manuel Torrecillas Martínez 134

14. *Programación colaborativa en ingeniería apoyada por inteligencia artificial: innovación educativa para un aprendizaje significativo*
J. Pérez Beltrán, R.P. Ñeco García, N.M. García Aracil, J.M. Catalán Orts, A. Muñoz Lozano 144

15. *AI-Assisted Active Learning and Scientific Competence in Sport Science Education*
Basilio Pueo, Alfonso Penichet-Tomas, Lamberto Villalon-Gasch, Jose Manuel Jimenez-Olmedo. 155

16. *Evaluación continua con validación formativa en Python: sostenibilidad, escalabilidad e integridad académica en la enseñanza del cálculo numérico en el Grado en Matemáticas*
J. A. Vargas Alemañy, L. Moreno Martínez, J.M. Sayol España, I. Vigo Aguiar, D. García García 163

Docencia universitaria en transformación: escenarios emergentes

Presentación

El libro que aquí presentamos constituye una nueva aportación al esfuerzo colectivo por comprender, renovar y fortalecer la enseñanza universitaria ante los profundos cambios que atraviesan la educación superior. Tras más de dos décadas de investigación, innovación y práctica docente compartida en el marco de las Jornadas de Redes de Investigación en Docencia Universitaria (REDES) y del Workshop Internacional de Innovación en Enseñanza Superior y TIC (INNOVAESTIC) de la Universidad de Alicante, esta obra da continuidad a una trayectoria que refleja el compromiso de nuestra comunidad universitaria con la mejora de la docencia.

Bajo el título *Docencia universitaria en transformación: escenarios emergentes*, el monográfico ofrece una mirada amplia y plural sobre los procesos que están configurando la enseñanza universitaria contemporánea. La transformación de las prácticas de aula, la incorporación de la inteligencia artificial y las tecnologías digitales, y la aparición de nuevos escenarios académicos plantean oportunidades y desafíos que requieren respuestas fundamentadas y una reflexión permanente sobre cómo enseñamos y cómo aprenden nuestros estudiantes.

La obra aborda también dimensiones esenciales de esta transformación, como la evaluación y el desarrollo de competencias, la inclusión, el bienestar y la responsabilidad social universitaria. Asimismo, presta atención a la formación docente y a la evolución de la identidad profesional, así como a la comunicación, la cultura y la transferencia del conocimiento. Finalmente, la interdisciplinariedad y los retos emergentes del conocimiento permiten ampliar la mirada hacia una universidad cada vez más conectada con las transformaciones sociales, profesionales y culturales de su entorno.

Los trabajos recogidos en esta obra han sido seleccionados de entre 41 propuestas presentadas, siguiendo un riguroso proceso de revisión por pares. Cada contribución es fruto de la dedicación de docentes e investigadores que, desde sus aulas experimentan y reflexionan sobre cómo mejorar la calidad de la enseñanza en la educación superior.

Este volumen invita a docentes, investigadores y estudiantes a acercarse a estos escenarios desde una perspectiva crítica, abierta y constructiva, entendiendo la transformación de la docencia universitaria como un proceso continuo de adaptación, innovación y construcción colectiva orientado a una universidad más inclusiva, responsable y conectada con los desafíos del presente y del futuro.

Rafael Molina Carmona
Vicerrector de Transformación Digital

Rosana Satorre Cuerda
Directora del Instituto de Ciencias de la Educación

Universidad de Alicante

1. Innovación docente en Fisiología Animal: adaptación de una práctica de laboratorio al aprendizaje autónomo del estudiante aplicando el principio de las tres R[1]

Armando Alberola Die, Isabel Ivorra Pastor, Andrés Morales Calderón

Área de Fisiología, Departamento de Fisiología, Genética y Microbiología
Universidad de Alicante

RESUMEN

El presente trabajo se presenta como una innovación educativa en la asignatura Fisiología Animal I del Grado en Biología de la Universidad de Alicante, motivada por las restricciones legales, materiales y económicas que impiden el uso tradicional de vertebrados en la docencia. El objetivo principal fue rediseñar la práctica titulada "Registro del potencial de acción compuesto" bajo el principio de las tres R (fundamento consistente en el empleo de animales de experimentación de la manera más responsable y humana posible, que promueve el reemplazo de vertebrados, su reducción y el refinamiento de los procedimientos), promoviendo además el aprendizaje activo y autónomo de los estudiantes. Para ello, en sesiones prácticas con un total de 172 estudiantes, durante los cursos 24-25 y 25-26, se sustituyó el modelo clásico de demostración en vertebrados por una práctica equivalente en lombrices de tierra y se emplearon bioamplificadores portátiles de bajo coste. Como recurso complementario, se implementó una simulación virtual interactiva en la plataforma UACloud. La evaluación mediante cuestionarios completados por el alumnado arrojó resultados altamente positivos en ambos periodos; el estudiantado destacó que el uso de los bioamplificadores y del material virtual facilitó notablemente la comprensión técnica y conceptual, fomentando significativamente su autonomía. Asimismo, los estudiantes veteranos confirmaron su preferencia por esta metodología activa frente a las demostraciones pasivas tradicionales. En conclusión, el proyecto demuestra que la combinación de modelos invertebrados, tecnología de código abierto y herramientas de simulación constituye una alternativa pedagógica viable, ética y sostenible que puede optimizar las competencias prácticas universitarias.

PALABRAS CLAVE: Aprendizaje autónomo, Innovación educativa, Fisiología Animal, Principio de las tres R, Electrofisiología

1. INTRODUCCIÓN

La Fisiología es la ciencia que estudia el funcionamiento de los organismos vivos y los mecanismos mediante los cuales células, tejidos, órganos y sistemas mantienen la vida y el equilibrio interno del cuerpo, conocido como homeostasis. Asimismo, tiene como objetivo comprender cómo los diferentes procesos biológicos se integran y regulan para permitir la adaptación del organismo frente a cambios internos y/o externos (Billman, 2020; Kuang, 2023). La enseñanza actual de la Fisiología promueve tanto la adquisición simultánea de conocimientos conceptuales como de habilidades prácticas mediante actividades de laboratorio con animales, modelos funcionales y metodologías activas que permitan al estudiante analizar e interpretar fenómenos fisiológicos en contextos reales (Balaguera

1 El presente trabajo ha contado con una ayuda del Programa de Redes de investigación en docencia universitaria del Instituto de Ciencias de la Educación de la Universidad de Alicante (convocatoria 2024). Ref.:6116

et al., 2023; Sefton, 2001). En el Espacio Europeo de Educación Superior (EEES), este enfoque se enmarca dentro de lo que se denomina formalmente "Enseñanza Centrada en el Estudiante" (Student-Centred Learning o SCL), donde el proceso de enseñanza-aprendizaje radica en formar al estudiante de forma autónoma, dotándolo de herramientas adecuadas para el estudio, con materiales didácticos que funcionen como recursos capaces de generar conocimientos de alto nivel (European Students' Union [ESU], 2015). Así, el objetivo principal de este proceso de aprendizaje, debe consistir en el desarrollo de una serie de competencias y habilidades, tanto académicas como interpersonales, definidas en correspondencia con la demanda social y los perfiles académicos y profesionales de los futuros egresados.

No obstante, este enfoque, centrado en una mayor autonomía del estudiante, supone un reto a la hora de implementar las prácticas de laboratorio de asignaturas como Fisiología Animal I (FA-I), impartida en el cuarto semestre del Grado en Biología de la Universidad de Alicante (UA). Esta dificultad radica, en primer lugar, en que el marco legal contemporáneo, regulado por la Directiva de la Unión Europea 2010/63/UE, el Real Decreto 53/2013 y las normativas derivadas de ellos, establece un límite estricto para las metodologías docentes tradicionales en Fisiología, defendidas históricamente por autores como Medler (2018), en las que se empleaban animales de laboratorio como ratas, ratones o ranas. La normativa actual estipula de forma taxativa que cualquier procedimiento con animales vertebrados o cefalópodos queda restringido a personal con acreditación oficial específica. Dado que los estudiantes de Grado en Biología carecen dicha capacitación regulatoria, las sesiones prácticas *in vivo* con estos modelos se vuelven inviables en el aula de Grado, obligando al profesorado a impartirlas a modo de demostración pasiva y mermando la adquisición de competencias activas por parte del estudiante. A esta restricción legal se suma que las preparaciones experimentales "clásicas", con vertebrados, conllevan una alta dificultad técnica de ejecución y costes prohibitivos en tiempo de preparación, mantenimiento y consumibles (Doke & Dhawale, 2015).

Por consiguiente, ante la confluencia de estas barreras económicas y la imposibilidad legal de manipulación por parte del alumnado, se plantean dos grandes vías alternativas que garantizan un aprendizaje práctico real, autónomo y eficiente. Por un lado, la implementación de simuladores interactivos de libre acceso y modelos *in silico*; por otro lado, el uso pedagógico de modelos adecuados en invertebrados no protegidos (como insectos, anélidos o moluscos no cefalópodos), los cuales eliminan los conflictos éticos y legales de la experimentación con vertebrados, reducen drásticamente los costes logísticos y devuelven al estudiante el protagonismo de la manipulación directa en el laboratorio (Derakhchan et al., 2026). No obstante, el empleo de invertebrados también debe llevarse a cabo reduciendo al mínimo el número de animales, adoptando protocolos de anestesia estandarizados y estableciendo criterios de sacrificio rápido y humanitario (Liang et al, 2026).

En este escenario, el objetivo principal del presente trabajo se centró en promover el aprendizaje autónomo de los estudiantes rediseñando e implementando una práctica de la asignatura FA-I, el registro del potencial de acción compuesto en un nervio, tradicionalmente impartida como una demostración con vertebrados (ranas) por parte del docente, mediante la aplicación rigurosa del "Principio de las tres R" (reemplazo, reducción y refinamiento).

Este postulado fue desarrollado originalmente por William Russell y Rex Burch en 1959, en su libro «The Principles of Humane Experimental Technique» (Russell y Burch, 1959), donde

propusieron que, en el caso de que se utilicen animales en procedimientos/experimentos, se haga todo lo posible por sustituirlos por métodos alternativos (como pueden ser animales invertebrados, técnicas y/o muestras *in vitro* o simulaciones *in silico*), reducir al mínimo el número de animales utilizados y perfeccionar los experimentos con animales para que causen el mínimo dolor y sufrimiento (Derakhchan et al., 2026).

Para alcanzar el propósito principal, se plantearon dos sub-objetivos. En primer lugar, fomentar el aprendizaje activo y autónomo del alumnado mediante la adquisición de nuevos equipos de electrofisiología portátiles (bioamplificadores básicos, de código abierto y bajo coste, adaptados para la docencia), los cuales permiten la ejecución de diferentes protocolos experimentales de forma totalmente independiente por parte de los estudiantes, bajo la constante supervisión y orientación del profesor responsable. En segundo lugar, aplicar el "Principio de las tres R" con el propósito de sustituir el uso de vertebrados por modelos invertebrados (no considerados legalmente animales de experimentación), eliminando así cualquier restricción legal para su manipulación directa por parte de los estudiantes. Concretamente, se seleccionaron lombrices de tierra, un modelo idóneo para la docencia debido a su fácil adquisición, bajo coste de mantenimiento, sencillo manejo y protocolo de anestesia. Como estrategia complementaria, atendiendo al principio de las tres R, se elaboraron materiales interactivos propios (simulaciones virtuales) a partir de registros electrofisiológicos obtenidos por los propios docentes en el laboratorio, en cursos previos.

2. MÉTODO

2.1. Descripción del contexto y de los participantes

Se implementó la práctica de laboratorio "Registro extracelular del potencial de acción compuesto" de 2,5 h de duración, incluida en los contenidos de la asignatura obligatoria FA-I (1.56 ECTS prácticos, 0.96 no presenciales) del Grado en Biología de la UA. Esta labor se llevó a cabo en el segundo cuatrimestre de los cursos 24-25 y 25-26, sobre un total de 96 y 76 estudiantes, respectivamente. Los alumnos pertenecieron a 5 grupos en cada curso.

2.2. Instrumentos

Para llevar a cabo la práctica de registro de biopotenciales en invertebrados se adquirieron bioamplificadores portátiles Neuron SpikerBox Pro de Backyard Brains (Marzullo y Gage, 2012; Shannon et al., 2014; Backyard Brains, EE.UU.; Figura 1A). Un bioamplificador es un dispositivo electrofisiológico, que se utiliza para adquirir y amplificar una señal eléctrica generada, por ejemplo, por los axones del cordón nervioso ventral de una lombriz. La digitalización de este tipo de señales permite, posteriormente, enviarla a diferentes dispositivos (ordenador o móvil) para poder analizar sus características mediante el empleo del software apropiado (Figura 1B). En este caso, los datos bioeléctricos fueron registrados y visualizados en tiempo real utilizando el software Spike Recorder (versión 1.9.38; Backyard Brains, EE.UU). La ventaja que ofrecen los instrumentos de Backyard Brains es que presentan un coste muy reducido (en comparación con el instrumental clásico), son portátiles y su manejo es relativamente sencillo, por lo que los estudiantes pueden familiarizarse rápidamente con su uso, fomentando la participación activa en su aprendizaje y permitiéndoles una amplia autonomía de trabajo. Asimismo, se utilizaron especímenes de lombriz de tierra

(*Lumbricus terrestris sp*., Figura 1C), adquiridos en formato de cebo vivo en el establecimiento comercial Decathlon (comercializados bajo la línea de cebos vivos de "The Worms Cebos"). Las principales razones que justifican su uso como modelo animal de elección, adicionalmente a que son invertebrados y no se requiere legalmente una capacitación por parte de quien los manipula, es que las lombrices son muy económicas, de fácil adquisición y que su cordón nervioso ventral incluye varios axones gigantes y de larga longitud. Además, estos animales pueden mantenerse en la nevera durante semanas sin la necesidad de ningún tipo de cuidado especial. Su anestesia es sencilla, mediante inmersión durante 2-3 minutos en una solución Ringer conteniendo 115 mM NaCl, 2 mM KCl, 1.8 mM $CaCl_2$, 5 mM HEPES (ácido 4-(2-hidroxietil)-1-piperazinaetanosulfónico), 20% etanol y ajustado el pH a 7.0. Por último, constituyen un modelo experimental conveniente y adecuado, debido a que sus neuronas generan las mismas señales eléctricas, denominadas potenciales de acción (PAs, Figura 1B), utilizadas por el sistema nervioso de todos los animales y que subyacen, entre otras funciones neuronales, a la generación de nuestras percepciones sensoriales, pensamientos y movimientos.

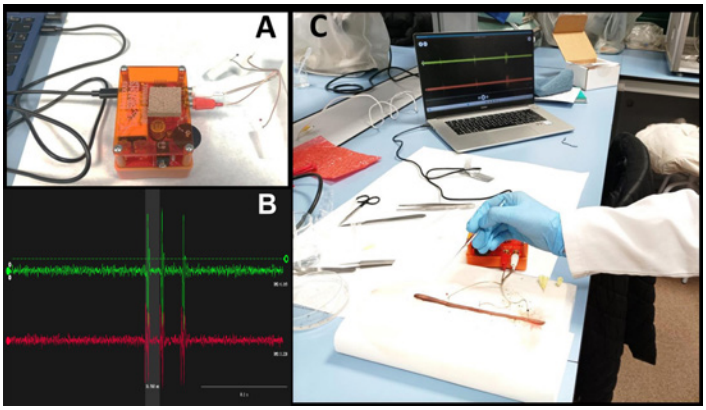


Figura 1. (A) Bioamplificador portátil Neuron SpikerBox Pro (Backyard Brains, EE.UU.) conectado a un PC. Obsérvense los tres electrodos de aguja que permiten registrar señales bioeléctricas cuando se insertan apropiadamente en el cuerpo de un animal. (B) Registro de PAs en una lombriz de tierra, evocados por la aplicación de un estímulo de presión. (C) Disposición de los equipos empleados y del animal para el registro de PAs tras la aplicación de un estímulo mecánico.

Paralelamente, se elaboró el material interactivo titulado "Simulación virtual de registro de potencial de acción compuesto" (Figura 2A) correspondiente a esa misma práctica, mediante el empleo de PowerPoint, en el que se incluyeron imágenes y registros obtenidos en cursos previos. Este recurso fue alojado en la plataforma UACloud, con el objetivo de que los estudiantes pudieran refinar los procedimientos antes de la experimentación y, posteriormente, tras haber realizado la práctica, consolidar de forma autónoma los conocimientos teóricos y prácticos adquiridos. La simulación consta de una parte introductoria y metodológica (Figura 2B y 2C) que pretende exponer de forma simple los principios básicos, técnicas y materiales necesarios para llevar a cabo registros de señales como los PAs. A continuación, los estudiantes pueden acceder a 6 ejercicios prácticos interactivos que utilizan registros reales (Figura 2D, 2E y 2F).

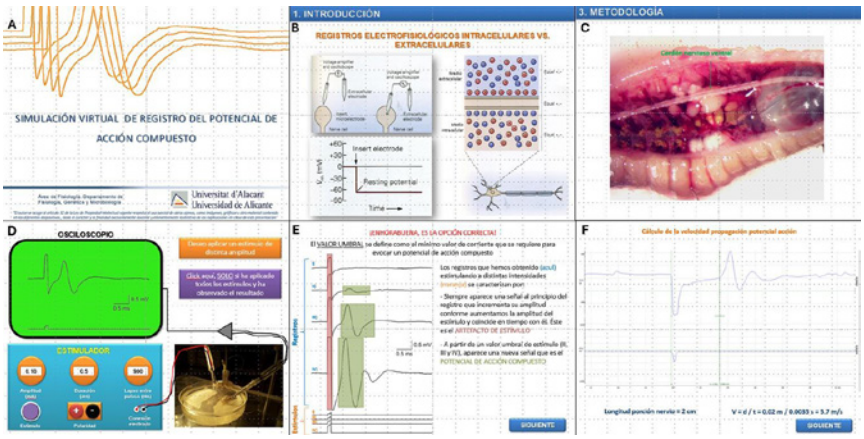


Figura 2. Capturas de pantalla del material interactivo titulado "Simulación virtual de registro de potencial de acción compuesto".

Como instrumento de análisis para evaluar la experiencia educativa se desarrolló un cuestionario de 6 preguntas con las 5 posibles respuestas siguientes: 1) muy en desacuerdo, 2) en desacuerdo, 3) ni de acuerdo, ni en desacuerdo, 4) de acuerdo, y 5) muy de acuerdo. Las 6 preguntas en cuestión fueron las siguientes:

1) El uso de bioamplificadores portátiles me ha facilitado el aprendizaje de los objetivos de la práctica.
2) El uso de bioamplificadores portátiles me ha facilitado a entender cómo funcionan los equipos electrofisiológicos.
3) El empleo de materiales interactivos me ha ayudado a comprender los conceptos teórico-prácticos explicados en la sesión práctica.
4) Considero que esta nueva forma de trabajo fomenta el aprendizaje autónomo.
5) Esta nueva forma de trabajo aumenta mi interés por la asignatura.
6) En caso de que hayas realizado esta práctica en cursos anteriores, ¿consideras que la nueva metodología de registro electrofisiológico aumenta el interés por la actividad y facilita el aprendizaje con respecto a las demostraciones realizadas, en cursos anteriores, por el profesor?

2.3. Procedimiento

En cada grupo de prácticas, los estudiantes fueron distribuidos en equipos de 3-4 personas, cada uno de ellos debía contar al menos con un PC portátil y/o un teléfono móvil en el que instalaron el software Spike Recorder (Backyard Brains, EE.UU.). El primer objetivo de cada grupo fue el de registrar PAs (Figura 1B), evocados en axones del cordón nervioso ventral del anélido, en respuesta a un estímulo mecánico de presión aplicado con una varilla de vidrio sobre la superficie del animal. Para ello, cada equipo anestesió un ejemplar de lombriz y, posteriormente, el animal fue colocado sobre una plataforma de corcho situando la porción dorsal hacia arriba. A continuación, los estudiantes insertaron en el cuerpo de la lombriz tres electrodos de aguja (dos electrodos activos de registro y uno de referencia) conectados a un bioamplificador Neuron SpikerBox Pro (Backyard Brains, EE.UU.). Los electrodos fueron implantados ligeramente desviados de la línea central del

anélido, con el fin de evitar perforar el intestino o el cordón nervioso ventral. Las figuras 1C y 3 muestran la posición de los electrodos y los equipos empleados, así como el esquema del procedimiento empleado en detalle. Para evocar el disparo de PAs se aplicaron uno o varios estímulos mecánicos presionando la superficie anterior del animal con una varilla de vidrio. Las señales eléctricas (PAs) evocadas en respuesta al estímulo viajaron a través del cordón nervioso ventral, desde la parte anterior o cefálica del animal hacia la posterior o caudal. Ello permitió registrar, en un primer momento, PAs a través del electrodo 1 y, posteriormente, transcurridos pocos milisegundos después, las mismas señales a través del electrodo 2 (Figura 3). A partir de los registros obtenidos de los PAs, los estudiantes determinaron, en segundo lugar, la velocidad de conducción de estas señales propagadas a través de los axones de la lombriz (Figura 3), a partir de la siguiente ecuación:

$$V = d / t$$

Donde V (en m/s) es la velocidad de conducción de PAs, d (en m) es la distancia entre el electrodo 1 y el electrodo 2, y t (en s) es la diferencia temporal entre los PAs registrados por el electrodo 1 y el 2.

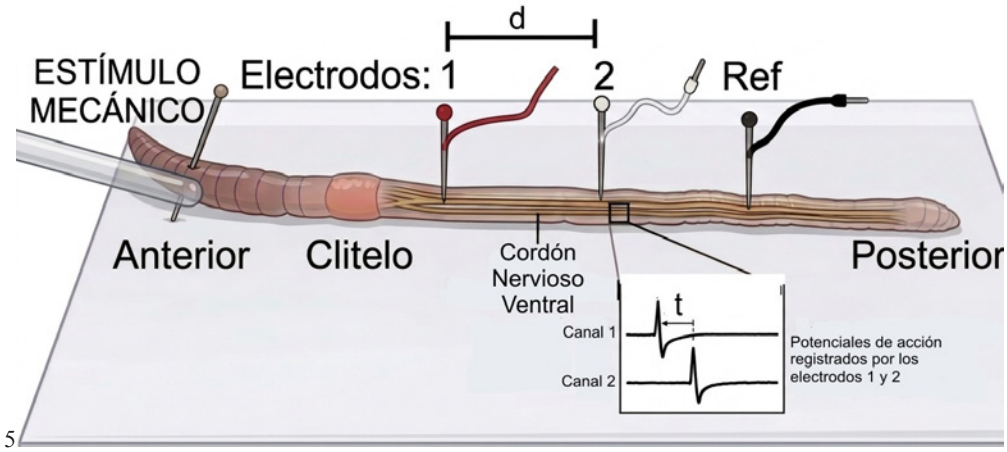


Figura 3. Disposición de los electrodos de registro (1 y 2) y de referencia (Ref) insertados en el cuerpo de la lombriz de tierra. La distancia (d) entre los electrodos 1 y 2 fue medida con una regla por los estudiantes de cada equipo. Imagen generada por los autores mediante la IA Gemini (Google, 2026).

Antes de finalizar cada sesión práctica, en los últimos 15 minutos, se entregó a cada estudiante el cuestionario descrito previamente, con el fin de evaluar la implementación de la experiencia educativa. En el curso 25-26, solo respondieron el cuestionario aquellos alumnos que realizaron la práctica implementada por primera vez, absteniéndose aquéllos que el curso anterior ya la habían realizado mediante este nuevo formato. Los cuestionarios fueron completados de forma totalmente anónima y entregados al profesor. Posteriormente, se contabilizaron las respuestas y se elaboraron gráficos de barras mediante el empleo del software Origin v6.1 (OriginLab Corp., Northampton, MA, EE.UU.). Los valores representados en los gráficos se corresponden con la media ± error estándar de la media (e.s.m.). El número de alumnos encuestados se representa como n. Se realizó el análisis de la varianza (ANOVA) para determinar diferencias entre grupos, cuando no se cumplió la normalidad se aplicó el test de Kruskal-Wallis y el test de Dunn para realizar comparaciones múltiples. Se consideró un nivel de significación de $p<0.05$.

3. RESULTADOS

La evaluación global de esta experiencia, tras el análisis de las respuestas de los cuestionarios por parte del alumnado, resultó positiva. En el Gráfico 1 se muestran las respuestas del estudiantado, durante el curso 24-25, a las 6 preguntas del cuestionario para evaluar la implementación de la práctica. Las barras negras representan las respuestas del total de estudiantes de ese curso (n=96), las rojas las de aquellos que cursaron por primera vez la asignatura (n=61) y las verdes las de los que ya habían cursado la asignatura al menos una vez (n=35). Tanto en el Gráfico 1 como en el 2, los valores del eje de ordenadas (Valoración) se corresponden con: 1 = "muy en desacuerdo", 2 = "en desacuerdo", 3 = "ni de acuerdo ni en desacuerdo", 4 = "de acuerdo" y 5 = "muy de acuerdo"). Así, los estudiantes del curso 24-25 indicaron que esta nueva forma de trabajo fomentó su aprendizaje autónomo con una puntuación de 4.51±0.08, siendo 5 la puntuación máxima (Gráfico 1, barra negra en pregunta 4; n=96). No obstante, la valoración de esta cuestión fue mayor por parte del alumnado que cursaba la asignatura por primera vez (4.61±0.09, Gráfico 1, barra roja en pregunta 4; n=61) que por el estudiantado que ya la había realizado en cursos anteriores, aunque en la modalidad de demostración por parte del profesor (4.34±0.15, Gráfico 1, barra verde en pregunta 4; n=35). Asimismo, señalaron que el uso de los bioamplificadores portátiles les facilitó tanto el aprendizaje de los objetivos de la práctica (4.65±0.06, n=96; Gráfico 1, barra negra en pregunta 1) como el entendimiento del funcionamiento de dichos equipos (4.48±0.08, n=96; Gráfico 1, barra negra en pregunta 2). De nuevo, en ambas cuestiones, la valoración fue superior por parte del alumnado novel (Gráfico 1, barras rojas en pregunta 1 y 2, 4.72±0.07 y 4.62±0.07, respectivamente; n=61) que por el veterano (Gráfico 1, barras verdes en pregunta 1 y 2, 4.51±0.10 y 4.23±0.16, respectivamente; n=35). Respecto al uso de materiales interactivos, los estudiantes indicaron de forma general que estos recursos les ayudaron a comprender los conceptos teórico-prácticos explicados en la sesión práctica (4.65±0.06, Gráfico 1, barra negra en pregunta 3; n=96). De hecho, los estudiantes de primer año valoraron esta cuestión con la mayor puntuación respecto al resto de preguntas (Gráfico 1, barra roja en pregunta 3, 4.79±0.06; n=61), mientras que los estudiantes veteranos la valoraron con un 4.40±0.11 (Gráfico 1, barra verde en pregunta 3; n=35). Además, de manera global, el alumnado manifestó que este modelo de trabajo innovador aumentó su interés por la asignatura (Gráfico 1, barra negra en pregunta 5; 4.15±0.10, n=96). Sin embargo, fue la pregunta menos valorada tanto por parte de los estudiantes noveles como por los veteranos (4.38±0.11 y 3.74±0.20, barras rojas y verdes en Gráfico 1, respectivamente). Finalmente, los 36 alumnos que habían realizado la práctica en cursos anteriores en forma de demostración consideraron que la nueva metodología aumentaba su interés y facilitaba el aprendizaje respecto a las prácticas clásicas realizadas por el profesor (Gráfico 1, barra verde en pregunta 6; 4.54±0.10).

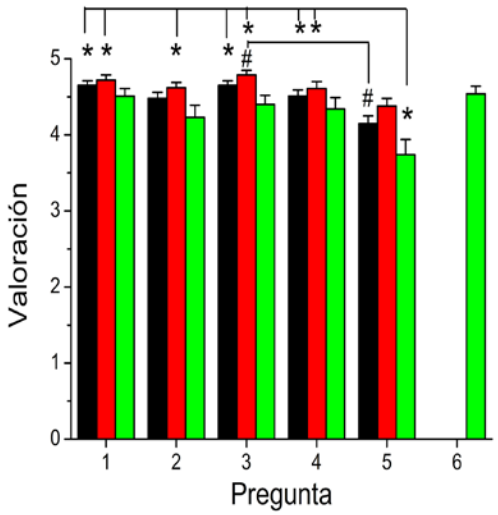


Gráfico 1. Valoración del estudiantado, durante el curso 24-25, de la implementación de la práctica. (#) muestra diferencias significativas entre los grupos señalados, (*) revela diferencias significativas entre los resultados de la pregunta en 5 en estudiantes veteranos y el resto de grupos señalados ($p<0.05$, ANOVA y test de Dunn).

El Gráfico 2 muestra las respuestas del estudiantado, durante el curso 25-26, a las preguntas del cuestionario para evaluar la implementación de la práctica. Las barras grises representan las respuestas del total de estudiantes de ese curso (n=76), las naranjas las de aquellos que cursaron por primera vez la asignatura (n=56) y las azules las de los que ya habían cursado la asignatura al menos una vez (n=20). En este curso, los resultados fueron muy similares al anterior. Así, los estudiantes de este año académico respondieron a la pregunta 4 que esta nueva forma de trabajo fomentó su aprendizaje autónomo con una puntuación de 4.63±0.09 (barra gris en pregunta 4 de Gráfico 2; n=76). Además, como ocurrió con los estudiantes del curso anterior, la valoración de esta cuestión fue mayor por parte del alumnado que cursaba la asignatura por primera vez (4.67±0.09, Gráfico 2, barra naranja en pregunta 4; n=56) que por el estudiantado que ya la había realizado en cursos anteriores, en la modalidad de demostración por parte del profesor (4.53±0.21, Gráfico 2, barra azul en pregunta 4; n=20). Asimismo, señalaron que el uso de los bioamplificadores portátiles les facilitó tanto el aprendizaje de los objetivos de la práctica (4.47±0.09, n=76; Gráfico 2, barra gris en pregunta 1) como el entendimiento del funcionamiento de dichos equipos (4.47±0.08, n=76; Gráfico 2, barra gris en pregunta 2). No obstante, a diferencia de los resultados del curso 24-25, en ambas cuestiones, la valoración fue inferior por parte del alumnado novel (Gráfico 2, barras naranjas en pregunta 1 y 2, 4.44±0.10 y 4.40±0.10, respectivamente; n=56) que por el veterano (Gráfico 2, barras azules en pregunta 1 y 2, 4.58±0.18 y 4.68±0.16, respectivamente; n=20). En relación con el empleo de materiales interactivos, los estudiantes señalaron de forma general que estos recursos les facilitaron la comprensión de los conceptos teórico-prácticos explicados en la sesión práctica (4.68±0.07, Gráfico 2, barra gris en pregunta 3; n=76). En esta cuestión, los estudiantes veteranos valoraron esta cuestión con la mayor puntuación respecto al resto de preguntas (Gráfico 2, barra azul en pregunta 3, 4.79±0.10; n=20), mientras que los estudiantes de primer año la valoraron con un 4.65±0.08 (Gráfico 2, barra naranja en pregunta 3; n=56). Asimismo, de manera global, el alumnado del curso 25-26 manifestó que este modelo de trabajo innovador aumentó su interés por la asignatura (Gráfico 2, barra gris en pregunta 5; 4.37±0.11, n=76). De nuevo, fue la pregunta menos valorada tanto por parte de los estudiantes

veteranos (4.11±0.25, barra azul en pregunta 5 de Gráfico 2; n=20) mientras que los noveles la puntuaron con un 4.46±0.11 (barra naranja en pregunta 5 de Gráfico 2; n=56). Finalmente, los 20 alumnos veteranos del año académico 25-26, que en cursos anteriores habían realizado la práctica en grupos en los que se impartía en forma de demostración, consideraron que la nueva metodología aumentaba su interés y facilitaba el aprendizaje respecto a las prácticas clásicas realizadas por el profesor (Gráfico 2, barra azul en pregunta 6; 4.50±0.19).

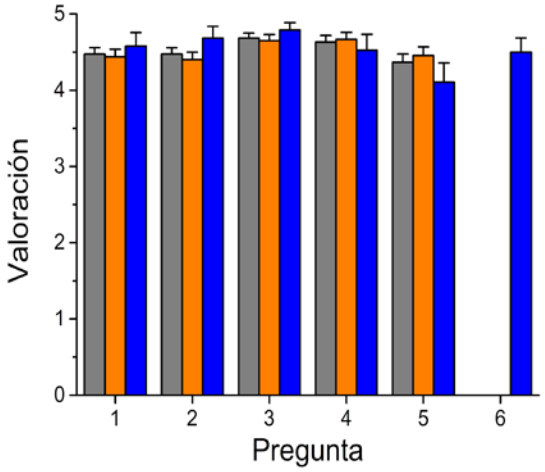


Gráfico 2. Valoración del estudiantado, durante el curso 25-26, de la implementación de la práctica p>0.05 en ANOVA y test Kruskal-Wallis.

4. DISCUSIÓN Y CONCLUSIONES

El rediseño e implementación de la práctica de laboratorio "Registro extracelular del potencial de acción compuesto" en la asignatura FA-I de la UA demuestra que es posible conjugar el cumplimiento estricto del marco legal actual con una mejora sustancial que aboga por la autonomía del estudiante en su aprendizaje. Los resultados obtenidos a lo largo de los cursos 24-25 y 25-26, a partir de las respuestas de los estudiantes a los cuestionarios, validan empíricamente que la transición desde un modelo de demostración pasiva hacia una metodología activa y autónoma es posible. Además, apoya la posibilidad y sirve como incentivo para que en el futuro se implementen otras prácticas de la misma asignatura que actualmente se imparten como demostración clásica.

El objetivo principal de promover el aprendizaje autónomo a través de la "Enseñanza Centrada en el Estudiante" (SCL), respaldado por la European Students' Union (2015), se vio reflejado en las elevadas puntuaciones otorgadas por el alumnado a la pregunta 4 (4.51 y 4.63, en los respectivos cursos). Al devolver la capacidad de manipulación directa al alumnado a través de un modelo alternativo, se rompe con la pasividad de las prácticas demostrativas, permitiendo que el estudiante se implique directamente en el análisis e interpretación de fenómenos fisiológicos reales, de manera independiente (Balaguera et al., 2023). Esto queda fehacientemente demostrado en las valoraciones del alumnado veterano (pregunta 6: 4.54 y 4.50, en los años académicos 24-25 y 25-26, respectivamente), quienes experimentaron ambos formatos y confirmaron de manera contundente que la nueva metodología activa facilita significativamente el aprendizaje en comparación con las demostraciones clásicas guiadas por el profesor.

Respecto al primer sub-objetivo, centrado en el uso de tecnología electrofisiológica portátil de bajo coste, las opiniones recogidas sobre los dispositivos empleados revelan un impacto metodológico muy favorable. Las preguntas 1 y 2 obtuvieron puntuaciones consistentemente altas (>4.40), lo que refrenda la viabilidad de la tecnología de código abierto para sortear las barreras económicas y los costes logísticos prohibitivos del instrumental de laboratorio clásico descritos por Doke y Dhawale (2015).

En cuanto al segundo sub-objetivo —la aplicación del "Principio de las tres R" (Russell y Burch, 1959; Derakhchan et al., 2026)—, la sustitución de vertebrados por invertebrados (lombrices) resolvió satisfactoriamente los dilemas éticos y las limitaciones legales sin que disminuya la calidad del aprendizaje o el rigor científico. A nivel biológico, el modelo permitió estudiar principios fundamentales de la transmisión de señales bioeléctricas de conducción nerviosa, cumpliendo con los propósitos teóricos descritos por Billman (2020) y Kuang (2023). El uso de invertebrados no protegidos no solo elimina los conflictos regulatorios, sino que además demostró ser una alternativa sumamente económica, duradera y de fácil mantenimiento.

Esta estrategia de reemplazo y refinamiento se vio complementada de forma sobresaliente por el material interactivo virtual desarrollado en UACloud. Los estudiantes valoraron de forma masiva este recurso (pregunta 3: 4.65 y 4.68, cursos 24-25 y 25-26, respectivamente), destacando la puntuación de 4.79 entre los estudiantes noveles del primer año y los veteranos del segundo. Esto demuestra que la simulación virtual, previa al laboratorio, disminuye la incertidumbre procedimental y puede servir como una herramienta de indudable valor para consolidar de forma autónoma los conceptos teórico-prácticos antes y después de la experimentación real.

Finalmente, es importante reseñar que la pregunta 5, relativa al aumento del interés por la asignatura, fue la que obtuvo de manera sistemática las puntuaciones más discretas en ambos cursos, oscilando entre 3.74 y 4.46. Este fenómeno es habitual en asignaturas con una alta carga conceptual y exigencia técnica como la Fisiología Animal. Sin embargo, el hecho de que las puntuaciones medias se mantengan firmemente en el rango de "de acuerdo" (valores superiores a 4 de manera global) sugiere que, si bien una única innovación práctica no transforma radicalmente la percepción general de una asignatura, sí actúa como un dinamizador motivacional clave dentro del plan de estudios del Grado en Biología.

A partir de los objetivos planteados y de la evaluación de la experiencia de innovación educativa implementada, se desprenden las siguientes conclusiones:

1) Cumplimiento del Enfoque SCL y Aprendizaje Activo: Se promovió con éxito el aprendizaje autónomo y activo de los estudiantes del Grado en Biología, logrando desplazar el antiguo modelo de demostración pasiva del docente hacia una metodología de manipulación directa, altamente valorada por el alumnado (especialmente por aquellos que conocían el formato clásico).
2) Eficacia del "Principio de las Tres R" mediante Reemplazo: El uso de la lombriz de tierra como modelo invertebrado no protegido demostró ser una alternativa didáctica y científica excelente para sustituir el uso de vertebrados. Se eliminaron los impedimentos legales y los costes prohibitivos, garantizando que el estudiante sea el verdadero protagonista de la experimentación electrofisiológica. Además se contribuye a que el estudiantado aprenda a refinar los procedimientos de laboratorio mediante la anestesia, cirugía y experimentación adecuada de estos invertebrados.
3) Validación Tecnológica de Bajo Coste: La adquisición e integración de los bioamplificadores portátiles demostró que es viable dotar a los laboratorios docentes de herramientas tecnológicas accesibles y de código abierto que faciliten la comprensión real de los registros extracelulares y el cálculo de la velocidad de conducción nerviosa sin mermar la calidad académica.

4) Sinergia Didáctica con Recursos Virtuales: El diseño y alojamiento del material interactivo virtual en UACloud funcionó como una estrategia sumamente eficaz. Consolidó una de las valoraciones más altas del proyecto, demostrando ser un recurso clave para la preparación autónoma previa y el afianzamiento conceptual posterior a la sesión práctica.
5) Impacto Positivo General: La innovación educativa fue evaluada de manera abrumadoramente positiva en dos cohortes consecutivas (96 y 76 estudiantes), confirmando que esta nueva estructura metodológica no solo aumenta el interés por la materia, sino que establece un precedente transferible para el rediseño de otras prácticas de laboratorio en ciencias biosanitarias.

5. REFERENCIAS

Balaguera Quinche, D. F., et al. (2023). Didactic biomodels in animal physiology learning: implementation in undergraduate veterinary medicine. Revista de Ciências Agroveterinárias, 22(1), 163–172.

Billman GE. (2020). Homeostasis: The Underappreciated and Far Too Often Ignored Central Organizing Principle of Physiology. Frontiers in Physiology. 11: 200. doi: 10.3389/fphys.2020.00200

Derakhchan K, Rockley K, Gossmann M, Fares R, Delpy E, Kanda Y, Passini E, Pike CM, Turner J and Delaunois A (2026). 25 years of applying the 3Rs principles in safety pharmacology: success stories and future perspectives. Front. Physiol. 17:1813708.doi: 10.3389/fphys.2026.1813708

Doke, S. K., & Dhawale, S. C. (2015). Alternatives to animal testing: A review. Saudi Pharmaceutical Journal, 23(3), 223-229.

European Students' Union (ESU). (2015). Student-Centred Learning: Toolkit for students, staff and higher education institutions. Bruselas: ESU / Education International.

Google. (2026). Gemini (versión Nano Banana 2, 26 de junio de 2026) [Modelo de IA generativa de imágenes]. https://gemini.google.com

Kuang SY. (2023) Advancing physiology education by understanding the multiple dimensions of homeostasis. Front. Physiol. 14: 1234214. doi: 10.3389/fphys.2023.1234214

Liang Y, Song Y, Chen M, Zhang X, Wang Y, Lin J, Zhou W, Zhang J, Wang W, Andriani LTS, Liang M, Khan SA, Carere C. (2026) Rethinking invertebrate welfare in science. Science, 392 (6801): 926. doi: 10.1126/science.aeh3585.

Marzullo, T. C., & Gage, G. J. (2012). The SpikerBox: A low cost, open-source bioamplifier for increasing public participation in neuroscience inquiry. PLoS ONE, 7(3), e30837. doi: 10.1371/journal.pone.0030837.

Medler, S. (2019) Anesthetic MS-222 eliminates nerve and muscle activity in frogs used for physiology teaching laboratories: Adv. Physiol. Edu. 43:69-75. doi:10.1152/advan.00114.2018.

Russell W.M.S. & Burch R.L. (1959). The Principles of Humane Experimental Technique. Methuen, London.

Sefton, A. (2001). International workshop: modern approaches to teaching and learning physiology. Adv. Physiol. Educ. 25: S64-S66.

Shannon KM, Gage GJ, Jankovic A, Wilson WJ, Marzullo TC. (2014). Portable conduction velocity experiments using earthworms for the college and high school neuroscience teaching laboratory. Adv Physiol Educ. 38 (1): 62-70. doi: 10.1152/advan.00088.2013.

2. Diagnosticar para innovar: perfil basal del engagement, pensamiento crítico y cohesión grupal en una muestra de estudiantes universitarios

J. A. Antón-Ruiz, V. Clement Carbonell; L. J. Viñuelas Barberá, N. Albaladejo Blázquez, N. Ruiz Robledillo, I. Portilla Tamarit

Universidad de Alicante

RESUMEN

El engagement académico, el pensamiento crítico y la cohesión grupal constituyen competencias transversales de especial relevancia en titulaciones sanitarias como Logopedia. Este estudio presenta el diagnóstico basal de dichas variables en estudiantes de primer curso del Grado en Logopedia de la Universidad de Alicante, con el fin de fundamentar una innovación docente basada en Aprendizaje Basado en Retos (ABR). Se realizó un estudio descriptivo-exploratorio con 57 participantes (93,0% mujeres; M = 19,98 años, DE = 3,58). La cohesión grupal se evaluó con la versión española del University Group Environment Questionnaire (UGEQ; M = 6,58, DE = 1,11), el engagement con la Utrecht Work Engagement Scale for Students (UWES-S; M = 3,57, DE = 0,86) y el pensamiento crítico con el Watson-Glaser Critical Thinking Appraisal, Form S (WGCTA-FS; M = 29,18, DE = 5,52). Los resultados muestran una cohesión orientada a la tarea grupal, un engagement de nivel medio con dedicación superior al vigor y la absorción, y un perfil de pensamiento crítico con mayor competencia en deducciones y menor dominio en reconocimiento de suposiciones. Se identificó una correlación negativa de efecto mediano-grande entre edad y cohesión grupal ($r_s = -0{,}469$, $p < 0{,}001$), así como una asociación positiva entre implicación individual en tareas grupales y vigor. Estos hallazgos fundamentan decisiones pedagógicas específicas para ajustar el ABR al perfil real del estudiantado.

PALABRAS CLAVE: engagement académico; pensamiento crítico; cohesión grupal; aprendizaje basado en retos; innovación docente.

1. INTRODUCCIÓN

La formación universitaria en ciencias de la salud plantea un doble desafío: transmitir conocimiento disciplinar especializado y desarrollar simultáneamente competencias transversales que habiliten al estudiantado para el ejercicio profesional en entornos complejos, multidisciplinares y cambiantes. En titulaciones como Logopedia, donde la intervención abarca desde la atención temprana hasta la rehabilitación en la tercera edad y donde las decisiones clínicas deben sustentarse en evidencia científica y coordinarse con otros profesionales, el engagement académico, el pensamiento crítico y la cohesión grupal adquieren una relevancia formativa especial (De la Cruz-Velazco et al., 2022; Ruiz-Robledillo et al., 2023).

El engagement académico ha sido conceptualizado como un estado mental positivo, persistente y vinculado específicamente a los estudios, caracterizado por vigor, energía mental y resiliencia ante las dificultades, dedicación, sentido de significado, entusiasmo e identificación con la carrera, y absorción, concentración profunda y dificultad para desconectarse de las tareas (Schaufeli et al., 2002). Estudios longitudinales han mostrado que el engagement se asocia positivamente con el

rendimiento académico, el bienestar psicológico y la satisfacción con los estudios, y actúa como factor protector frente al burnout (Carmona-Halty et al., 2019; Salanova et al., 2010). Resulta especialmente relevante que las tres subescalas presentan perfiles diferenciados en función del curso y del perfil del estudiante, lo que convierte su evaluación basal en un instrumento diagnóstico de gran utilidad pedagógica (Bresó et al., 2007).

El pensamiento crítico constituye una competencia transversal reconocida en los marcos europeos de educación superior y en los criterios de acreditación de titulaciones sanitarias. En las profesiones de la salud implica detectar supuestos implícitos en los razonamientos diagnósticos, evaluar la calidad de la evidencia disponible, extraer conclusiones válidas y justificar decisiones clínicas ante equipos interdisciplinares (Facione, 1990; Profetto-McGrath, 2005). El Watson-Glaser Critical Thinking Appraisal, Form S (WGCTA-FS; Watson & Glaser, 1994), validado en población universitaria española (De Mangione & De Anglat, 2007), evalúa cinco dimensiones: inferencia, reconocimiento de suposiciones, deducciones, interpretación y evaluación de argumentos.

La cohesión grupal en el aula universitaria puede conceptualizarse a través de dos ejes: la implicación individual, en sus dimensiones de tarea (ATG-T) y social (ATG-S), y la percepción de integración grupal, igualmente en tarea (GI-T) y social (GI-S), según el modelo del University Group Environment Questionnaire (UGEQ), validado en español por Ruiz-Robledillo et al. (2023). La investigación previa ha mostrado que una cohesión orientada a la tarea facilita la coordinación y la adquisición colectiva de competencias, mientras que la dimensión social refuerza el bienestar emocional en el aula (Bosselut et al., 2018).

En este marco, el Aprendizaje Basado en Retos (ABR) emerge como una metodología activa con potencial para fortalecer simultáneamente estas tres competencias. Originariamente propuesta por Apple (2011), el ABR involucra al estudiantado en situaciones problemáticas reales, relevantes y vinculadas a su entorno, promoviendo respuestas colaborativas, innovadoras y reflexivas que integran conocimiento teórico y práctica profesional. La evidencia disponible sobre su eficacia en educación superior, aún incipiente en ciencias de la salud, sugiere mejoras en motivación, pensamiento crítico y trabajo en equipo (Agüero et al., 2019; De la Cruz-Velazco et al., 2022). Un prerrequisito metodológico ineludible para la implementación eficaz del ABR es conocer el perfil real del grupo antes de la intervención. Sin ese diagnóstico basal, el diseño pedagógico se ajusta a un estudiante ideal en lugar de al estudiantado concreto (Bolaños, 2019).

El presente trabajo tiene como objetivo describir el perfil basal de engagement, pensamiento crítico y cohesión grupal en estudiantes de primer curso del Grado en Logopedia de la Universidad de Alicante, analizar las relaciones entre estas variables y explorar su asociación con características sociodemográficas relevantes, especialmente la edad y la situación de convivencia, como fundamento empírico para el diseño de una innovación docente basada en ABR. Más allá de la intención descriptiva, el estudio pretende aportar un conocimiento no documentado previamente en esta población: la relación entre la edad del estudiantado y su percepción de cohesión grupal, una cuestión escasamente explorada en la literatura sobre competencias transversales en el área de ciencias de la salud y con implicaciones directas para la integración del alumnado.

2. MÉTODO

2.1. Descripción del contexto y de los participantes

El estudio se desarrolló en el contexto de la asignatura Psicología, materia de formación básica del primer curso y primer cuatrimestre del Grado en Logopedia de la Universidad de Alicante, mediante muestreo no probabilístico por conveniencia con participación voluntaria y consentimiento informado.

Inicialmente participaron 60 estudiantes. Tras excluir un registro de prueba y dos respuestas duplicadas identificadas mediante el código de seguimiento, la muestra final quedó constituida por N = 57 participantes (93,0% mujeres; 7,0% varones), con edades comprendidas entre 17 y 42 años (M = 19,98, DE = 3,58). La Figura 1 muestra la distribución según situación académico-laboral y modalidad de convivencia.

2.2. Instrumentos

- Cuestionario sociodemográfico ad hoc para la recogida de edad, sexo, situación de convivencia y ocupación académico-laboral.
- Versión española del University Group Environment Questionnaire (UGEQ; Ruiz-Robledillo et al., 2023). Instrumento de 16 ítems con escala Likert de 9 puntos que evalúa la cohesión grupal en cuatro dimensiones: ATG-T (4 ítems), ATG-S (5 ítems), GI-T (3 ítems) y GI-S (4 ítems). En la validación española, los coeficientes alfa oscilan entre 0,72 y 0,86. En la muestra de este trabajo se obtuvieron: αtotal = 0,859; αATG-T = 0,761; αATG-S = 0,740; αGI-T = 0,566; αGI-S = 0,817. El valor moderado de α en GI-T es esperable dado el número reducido de ítems (k = 3).
- Utrecht Work Engagement Scale for Students (UWES-S; Schaufeli et al., 2002). Compuesta por 17 ítems con escala de respuesta de 7 puntos (0 = nunca; 6 = siempre), distribuidos en Vigor (VI, 6 ítems), Dedicación (DE, 5 ítems) y Absorción (AB, 6 ítems). En la validación original, α oscila entre 0,63 y 0,81. En la muestra de este estudio: αtotal = 0,875; αVI = 0,735; αDE = 0,849; αAB = 0,750.
- Watson-Glaser Critical Thinking Appraisal, Form S (WGCTA-FS; Watson & Glaser, 1994). Instrumento de 40 ítems dicotómicos distribuidos en cinco dimensiones de 8 ítems cada una: inferencia, reconocimiento de suposiciones, deducciones, interpretación y evaluación de argumentos. La consistencia interna se evalúa mediante el coeficiente KR-20, equivalente al alfa de Cronbach para ítems dicotómicos. En muestras universitarias españolas se reporta α = 0,82 (De Mangione & De Anglat, 2007). En la presente muestra: KR-20total = 0,810; por dimensiones, los valores oscilan entre 0,40 (Suposiciones) y 0,73 (Evaluación de argumentos), rango esperable con k = 8 ítems por subescala.

2.3. Procedimiento

La recogida de datos se realizó mediante un formulario online anónimo (Google Forms) durante el primer cuatrimestre del curso 2025-26. Cada participante generó un código alfanumérico de seguimiento que garantizaba el anonimato. El análisis estadístico se realizó con IBM SPSS Statistics v.29. La normalidad se evaluó con la prueba de Shapiro-Wilk, y dado que no se cumplía dicho supuesto en varias variables, se emplearon la correlación de Spearman (rs), la prueba U de Mann-Whitney y la prueba de Kruskal-Wallis para los análisis comparativos y relacionales, reportando los tamaños del efecto rs y η^2H. El nivel de significación se fijó en α = 0,05.

3. RESULTADOS

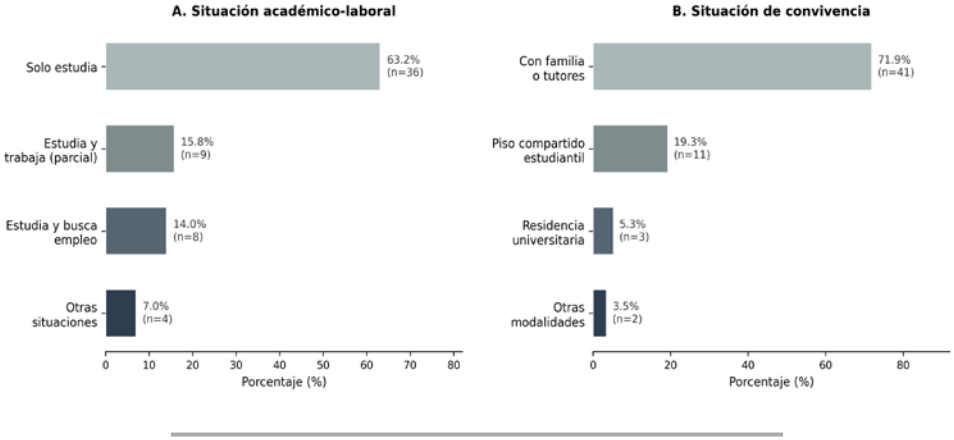


Figura 1. Perfil sociodemográfico de la muestra (N = 57)

Nota. Panel A: distribución por situación académico-laboral. Panel B: distribución por modalidad de convivencia.

La prueba de Shapiro-Wilk indicó distribución no significativamente distinta de la normal para UGEQ (W = 0,98, p = 0,52) y UWES-S global (W = 0,96, p = 0,06), pero sí para el WGCTA-FS total (W = 0,93, p = 0,002) y varias subescalas. Se emplearon estadísticos no paramétricos en todos los análisis comparativos y correlacionales. La Tabla 1 recoge los estadísticos descriptivos.

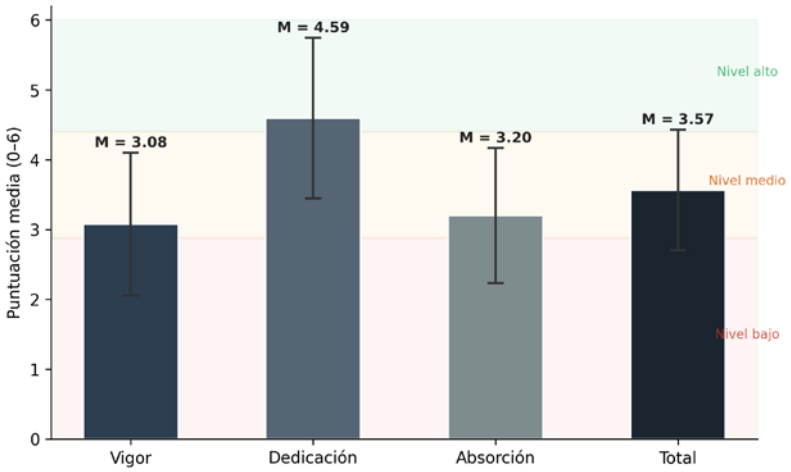
Tabla 1. Estadísticos descriptivos de los instrumentos (N = 57)

Variable	M	DE	Rango observado
UGEQ (cohesión grupal global)	6,58	1,11	3,8 – 8,8
ATG-T (implicación individual – tarea)	6,94	1,34	2,5 – 9,0
ATG-S (implicación individual – social)	6,36	1,50	1,6 – 8,8
GI-T (integración grupal – tarea)	7,68	1,10	5,0 – 9,0
GI-S (integración grupal – social)	5,67	1,75	1,5 – 9,0
UWES-S (engagement global)	3,57	0,86	2,00 – 5,29
Vigor	3,08	1,03	1,17 – 5,17
Dedicación	4,59	1,15	1,00 – 6,00
Absorción	3,20	0,97	1,67 – 5,50
WGCTA-FS (pensamiento crítico global)	29,18	5,52	13 – 40
Inferencia	5,68	1,63	1 – 8
Reconocimiento de suposiciones	4,51	1,66	1 – 8
Deducciones	7,28	1,13	3 – 8
Interpretación	5,60	1,61	1 – 8
Evaluación de argumentos	6,11	1,82	1 – 8

Nota. UGEQ: escala de respuesta 1-9. UWES-S: escala 0-6; niveles normativos: bajo < 2,88; medio = 2,88-4,40; alto > 4,40 (Schaufeli & Bakker, 2004). WGCTA-FS: cada subescala con rango teórico 0-8; total 0-40.

En cuanto a la cohesión grupal (Tabla 1), la puntuación global del UGEQ refleja un clima de grupo positivo. El perfil por subescalas muestra que la integración grupal orientada a la tarea (GI-T: M = 7,68) y la implicación individual en las tareas del grupo (ATG-T: M = 6,94) son las dimensiones más desarrolladas, mientras que la integración social (GI-S: M = 5,67) se sitúa en valores más moderados, coherentes con el inicio del curso.

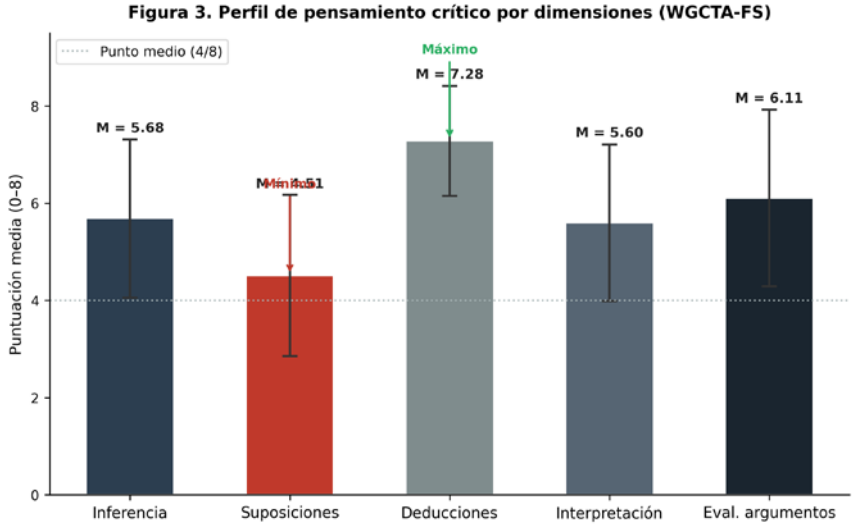
Figura 2. Perfil de engagement académico (UWES-S)



Nota. Las bandas de color indican los niveles normativos (Schaufeli & Bakker, 2004).
Las barras de error representan ±1 DE.

El engagement académico se sitúa en nivel medio (M = 3,57, DE = 0,86), con el 54,4% de los participantes en el rango normativo medio, el 22,8% en el bajo y el 22,8% en el alto. La Figura 2 ilustra el perfil asimétrico entre subescalas: la dedicación (M = 4,59) se aproxima al nivel alto, mientras el vigor (M = 3,08) y la absorción (M = 3,20) permanecen en niveles intermedios.

Figura 3. Perfil de pensamiento crítico por dimensiones (WGCTA-FS)



Nota. La línea punteada indica el punto medio de la escala (4/8). Las barras de error representan ±1 DE. La dimensión Suposiciones (en rojo) obtiene la puntuación más baja.

El perfil del WGCTA-FS (Figura 3) muestra una competencia especialmente elevada en deducciones (M = 7,28) y, en menor medida, en evaluación de argumentos (M = 6,11), inferencia (M = 5,68) e interpretación (M = 5,60). La dimensión de reconocimiento de suposiciones obtiene la puntuación

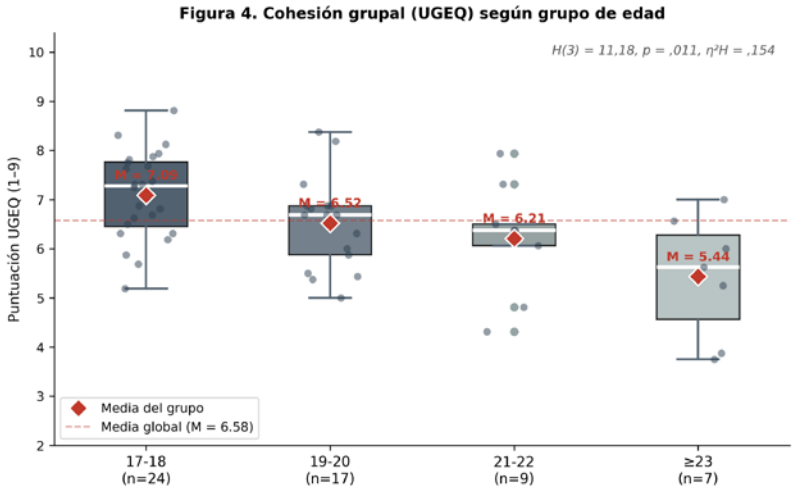
más baja (M = 4,51), ligeramente por encima del punto medio de la escala, señalando el área de mayor vulnerabilidad en el razonamiento crítico del grupo.

Tabla 2. Correlaciones de Spearman entre variables principales y con la edad (N = 57)

Variable	1	2	3	4	5	6. Edad
1. UGEQ (total)	—					−0,47**
2. ATG-T	0,82**	—				−0,34**
3. GI-T	0,73**	0,44**	—			−0,46**
4. UWES-S (total)	0,22	0,34**	0,10	—		0,12
5. Vigor	0,19	0,31*	0,07	0,83**	—	−0,02
6. WGCTA-FS (total)	−0,02	0,01	−0,04	0,01	0,06	−0,10

Nota. * p < 0,05; ** p < 0,01. La columna 6 recoge las correlaciones de Spearman de cada variable con la edad del participante. Los signos negativos se muestran explícitamente.

Las tres escalas principales no mostraron correlaciones significativas entre sí (todos rs < 0,22, p > 0,10), confirmando su independencia como constructos. A nivel de subescalas destaca la asociación positiva entre ATG-T y vigor (rs = 0,31, p = 0,018). La edad correlacionó negativamente con todas las subescalas de cohesión grupal (todas p < 0,05), siendo la relación especialmente intensa con UGEQ total (rs = −0,469, p < 0,001).



Nota. Diagrama de caja con puntos individuales y medias grupales (rombos rojos). Línea discontinua: media global (M = 6,58). H(3) = 11,18, p = ,011, η²H = ,154.

La Figura 4 ilustra el descenso progresivo de la cohesión grupal conforme aumenta la edad: el grupo de 17-18 años presenta la puntuación más alta (M = 7,09, DE = 0,90) frente al grupo de ≥23 años (M = 5,44, DE = 1,25). La prueba de Kruskal-Wallis confirmó diferencias estadísticamente significativas entre grupos de edad (H(3) = 11,18, p = 0,011, η²H = 0,154, efecto mediano). Debe señalarse que la situación de convivencia también se asocia a la edad (familia: Medad = 19,37; piso compartido: Medad = 20,73; U = 136,5, p = 0,040), lo que sugiere que la tendencia hacia menor cohesión en estudiantes en piso compartido puede ser parcialmente atribuible a la edad. Las diferencias

según sexo y situación laboral no alcanzaron significación estadística (todos $p > 0,19$). No obstante, el desequilibrio de la muestra por sexo (93,0% mujeres, frente a 7,0% varones) confiere una potencia estadística muy baja a este contraste, por lo que la ausencia de significación no debe interpretarse como evidencia de ausencia de diferencias.

4. DISCUSIÓN Y CONCLUSIONES

El presente estudio ha ofrecido un diagnóstico basal riguroso de tres competencias transversales en estudiantes de primer curso del Grado en Logopedia, aportando una base empírica diferenciada para el diseño de innovación docente mediante ABR. Los resultados revelan un perfil caracterizado por cohesión orientada a la tarea, engagement funcional con asimetría entre subescalas y pensamiento crítico que combina fortaleza lógico-deductiva con dificultades en el razonamiento epistémico implícito.

La cohesión grupal muestra valores positivos y un perfil consistente con los datos normativos obtenidos por Ruiz-Robledillo et al. (2023) en muestras universitarias españolas. El predominio de las dimensiones de tarea sobre las sociales es esperable al inicio del curso y resulta funcionalmente favorable para el ABR, que exige principalmente coordinación en torno a objetivos académicos compartidos (Bosselut et al., 2018). El fortalecimiento deliberado de la dimensión social requerirá incorporar componentes vivenciales e interacción informal en el diseño de los retos.

El nivel intermedio de engagement, con dedicación marcadamente superior al vigor y la absorción, ofrece una imagen matizada del estudiantado de nuevo ingreso. La alta dedicación refleja un fuerte sentido de propósito vocacional, recurso motivacional inicial valioso. Sin embargo, el vigor y la absorción moderados sugieren que aún no se han consolidado los recursos energéticos y la implicación cognitiva profunda necesarios para un engagement pleno, posiblemente por la incertidumbre propia de la transición universitaria (Bresó et al., 2007; Salanova et al., 2010). Este diagnóstico apoya el diseño de retos que incrementen la relevancia percibida de las tareas, para elevar la dedicación ya presente, y la experiencia de competencia y autonomía, factores clave para activar el vigor y la absorción (Carmona-Halty et al., 2019).

El perfil del WGCTA-FS presenta implicaciones pedagógicas específicas. El dominio elevado en deducciones ($M = 7,28$) indica que el estudiantado maneja bien el razonamiento lógico formal. Sin embargo, el rendimiento más bajo en reconocimiento de suposiciones ($M = 4,51$) apunta a una debilidad en la habilidad de explicitar premisas tácitas, de consecuencias clínicas directas: en la práctica logopédica, muchos errores diagnósticos derivan de asumir supuestos no examinados (Facione, 1990; Profetto-McGrath, 2005). El ABR, al obligar al estudiantado a identificar la idea central de cada reto y a justificar públicamente sus decisiones, proporciona el andamiaje pedagógico adecuado para trabajar esta dimensión (Apple, 2011; Bolaños, 2019).

El hallazgo más novedoso es la correlación negativa de efecto mediano-grande entre edad y cohesión grupal ($r_s = -0,469$, $p < 0,001$), que se mantiene en todas las subescalas del UGEQ. Los estudiantes de mayor edad muestran sistemáticamente menor cohesión percibida, especialmente en las dimensiones social y de tarea. Dado el carácter transversal y correlacional del diseño, esta asociación no permite establecer relaciones explicativas. Cabe plantear, a modo de hipótesis de trabajo, que la experiencia de inserción grupal podría diferir entre el estudiantado más joven y quienes acceden más tardíamente al grado, posibilidad que la modalidad de convivencia no agota y que deberá contrastarse en diseños específicos. La asociación entre ATG-T y vigor ($r_s = 0,31$, $p = 0,018$) complementa este panorama: la implicación energética en los estudios y en las tareas del grupo constituyen expresiones

parcialmente compartidas de un mismo patrón motivacional, lo que apoya el diseño de intervenciones que trabajen ambas de forma integrada.

Este estudio aporta varias contribuciones específicas: 1) En el plano académico, facilita datos de referencia para la escasa literatura sobre diagnóstico basal de competencias transversales en logopedia y ciencias de la salud, donde la evidencia sobre ABR es notablemente menor que en ingenierías (De la Cruz-Velazco et al., 2022). 2) En el plano clínico-formativo, los resultados refuerzan la necesidad de entrenar explícitamente el reconocimiento de supuestos implícitos en los currículos sanitarios, habilidad directamente vinculada a la calidad del razonamiento diagnóstico (Facione, 1990; Profetto-McGrath, 2005). 3) En el plano social, la relación inversa entre edad y cohesión visibiliza una brecha de integración que afecta al estudiantado no tradicional, con responsabilidades laborales o familiares, colectivo cuya presencia en las aulas crece de forma sostenida y cuyas necesidades de inclusión raramente son atendidas por los diseños pedagógicos convencionales.

Las limitaciones del estudio incluyen el tamaño muestral reducido (N = 57), el muestreo por conveniencia, la circunscripción a una titulación y universidad, y el diseño transversal que impide establecer relaciones causales. El tamaño muestral limita la potencia estadística de los contrastes de subgrupos y la precisión de las estimaciones, por lo que los resultados deben considerarse preliminares. Asimismo, la elevada proporción de mujeres (93,0%), característica del Grado en Logopedia, pero especialmente acusada en esta muestra, restringe la generalización de los hallazgos al conjunto del estudiantado e impide un análisis fiable de las diferencias por sexo. Ambas circunstancias exigen prudencia en la extrapolación de los resultados y su replicación en muestras más amplias y equilibradas. Estas limitaciones serán abordadas en el análisis longitudinal pretest-postest, reservado para su publicación en revista de impacto, que evaluará la eficacia del ABR sobre las tres competencias aquí diagnosticadas.

El diagnóstico basal proporciona orientaciones pedagógicas que, dado el diseño transversal, conviene formular como hipótesis de trabajo para la fase de intervención: diseñar retos que fortalezcan el reconocimiento de supuestos implícitos; aprovechar la cohesión orientada a la tarea mientras se estimula la dimensión social; activar el vigor y la absorción mediante tareas retadoras y significativas; e incorporar estrategias específicas de integración grupal para el estudiantado de mayor edad. La solidez de estas decisiones reside en que derivan de los datos reales del grupo, que es precisamente el valor añadido de un diagnóstico basal riguroso como paso previo a cualquier innovación docente. El valor de estas orientaciones reside en que derivan de los datos directos y reales del grupo; su eficacia, no obstante, deberá verificarse empíricamente en la fase de implementación, que es precisamente la aportación de un diagnóstico basal riguroso como paso previo a cualquier innovación docente.

REFERENCIAS

Agüero, M., López, L., y Pérez, J. (2019). Challenge based learning como modelo de aprendizaje profesionalizante: Caso del programa Universidad Europea con Comunica +A. Vivat Academia, (149), 1–24. https://doi.org/10.15178/va.2019.149.1-24

Apple. (2011). Challenge based learning: A classroom guide. http://www.apple.com/br/education/docs/CBL_Classroom_Guide_Jan_2011.pdf

Bolaños, O. (2019). Aprendizaje basado en retos. Centro de Recursos para el Aprendizaje. https://eduteka.icesi.edu.co/pdfdir/crea-rutatic-aprendizaje-basado-en-retos.pdf

Bosselut, G., Heuze, J., Castro, O., Fouquereau, E., y Chevalier, S. (2018). Using exploratory struc-

tural equation modeling to validate a new measure of cohesion in the university classroom setting: The University Group Environment Questionnaire (UGEQ). International Journal of Educational Research, 89, 1–9. https://doi.org/10.1016/j.ijer.2018.03.003

Bresó, E., Salanova, M., Schaufeli, W. B., y Nogareda, C. (2007). NTP 732: Síndrome de estar quemado por el trabajo «Burnout» (III): Instrumento de medición. Instituto Nacional de Seguridad e Higiene en el Trabajo.

Carmona-Halty, M., Salanova, M., Llorens, S., y Schaufeli, W. B. (2019). How psychological capital mediates between study-related positive emotions and academic performance. Journal of Happiness Studies, 20(2), 605–617. https://doi.org/10.1007/s10902-018-9963-5

Carron, A. V. (1982). Cohesiveness in sport groups: Interpretations and considerations. Journal of Sport Psychology, 4(2), 123–138. https://doi.org/10.1123/jsp.4.2.123

De la Cruz-Velazco, P. H., Poquis-Velasquez, E., Valle-Chavez, R. A., Castañeda Sánchez, M. I., y Sánchez Anastacio, K. R. (2022). Aprendizaje basado en retos en la educación superior: Una revisión bibliográfica. Horizontes. Revista de Investigación en Ciencias de la Educación, 6(25), 1409–1421. https://doi.org/10.33996/revistahorizontes.v6i25.422

De Mangione, E., y De Anglat, H. (2007). Evaluación de la competencia crítica a través del test de Watson-Glaser: Exploración de sus cualidades psicométricas. Revista de Psicología, 3(6), 1–14. http://bibliotecadigital.uca.edu.ar/repositorio/revistas/revistapsicologia06.pdf

Ennis, R. H. (1993). Critical thinking assessment. Theory into Practice, 32(3), 179–186. https://doi.org/10.1080/00405849309543594

Facione, P. A. (1990). Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction. The California Academic Press.

Profetto-McGrath, J. (2005). The relationship of critical thinking skills and critical thinking dispositions of baccalaureate nursing students. Journal of Advanced Nursing, 53(3), 306–315. https://doi.org/10.1111/j.1365-2648.2006.03727.x

Ruiz-Robledillo, N., Albaladejo-Blázquez, N., Ferrer-Cascales, R., Rubio-Aparicio, M., Madrid-Valero, J. J., Vela-Bermejo, J., y Bosselut, G. (2023). Spanish adaptation and validation of the UGEQ questionnaire to assess group cohesion in university students. Revista Digital de Investigación en Docencia Universitaria, 17(1), e1618. https://doi.org/10.19083/ridu.2023.1618

Salanova, M., Schaufeli, W. B., Martínez, I., y Bresó, E. (2010). How obstacles and facilitators predict academic performance: The mediating role of study burnout and engagement. Anxiety, Stress & Coping, 23, 53–70. https://doi.org/10.1080/10615800802609965

Schaufeli, W. B., & Bakker, A. B. (2004). Job demands, job resources, and their relationship with burnout and engagement: A multi-sample study. Journal of Organizational Behavior, 25(3), 293–315. https://doi.org/10.1002/job.248

Schaufeli, W. B., Salanova, M., González-Romá, V., y Bakker, A. B. (2002). The measurement of engagement and burnout: A two-sample confirmatory factor analytic approach. Journal of Happiness Studies, 3, 71–92. https://doi.org/10.1023/A:1015630930326

Watson, G., y Glaser, E. M. (1994). Watson-Glaser Critical Thinking Appraisal, Short Form manual. Psychological Corporation.

3. Aprendizaje adaptativo con inteligencia artificial para reforzar la enseñanza de Anatomía Humana en Ciencias de la Salud

Eva Ausó[1], Gema Esquiva Sobrino[1], Ana Cifuentes[1], Damián R. Martínez St. John[1], J. Víctor García Velasco[1], Javier Sanz Reig[1], María de las Nieves Pérez Climent[1], Erika Meléndez-Oliva[1]*

[1]Departamento de Óptica, Farmacología y Anatomía, Facultad de Ciencias, Universidad de Alicante, Alicante, España
**Autora de correspondencia: Erika Meléndez-Oliva*

RESUMEN

La enseñanza de la Anatomía Humana en titulaciones de Ciencias de la Salud implica una elevada carga conceptual, terminológica y visual que dificulta el aprendizaje autónomo del alumnado de primer curso. El objetivo de este trabajo fue evaluar la implementación de una plataforma digital basada en inteligencia artificial y aprendizaje adaptativo como recurso complementario para el repaso y la consolidación de contenidos anatómicos. Se desarrolló una experiencia de innovación docente con diseño cuasi-experimental y alcance descriptivo en tres asignaturas de Anatomía de una universidad pública española. La cohorte de implantación estuvo formada por 417 estudiantes de Medicina, Enfermería y Nutrición Humana y Dietética. La evaluación combinó una encuesta de satisfacción respondida por 322 estudiantes y la comparación agregada de resultados académicos entre los cursos 2024-2025 y 2025-2026. El uso declarado de la plataforma fue del 79 % en Enfermería y Medicina, y del 47,9 % en Nutrición. La valoración global fue mayoritariamente positiva, especialmente en utilidad para repasar, preparar exámenes y organizar el estudio. Los resultados académicos mostraron mejoras en Nutrición y Enfermería, mientras que en Medicina fueron heterogéneos. La experiencia sugiere que el aprendizaje adaptativo puede reforzar la enseñanza anatómica, siempre que se acompañe de integración docente y mejora técnica.

PALABRAS CLAVE: inteligencia artificial; aprendizaje adaptativo; Anatomía Humana; innovación docente; educación superior.

1. INTRODUCCIÓN

La Anatomía Humana constituye una disciplina básica en la formación del alumnado de Ciencias de la Salud. Su aprendizaje resulta imprescindible para comprender la organización estructural del cuerpo humano, interpretar procesos fisiológicos y patológicos, y desarrollar competencias profesionales vinculadas a la práctica clínica, asistencial y sanitaria. Sin embargo, se trata de una materia compleja, caracterizada por una elevada densidad terminológica, una importante carga memorística y la necesidad de integrar relaciones espaciales entre estructuras anatómicas. Esta combinación de elementos convierte el aprendizaje anatómico en un reto especialmente relevante durante el primer curso universitario.

La dificultad de la asignatura no se limita a la memorización de nombres de estructuras. El alumnado debe reconocer relaciones topográficas, interpretar planos corporales, comprender la organización de sistemas orgánicos y aplicar el conocimiento anatómico a preguntas de razonamiento o situaciones pro-

fesionales. Además, los estudiantes acceden a la universidad con niveles de preparación previa, ritmos de estudio y estrategias de aprendizaje muy diferentes. Esta heterogeneidad dificulta que los modelos docentes tradicionales respondan de forma personalizada a todas las necesidades de aprendizaje.

En este contexto, la docencia universitaria actual requiere recursos flexibles, accesibles y compatibles con el estudio autónomo. Las clases teóricas, las prácticas y los seminarios continúan siendo esenciales, pero pueden complementarse con herramientas digitales que permitan practicar de forma repetida, recibir retroalimentación, detectar errores y organizar el repaso antes de las pruebas de evaluación. La literatura sobre *e-learning* en educación médica ha señalado que las tecnologías digitales pueden ampliar las oportunidades de práctica y favorecer entornos de aprendizaje más activos (Ellaway & Masters, 2008; Ruiz et al., 2006).

La inteligencia artificial aplicada a la educación superior permite avanzar hacia modelos de aprendizaje más personalizados. Las plataformas basadas en IA pueden adaptar itinerarios, secuenciar actividades, identificar patrones de error y ofrecer recursos ajustados al progreso del estudiante. Las revisiones recientes sobre IA en educación superior destacan el potencial de los sistemas adaptativos y de personalización, pero también subrayan la necesidad de evaluar su impacto con rigor, atender a las implicaciones éticas y situar la tecnología dentro de diseños pedagógicos claros (Bond et al., 2024; Zawacki-Richter et al., 2019).

El aprendizaje adaptativo resulta especialmente pertinente en asignaturas anatómicas, donde el alumnado puede necesitar refuerzos distintos: terminología, reconocimiento de estructuras, relaciones espaciales o aplicación de contenidos. En ciencias de la salud, la literatura sobre tecnologías adaptativas muestra que su potencial depende de la calidad del diseño instruccional, de la integración temporal de las actividades y del grado de agencia que conserva el estudiante (Andersen et al., 2022; Cavanagh et al., 2020). En Anatomía Humana, además, la incorporación de recursos digitales debe entenderse como una estrategia complementaria, no como una simple sustitución de la experiencia presencial o práctica (Wickramasinghe et al., 2022).

La plataforma Khipulearn se incorporó como recurso digital de apoyo al aprendizaje anatómico. Se trata de un entorno basado en inteligencia artificial y aprendizaje adaptativo que permite organizar contenidos por competencias, crear actividades interactivas, establecer umbrales de progreso y realizar seguimiento del avance del alumnado. Su utilización se planteó como una estrategia complementaria para reforzar el estudio autónomo, favorecer el repaso progresivo y facilitar la identificación de lagunas de conocimiento.

El objetivo de este trabajo fue evaluar la implementación de Khipulearn como herramienta complementaria de apoyo al aprendizaje de la Anatomía Humana en tres titulaciones de Ciencias de la Salud. De forma específica, se analizó el uso declarado de la plataforma, la satisfacción del alumnado, la utilidad percibida para el aprendizaje y la preparación de la asignatura, las propuestas de mejora y la evolución descriptiva de los resultados académicos respecto al curso previo a la implantación.

2. MÉTODO

Descripción del contexto y de los participantes

Se desarrolló una experiencia de innovación docente con diseño cuasi-experimental, prospectivo y alcance descriptivo en tres asignaturas de Anatomía Humana de una universidad pública española. La experiencia se enmarcó en un proyecto institucional de innovación docente orientado a incorporar herramientas digitales basadas en inteligencia artificial y aprendizaje adaptativo en asignaturas biomédicas de primer curso.

La cohorte de implantación estuvo constituida por 417 estudiantes matriculados en tres titulaciones de Ciencias de la Salud: 222 estudiantes del Grado en Enfermería, 117 del Grado en Nutrición Humana y Dietética y 78 del Grado en Medicina. Las asignaturas implicadas fueron Anatomía Humana, Anatomía y Anatomía General y Embriología Humana, todas ellas con contenidos anatómicos básicos y elevada carga conceptual.

La participación del alumnado en el uso de la plataforma fue voluntaria. La encuesta de satisfacción fue respondida por 322 estudiantes, lo que representa el 77,2 % de la cohorte total. De estas respuestas, 186 correspondieron a Enfermería, 74 a Nutrición Humana y Dietética y 62 a Medicina. El cuestionario incluyó consentimiento informado inicial y el análisis se realizó de forma agregada, sin identificación individual.

Tabla 1. Participantes incluidos en la implantación y en la encuesta de satisfacción

Titulación	Asignatura	Estudiantes matriculados	Respuestas encuesta	Uso declarado
Enfermería	Anatomía Humana	222	186	147 (79 %)
Nutrición Humana y Dietética	Anatomía	117	74 respuestas; 73 válidas	35 (47,9 %)
Medicina	Anatomía General y Embriología Humana	78	62	49 (79 %)
Total	Tres asignaturas de Anatomía	417	322	231 usuarios

Fuente: elaboración propia.

INSTRUMENTOS

El instrumento principal fue la plataforma Khipulearn, configurada como recurso complementario de estudio y repaso. El entorno permitió estructurar contenidos por competencias, generar actividades interactivas, organizar preguntas de autoevaluación y facilitar el seguimiento del progreso del estudiante. Los materiales fueron diseñados y revisados por profesorado del área de Anatomía para garantizar su adecuación a los contenidos y resultados de aprendizaje de cada asignatura.

Para evaluar la experiencia se elaboró una encuesta de satisfacción dirigida al alumnado. El cuestionario recogió información sobre datos sociodemográficos, uso de la plataforma, tiempo aproximado de utilización, momento de uso, utilidad percibida, diseño y funcionamiento, impacto en el aprendizaje, valoración global y sugerencias de mejora. Los ítems de valoración se organizaron mediante escalas tipo Likert y preguntas cerradas, complementadas con respuestas abiertas para recoger propuestas cualitativas.

Además, para explorar la posible relación entre la implantación de la plataforma y el rendimiento académico, se analizaron de forma agregada las actas de calificación de las asignaturas implicadas. Se compararon las categorías de calificación del curso 2024-2025, previo a la incorporación de la plataforma, con las del curso 2025-2026, en el que Khipulearn estuvo disponible como recurso complementario.

PROCEDIMIENTO

El procedimiento se desarrolló en cinco fases. En primer lugar, entre noviembre de 2024 y julio de 2025, el profesorado participante recibió formación en el uso de Khipulearn y se familiarizó con sus posibilidades docentes. Esta fase permitió definir la estructura general de la plataforma, las competencias y la organización de los contenidos.

En segundo lugar, se diseñaron los materiales docentes. El equipo elaboró preguntas, casos y actividades interactivas adaptadas a los contenidos de cada asignatura. La finalidad fue crear un banco de recursos que permitiera al alumnado practicar de forma autónoma, identificar errores y reforzar los bloques de contenido trabajados en clase.

En tercer lugar, durante el primer cuatrimestre del curso 2025-2026, Khipulearn se implementó como recurso complementario en las tres asignaturas. El alumnado pudo acceder a la plataforma para repasar contenidos, preparar exámenes y realizar actividades de autoevaluación. El uso no sustituyó a la docencia presencial ni modificó los sistemas ordinarios de evaluación.

En cuarto lugar, tras la finalización del periodo de uso, se administró la encuesta de satisfacción y se analizaron los datos de forma descriptiva. Finalmente, se compararon las calificaciones agregadas de los cursos 2024-2025 y 2025-2026. Esta comparación debe interpretarse como exploratoria, ya que no permite atribuir causalidad directa a la plataforma y puede estar influida por características de cohorte, diferencias en pruebas de evaluación y otros recursos docentes disponibles.

3. RESULTADOS

La plataforma fue utilizada como herramienta de repaso por 231 estudiantes de los 322 que respondieron a la encuesta. El uso declarado fue más elevado en Enfermería y Medicina, donde el 79 % del alumnado encuestado indicó haber utilizado Khipulearn. En Nutrición Humana y Dietética, el porcentaje de uso fue del 47,9 %. Esta diferencia sugiere que la comunicación inicial, la integración de la herramienta en la dinámica de la asignatura y la percepción de necesidad pueden influir en la adopción real de la plataforma.

Tabla 2. Uso, utilidad percibida y recomendación de Khipulearn por titulación

Titulación	Uso declarado	Útil para repasar	Contribuye al aprendizaje	Recomendaría su uso
Enfermería	147/186 (79 %)	84,4 %	73,5 %	75,5 %
Medicina	49/62 (79 %)	71,4 %	65,3 %	71,4 %
Nutrición Humana y Dietética	35/73 (47,9 %)	62,9 %	60,0 %	60,0 %

Fuente: elaboración propia.

Entre quienes utilizaron Khipulearn, la valoración global fue mayoritariamente positiva, aunque con diferencias entre titulaciones. En Enfermería, 97 de 147 estudiantes usuarios valoraron la experiencia como satisfactoria o muy satisfactoria. En Medicina, esta valoración fue expresada por 30 de 49 estudiantes usuarios, y en Nutrición Humana y Dietética por 18 de 35. Los aspectos mejor valorados fueron la utilidad para repasar contenidos anatómicos, consolidar conceptos y preparar las pruebas de evaluación.

La utilidad percibida fue especialmente elevada en Enfermería, donde el 84,4 % del alumnado usuario consideró que la plataforma resultó útil para repasar contenidos, el 73,5 % indicó que contribuyó positivamente a su aprendizaje y el 75,5 % recomendaría su uso a otros estudiantes. En Medicina, el 71,4 % la consideró útil para repasar, el 65,3 % señaló que contribuyó positivamente a su aprendizaje y el 71,4 % la recomendaría. En Nutrición Humana y Dietética, estos porcentajes fueron del 62,9 %, 60,0 % y 60,0 %, respectivamente.

El análisis comparativo de las actas académicas mostró resultados heterogéneos entre titulaciones. En Nutrición Humana y Dietética se observó una mejora relevante: el porcentaje de suspensos descendió del 38,7 % en 2024-2025 al 15,6 % en 2025-2026, mientras que las calificaciones de notable y matrícula de honor aumentaron del 15,1 % al 38,5 %. En Enfermería también se identificó una reducción de suspensos, del 64,5 % al 57,1 %, acompañada de un ligero incremento de calificaciones positivas.

En Medicina, los resultados académicos fueron menos homogéneos. Aunque se mantuvo una proporción elevada de calificaciones de notable y matrícula de honor, el porcentaje de suspensos aumentó del 28,7 % al 40,0 %. Por tanto, no puede afirmarse una mejora global del rendimiento académico en esta titulación. Este resultado refuerza la necesidad de interpretar la comparación entre cohortes con cautela y de considerar factores contextuales, como la integración de la plataforma, el tipo de evaluación, el tiempo efectivo de uso y las características del grupo.

Tabla 3. Comparación agregada del rendimiento académico antes y después de la implantación

Titulación	Curso	Suspensos	Notable + matrícula de honor	Interpretación descriptiva
Nutrición Humana y Dietética	2024-2025	41/106 (38,7 %)	16/106 (15,1 %)	Curso previo
Nutrición Humana y Dietética	2025-2026	17/109 (15,6 %)	42/109 (38,5 %)	Mejora marcada
Enfermería	2024-2025	142/220 (64,5 %)	35/220 (15,9 %)	Curso previo
Enfermería	2025-2026	125/219 (57,1 %)	43/219 (19,6 %)	Mejora moderada
Medicina	2024-2025	29/101 (28,7 %)	43/101 (42,6 %)	Curso previo
Medicina	2025-2026	42/105 (40,0 %)	44/105 (41,9 %)	Resultados heterogéneos

Fuente: elaboración propia.

Las respuestas abiertas permitieron identificar fortalezas y áreas de mejora. El alumnado destacó la utilidad de Khipulearn para repasar, practicar preguntas tipo test, orientar el estudio y consolidar contenidos. Sin embargo, también se señalaron limitaciones técnicas recurrentes: lentitud en la carga, bloqueos, repetición de preguntas, dificultad para rehacer cuestionarios y ausencia de retroalimentación suficiente tras las respuestas incorrectas. La mejora técnica y pedagógica de la retroalimentación fue una de las demandas más reiteradas.

Las áreas de mejora seleccionadas con mayor frecuencia fueron el funcionamiento técnico, la velocidad de carga y la estabilidad de la plataforma. En Enfermería, el 73,5 % del alumnado usuario señaló la necesidad de mejorar el funcionamiento técnico y el 72,1 % la velocidad de carga. En

Medicina, estos porcentajes fueron del 67,3 % y 65,3 %, respectivamente; y en Nutrición Humana y Dietética, del 57,1 % y 54,3 %. Estos datos muestran que la aceptación pedagógica de la herramienta estuvo condicionada por la experiencia técnica de uso.

Tabla 4. Fortalezas y áreas de mejora identificadas por el alumnado

Dimensión	Fortalezas identificadas	Áreas de mejora
Utilidad para el aprendizaje	Repaso de contenidos, consolidación de conceptos y preparación de exámenes	Incrementar la retroalimentación tras errores y vincular respuestas con explicaciones
Diseño de actividades	Preguntas útiles para orientar el estudio y practicar contenidos anatómicos	Permitir repetir cuestionarios y evitar la repetición automática de preguntas
Funcionamiento técnico	Disponibilidad de un entorno digital complementario	Mejorar velocidad de carga, estabilidad, acceso y navegación
Integración docente	Recurso compatible con el estudio autónomo y la docencia presencial	Presentar la herramienta desde el inicio y reforzar su uso durante el curso

Fuente: elaboración propia.

A partir de los resultados obtenidos se definieron varias líneas de continuidad orientadas a consolidar la experiencia. Estas propuestas se centran en mejorar la integración docente de la plataforma, optimizar su rendimiento técnico, ampliar la retroalimentación formativa y diseñar evaluaciones futuras que permitan analizar con mayor precisión la relación entre uso individual y resultados de aprendizaje. La finalidad no es únicamente mantener el recurso, sino transformarlo en un componente estable y pedagógicamente integrado dentro de las asignaturas de Anatomía Humana.

Tabla 5. Propuestas de continuidad derivadas de la experiencia

Línea de continuidad	Necesidad identificada	Acción propuesta
Integración docente	Uso desigual entre titulaciones y necesidad de mayor visibilidad	Presentar la plataforma desde el inicio y vincular actividades con cada bloque temático
Mejora técnica	Lentitud, bloqueos y repetición de preguntas	Optimizar velocidad, estabilidad y accesibilidad desde distintos dispositivos
Retroalimentación	Demanda de visualizar respuestas correctas y comprender errores	Incorporar explicaciones breves y revisión de respuestas falladas
Evaluación futura	Necesidad de analizar impacto con mayor precisión	Relacionar uso individual, tiempo de conexión y rendimiento académico
Transferencia	Potencial aplicación en otras materias biomédicas	Adaptar el modelo a otras asignaturas y compartir la experiencia con otros equipos docentes

Fuente: elaboración propia.

4. DISCUSIÓN Y CONCLUSIONES

El objetivo de este trabajo fue evaluar la implementación de una plataforma de aprendizaje adaptativo basada en inteligencia artificial como recurso complementario en asignaturas de Anatomía Humana. Los resultados muestran una aceptación global positiva por parte del alumnado usuario, especialmente en relación con el repaso de contenidos, la preparación de exámenes y la organización del estudio. Estos hallazgos son coherentes con la literatura que describe el *e-learning* como un complemento útil para ampliar oportunidades de práctica y autoevaluación en educación médica y sanitaria (Ellaway & Masters, 2008; Ruiz et al., 2006).

La experiencia confirma que la tecnología puede aportar valor cuando se vincula a una necesidad docente concreta. En este caso, la necesidad era reforzar el aprendizaje autónomo en una materia con elevada carga conceptual, terminológica y visual. La plataforma permitió organizar contenidos por competencias y ofrecer actividades de repaso, lo que se alinea con los enfoques de aprendizaje adaptativo orientados a personalizar la progresión del estudiante (Andersen et al., 2022; Cavanagh et al., 2020).

No obstante, los resultados también muestran que la disponibilidad de una herramienta digital no garantiza por sí misma una mejora uniforme del rendimiento académico. La evolución de las calificaciones fue favorable en Nutrición Humana y Dietética y moderadamente positiva en Enfermería, pero heterogénea en Medicina. Esta diferencia puede estar relacionada con el grado de integración de la herramienta en la asignatura, el perfil del alumnado, la dificultad de las pruebas o la forma en que cada cohorte utilizó el recurso. Por ello, la comparación de actas debe interpretarse como un indicador descriptivo y no como evidencia causal.

Una aportación relevante del estudio es la identificación de condiciones necesarias para que una plataforma adaptativa sea percibida como útil. El alumnado valoró positivamente el concepto pedagógico de la herramienta, pero las incidencias técnicas limitaron su experiencia. Este hallazgo coincide con la literatura sobre *digital anatomy*, que advierte que las herramientas digitales no deben entenderse como una simple conversión tecnológica, sino como recursos que requieren infraestructura, diseño pedagógico e integración adecuada (Wickramasinghe et al., 2022).

Desde el punto de vista docente, los resultados sugieren varias implicaciones. En primer lugar, la plataforma debe presentarse desde el inicio del curso y vincularse de forma explícita con los bloques temáticos impartidos. En segundo lugar, conviene orientar al alumnado sobre cuándo y cómo utilizarla, evitando que su uso se concentre únicamente en los días previos al examen. En tercer lugar, las actividades deben revisarse de forma continua para asegurar su alineación con los resultados de aprendizaje y con el nivel de profundidad exigido en cada asignatura.

La retroalimentación emerge como una dimensión clave de mejora. Muchos estudiantes solicitaron visualizar la respuesta correcta, conocer por qué una opción era incorrecta y poder repetir actividades ya completadas. Incorporar explicaciones breves tras los errores permitiría transformar la plataforma en un recurso más formativo, especialmente en Anatomía Humana, donde las confusiones terminológicas y topográficas pueden mantenerse si el estudiante no identifica con precisión el origen del fallo.

El estudio presenta limitaciones. En primer lugar, el uso de Khipulearn fue voluntario, por lo que no se dispone de una exposición homogénea entre estudiantes. En segundo lugar, la comparación académica se realizó entre cohortes de cursos diferentes, sin control de variables como nivel previo, tiempo real de uso o dificultad de las pruebas. En tercer lugar, los datos de satisfacción proceden de autoinforme y pueden estar condicionados por la experiencia técnica de la plataforma. Finalmente, el análisis se centró en resultados agregados, por lo que no fue posible relacionar de forma individual el uso de la herramienta con el rendimiento final.

A pesar de estas limitaciones, la experiencia aporta evidencias prácticas sobre la viabilidad de incorporar aprendizaje adaptativo en asignaturas anatómicas. La combinación de datos cuantitativos, respuestas cualitativas y comparación académica descriptiva permite identificar tanto el potencial de la plataforma como los aspectos que deben mejorarse antes de una implantación más amplia. Futuros estudios deberían analizar el tiempo efectivo de conexión, la progresión por competencias y la relación individual entre uso, satisfacción y rendimiento.

En conclusión, Khipulearn fue percibida como una herramienta útil para apoyar el estudio autónomo de la Anatomía Humana en titulaciones de Ciencias de la Salud. Su implementación mostró resultados positivos en términos de satisfacción y utilidad percibida, y se asoció a mejoras académicas en algunas titulaciones, aunque no de forma uniforme. La experiencia sugiere que las plataformas de inteligencia artificial y aprendizaje adaptativo pueden reforzar la enseñanza anatómica siempre que se integren pedagógicamente, funcionen de manera estable y ofrezcan retroalimentación suficiente al alumnado.

Financiación: este trabajo se deriva de un proyecto de innovación docente financiado en el marco del Programa de Redes de Investigación en Docencia Universitaria de la Universidad de Alicante, red 6288.

Conflicto de intereses: las autoras y los autores declaran no tener conflictos de intereses relacionados con este trabajo.

5. REFERENCIAS

Andersen, B. L., Jørnø, R. L., & Nortvig, A.-M. (2022). Blending adaptive learning technology into nursing education: A scoping review. Contemporary Educational Technology, 14(1), Article ep333. https://doi.org/10.30935/cedtech/11370

Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., Pham, P., Chong, S. W., & Siemens, G. (2024). A meta systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigour. International Journal of Educational Technology in Higher Education, 21, Article 4. https://doi.org/10.1186/s41239-023-00436-z

Cavanagh, T., Chen, B., Lahcen, R. A. M., & Paradiso, J. R. (2020). Constructing a design framework and pedagogical approach for adaptive learning in higher education: A practitioner's perspective. International Review of Research in Open and Distributed Learning, 21(1), 172-196. https://doi.org/10.19173/irrodl.v21i1.4557

Ellaway, R. H., & Masters, K. (2008). AMEE Guide 32: E-learning in medical education. Part 1: Learning, teaching and assessment. Medical Teacher, 30(5), 455-473. https://doi.org/10.1080/01421590802108331

Ruiz, J. G., Mintzer, M. J., & Leipzig, R. M. (2006). The impact of e-learning in medical education. Academic Medicine, 81(3), 207-212. https://doi.org/10.1097/00001888-200603000-00002

Wickramasinghe, N., Thompson, B. R., & Xiao, J. (2022). The opportunities and challenges of digital anatomy for medical sciences: Narrative review. JMIR Medical Education, 8(2), e34687. https://doi.org/10.2196/34687

Zawacki-Richter, O., Marín, V. I., Bond, M., & 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, Article 39. https://doi.org/10.1186/s41239-019-0171-0

4. De la rigurosidad escrita a la síntesis visual: el papel del informe y el póster en ingeniería ante la IA generativa

Fidel Aznar Gregori[1], Ramón Rizo Aldeguer[1], Pilar Arques Corrales[1], Mar Pujol López[1], Patricia Compañ Rosique[1], Javier Botana Gómez[3], Miguel Angel Lozano Ortega[1], Francisco José Mora Lizán[1], Juan Antonio Puchol García[1], Mª José Pujol López[2]

[1]*Dpto. de Ciencia de la Computación e Inteligencia Artificial, Universidad de Alicante*
[2]*Dpto. de Matemática Aplicada, Universidad de Alicante*
[3]*Servicio de informática - Área de desarrollo de aplicaciones, Universidad de Alicante*

RESUMEN

La comunicación científica constituye una competencia transversal esencial en la formación de ingenieros. Este estudio analiza comparativamente dos formatos de divulgación, el informe técnico y el póster científico, implementados en la asignatura *Sistemas Inteligentes* del Grado en Ingeniería Informática de la Universidad de Alicante. A partir de una muestra de 114 estudiantes que participaron en un mini-congreso sobre proyectos de Inteligencia Artificial, se evaluó, mediante un cuestionario mixto (escalas Likert 1-5 y preguntas abiertas), la percepción del alumnado sobre la contribución de cada formato al desarrollo de competencias comunicativas. Los resultados revelan una clara complementariedad: el informe fue percibido como el formato que más profundidad técnica exige (81,6 % de acuerdo, M = 3,98/5), mientras que el póster destacó como la herramienta que mejor entrena la capacidad de síntesis y la adaptación del discurso a audiencias no especializadas (64,0 %, M = 3,76). El 66,7 % elegiría el informe como único medio para acreditar formalmente su aprendizaje, si bien el 53,5 % lo consideró la tarea más desafiante. Se analiza asimismo el papel de la inteligencia artificial generativa, empleada para tareas mecánicas (53,5 %) pero sometida a verificación rigurosa por el 70,2 % de los encuestados. En conjunto, los resultados sugieren que la integración de ambos formatos puede constituir una estrategia didáctica sinérgica y replicable para la formación integral en ingeniería.

PALABRAS CLAVE: comunicación científica, póster científico, informe técnico, docencia en ingeniería, inteligencia artificial generativa.

1. INTRODUCCIÓN

La capacidad para comunicar resultados científicos y técnicos de forma eficaz constituye una de las competencias transversales más demandadas en el perfil profesional del ingeniero contemporáneo (García-Peñalvo et al., 2020). Sin embargo, la formación universitaria en ingeniería ha priorizado tradicionalmente la adquisición de conocimientos técnicos, relegando el desarrollo de habilidades comunicativas a un segundo plano (Jonassen et al., 2006). Esta brecha entre la formación recibida y las demandas del entorno profesional se agrava en un contexto donde la complejidad de los proyectos tecnológicos exige ingenieros capaces de adaptar su discurso a audiencias diversas, desde comités técnicos hasta públicos no especializados.

En el ámbito de la enseñanza de la Ingeniería Informática, esta problemática adquiere particular relevancia. La asignatura Sistemas Inteligentes, perteneciente al Grado en Ingeniería Informática de la Universidad de Alicante, representa para más de 200 estudiantes anuales su primera inmersión

formal en los fundamentos de la Inteligencia Artificial (IA). Si bien la materia aborda contenidos técnicos avanzados (desde MÉTODOs de búsqueda y representación del conocimiento hasta aprendizaje automático y visión artificial), el formato tradicional de evaluación, basado en memorias escritas y exámenes, no resultaba suficiente para desarrollar las competencias comunicativas que el plan de estudios establece como resultado de aprendizaje explícito (Autor, 2025a).

La literatura reciente sobre innovación docente en ingeniería ha señalado el potencial de los formatos de congreso como metodología activa para el desarrollo de competencias transversales. La simulación de eventos científicos, incluyendo la elaboración de informes, pósteres y presentaciones orales, permite al alumnado experimentar de forma auténtica los procesos de comunicación académica y profesional (Walkington et al., 2017). Paralelamente, la irrupción de la inteligencia artificial generativa (IAG) en el ecosistema educativo ha introducido nuevas variables que merecen un análisis riguroso. Herramientas como los grandes modelos de lenguaje (LLM) ofrecen un potencial significativo para apoyar la redacción y síntesis de contenidos, pero también plantean interrogantes sobre la profundidad real del aprendizaje y la necesidad de fomentar un uso crítico y ético de estas tecnologías (Crawford et al., 2023; UNESCO, 2023).

En este contexto, el presente estudio aborda dos formatos fundamentales de comunicación científica, el informe técnico escrito y el póster científico, evaluando su contribución diferencial al desarrollo de competencias en el alumnado de ingeniería. Los objetivos específicos de la investigación son: a) analizar cómo cada formato contribuye al desarrollo de la capacidad de síntesis, la profundización técnica y el pensamiento crítico del alumnado; b) evaluar el papel de la inteligencia artificial generativa como herramienta de asistencia durante el proceso de elaboración de ambos formatos; y c) examinar la utilidad percibida de las rúbricas de evaluación y de la revisión por pares como instrumentos de apoyo al aprendizaje.

2. MÉTODO

2.1. Descripción del contexto y de los participantes

La experiencia se desarrolló en la asignatura *Sistemas Inteligentes*, una materia obligatoria de tercer curso del Grado en Ingeniería Informática de la Universidad de Alicante, cursada por más de 200 estudiantes en cada edición. Como innovación metodológica, se organizó un mini-congreso estudiantil sobre proyectos de Inteligencia Artificial, en el cual el alumnado debía asumir el rol de investigadores, elaborando tanto un informe técnico escrito como un póster científico para comunicar los resultados de su trabajo. La experiencia incluyó además una sesión formativa sobre diseño de pósteres científicos, dinámicas de evaluación por pares (*peer review*) y la presentación pública de los pósteres en una sesión plenaria.

La muestra del estudio estuvo compuesta por 114 estudiantes que respondieron voluntariamente al cuestionario de evaluación de la experiencia. Todos los participantes habían completado ambas tareas (informe escrito y póster científico) y participado en las actividades de evaluación por pares asociadas. La participación en la encuesta fue anónima y no condicionó la calificación de la asignatura.

2.2. Instrumentos

Se diseñó un cuestionario mixto estructurado en tres bloques. El primer bloque (ítems 1A-4B) evaluó, mediante escalas Likert de 5 puntos (1 = “Totalmente en desacuerdo”, 5 = “Totalmente de acuerdo”),

la percepción del alumnado sobre la contribución de cada formato —informe y póster— a cuatro dimensiones: profundización en detalles técnicos, identificación del mensaje clave, organización lógica de la investigación y mejora de la competencia comunicativa ante distintos tipos de audiencia. El segundo bloque (ítems 5-7) valoró la utilidad de los instrumentos de apoyo: la sesión formativa sobre pósteres, la actividad de revisión por pares y la claridad de las rúbricas de evaluación. El tercer bloque (ítems IA1-IA6) exploró el uso de la inteligencia artificial generativa durante la elaboración de ambos formatos, incluyendo su papel en tareas de redacción, síntesis, corrección de estilo y diseño visual, así como la necesidad percibida de verificar los contenidos generados.

Complementariamente, el cuestionario incluyó preguntas abiertas sobre tres aspectos: los contenidos sacrificados al trasladar el informe al formato póster (ítem 11), los aprendizajes específicos atribuibles al diseño del póster (ítem 12) y la percepción sobre el ahorro de tiempo proporcionado por la IA generativa en cada formato. Estas respuestas se analizaron mediante un proceso de codificación temática, identificando patrones recurrentes en las narrativas del alumnado. El cuestionario se elaboró específicamente para este estudio por el equipo docente. Los ítems de las escalas se redactaron a partir de los objetivos de la investigación y de las dimensiones competenciales que la experiencia pretendía activar (profundización técnica, identificación del mensaje clave, organización lógica, comunicación a distintas audiencias, utilidad de los instrumentos de apoyo y uso de la IA generativa), y fueron revisados y consensuados por el equipo docente antes de su administración, con el fin de garantizar la claridad, la pertinencia y la ausencia de ambigüedad de cada enunciado.

2.3. Procedimiento

La intervención se desarrolló en cinco fases a lo largo del cuatrimestre. En la primera fase, los estudiantes recibieron una sesión formativa sobre comunicación científica y diseño de pósteres, orientada a modificar la mentalidad de "copiar y pegar texto" hacia un enfoque de diseño visual y síntesis gráfica. En la segunda fase, cada grupo elaboró un informe técnico escrito (extensión máxima de 15 páginas) documentando el desarrollo de su proyecto de IA, incluyendo fundamentación teórica, metodología, experimentación y resultados. En la tercera fase, cada grupo diseñó un póster científico a partir del contenido del informe, enfrentándose al reto cognitivo de seleccionar, jerarquizar y representar visualmente la información esencial.

En la cuarta fase, se implementó una dinámica de revisión por pares mediante los talleres de Moodle: el alumnado evaluó los pósteres y informes de otros compañeros utilizando las rúbricas proporcionadas por el profesorado, identificando fortalezas y áreas de mejora. La herramienta de talleres de Moodle permitió gestionar de forma automatizada la asignación aleatoria de revisores, la recogida de evaluaciones y el cálculo de calificaciones, garantizando el anonimato entre evaluadores y evaluados. Finalmente, en la quinta fase, los estudiantes presentaron sus pósteres en una sesión plenaria del minicongreso, simulando el formato de una comunicación científica real. El cuestionario de evaluación se administró de forma digital tras la finalización de todas las actividades, garantizando el anonimato de las respuestas.

Los datos cuantitativos procedentes de las escalas Likert se sometieron a un análisis estadístico descriptivo. Para cada ítem se calculó la media (M) y el porcentaje de acuerdo, definido como la proporción de respuestas situadas en los valores 4 ("De acuerdo") y 5 ("Totalmente de acuerdo") de la escala de cinco puntos. Con el fin de explorar las relaciones entre las dimensiones evaluadas, se calculó la matriz de correlaciones de Pearson (r) entre la totalidad de los ítems del cuestionario,

interpretándose la magnitud de los coeficientes según criterios convencionales: correlación débil ($r < 0,30$), moderada ($0,30 \leq r < 0,50$) y fuerte ($r \geq 0,50$). Los análisis se efectuaron con Python 3.12, utilizando el módulo estándar csv para el procesamiento de los datos, NumPy 2.2 para el cálculo de la matriz de correlaciones de Pearson y Matplotlib 3.9 para la generación de las figuras.

Las respuestas a las preguntas abiertas (ítems 11 y 12 y la valoración del ahorro de tiempo) se examinaron mediante un análisis de contenido de carácter inductivo. Tras una lectura preliminar del corpus completo, se identificaron categorías temáticas emergentes; a continuación, cada respuesta se codificó asignándola a una o varias categorías, y se cuantificó la frecuencia relativa de cada categoría sobre el total de la muestra ($N = 114$).

La fiabilidad del cuestionario se estimó mediante el coeficiente alfa de Cronbach sobre las 114 respuestas completas a los 17 ítems de escala. La consistencia interna del cuestionario fue adecuada ($\alpha = 0,81$). Los bloques correspondientes a percepción de los formatos, instrumentos de apoyo y uso de la IA generativa obtuvieron coeficientes α de 0,82, 0,57 y 0,82, respectivamente; el valor más moderado del bloque de apoyo resulta esperable, pues agrupa tres recursos didácticos heterogéneos (sesión formativa, revisión por pares y rúbricas) que no configuran una escala unidimensional.

3. RESULTADOS

Los resultados se presentan organizados en tres ejes: la contribución diferencial de cada formato a las competencias comunicativas, la valoración de los instrumentos de apoyo al aprendizaje y el papel transversal de la inteligencia artificial generativa en el proceso de elaboración.

Tabla 1. Percepción del alumnado sobre la contribución de cada formato a las competencias comunicativas ($N = 114$). Escala Likert de 1 (Totalmente en desacuerdo) a 5 (Totalmente de acuerdo). % Acuerdo = respuestas 4 o 5.

Dimensión evaluada	Informe (M)	Póster (M)	% Acuerdo Inf.	% Acuerdo Pós.
Profundización en detalles técnicos	3,98	3,32	81,6 %	41,2 %
Identificación del mensaje clave	4,07	3,96	79,8 %	76,3 %
Organización de la estructura lógica	4,04	3,60	79,8 %	57,9 %
Comunicación a audiencia técnica	3,61	—	58,8 %	—
Comunicación a audiencia general	—	3,76	—	64,0 %

La Tabla 1 sintetiza los resultados del primer bloque del cuestionario. El informe escrito obtuvo puntuaciones consistentemente superiores en las dimensiones vinculadas al rigor analítico y la estructuración lógica del contenido. El 81,6 % del alumnado afirmó que la redacción del informe les obligó a profundizar en los detalles técnicos de su proyecto ($M = 3,98$), frente al 41,2 % que atribuyó este efecto al diseño del póster ($M = 3,32$). Esta diferencia de 40,4 puntos porcentuales refleja la naturaleza complementaria de ambos formatos: el informe demanda exhaustividad, mientras que el póster exige selección y economía expresiva.

En contraste, ambas herramientas resultaron altamente eficaces para ayudar al alumnado a identificar el mensaje central de su investigación, con valores similares: 79,8 % para el informe ($M = 4,07$) y 76,3 % para el póster ($M = 3,96$). Este hallazgo sugiere que la clarificación del núcleo argumental del trabajo puede lograrse por vías distintas y complementarias: el informe mediante el desarrollo secuencial de la argumentación, y el póster a través del ejercicio de síntesis y jerarquización visual.

En cuanto a la competencia comunicativa ante distintos tipos de audiencia, los datos revelan un patrón diferencial coherente con los objetivos de cada formato. El alumnado percibió una mayor mejora en su capacidad para comunicar a una audiencia general o de negocio mediante el póster (64,0 %, M = 3,76), mientras que la mejora en la comunicación a audiencias técnicas se atribuyó en mayor medida al informe (58,8 %, M = 3,61).

Tabla 2. Valoración de los instrumentos de apoyo al aprendizaje (N = 114).
% Acuerdo = respuestas 4 o 5 en escala Likert de 5 puntos.

Instrumento de apoyo	Media (1-5)	% Acuerdo
Rúbricas de evaluación claras y orientadoras	4,18	83,3 %
Peer review: identificación de mejoras	4,09	77,2 %
Sesión formativa: cambio de mentalidad	3,82	67,5 %

La Tabla 2 presenta la valoración de los instrumentos de apoyo. Las rúbricas de evaluación fueron el recurso mejor valorado, con un 83,3 % de acuerdo y una media de 4,18 sobre 5, lo que indica que el alumnado las percibió como herramientas claras y efectivas para orientar su trabajo. La actividad de revisión por pares también fue muy bien recibida: el 77,2 % afirmó que evaluar los trabajos de sus compañeros les ayudó a identificar errores y posibles mejoras para su propio proyecto (M = 4,09). La sesión formativa sobre diseño de pósteres, aunque con una valoración algo inferior (M = 3,82; 67,5 %), fue reconocida como un elemento facilitador del cambio de mentalidad desde un enfoque textual hacia uno visual.

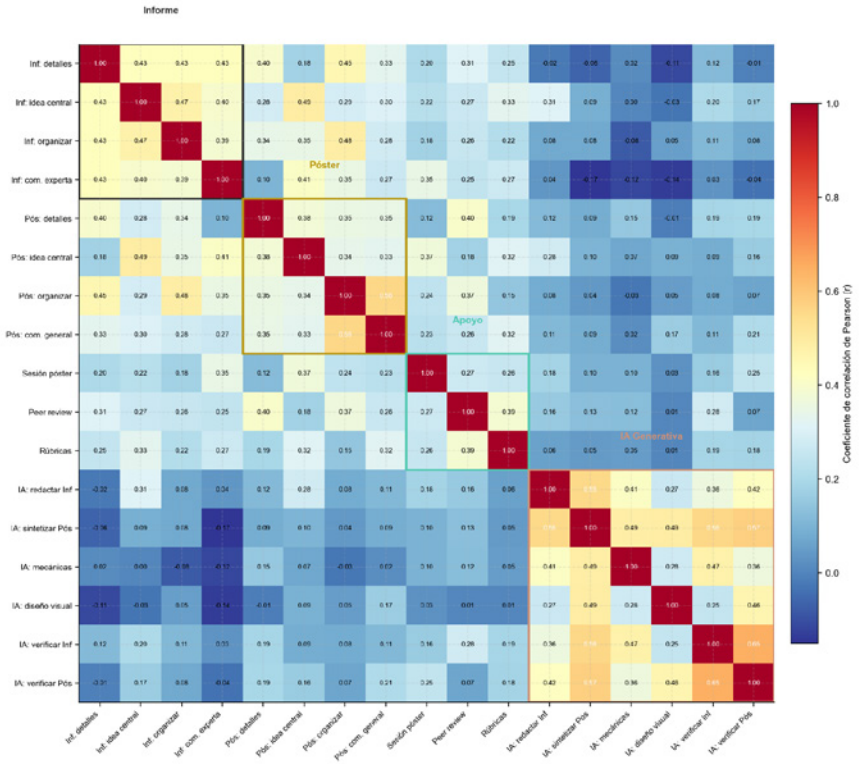


Figura 1. Matriz de correlaciones de Pearson entre los ítems del cuestionario (N = 114). Los recuadros agrupan conceptualmente los ítems por bloques: formato Informe, formato Póster, instrumentos de apoyo e IA generativa.

Para profundizar en las relaciones entre las distintas dimensiones evaluadas, la Figura 1 presenta la matriz de correlaciones de Pearson calculada a partir de las 114 respuestas completas al cuestionario. Se observan correlaciones positivas moderadas-fuertes dentro de cada bloque conceptual. Los ítems del bloque Informe correlacionan entre sí con valores de r entre 0,50 y 0,72, indicando una consistencia interna adecuada en la percepción del alumnado sobre el valor formativo del informe escrito. De forma similar, los ítems del bloque Póster muestran correlaciones de r entre 0,35 y 0,58. Destaca la correlación positiva entre la utilidad percibida de las rúbricas y la valoración de la revisión por pares ($r = 0,52$), así como entre la necesidad de verificar la IA en el informe y la misma necesidad en el póster ($r = 0,62$), lo que sugiere un patrón actitudinal consistente hacia la verificación crítica independientemente del formato. Las correlaciones cruzadas entre bloques son débiles ($r < 0,30$), confirmando que los constructos medidos (informe, póster, apoyo e IA) constituyen dimensiones perceptivas diferenciadas en la experiencia del alumnado.

Tabla 3. Percepción sobre el uso de IA generativa en la elaboración de los formatos (N = 114). Escala Likert de 1 (totalmente en desacuerdo) a 5 (totalmente de acuerdo).

Afirmación sobre uso de IA generativa	Media (1-5)	% Acuerdo
La IA fue fundamental para redactar el Informe	3,32	47,4 %
La IA fue fundamental para sintetizar el Póster	3,28	50,0 %
Usé la IA para tareas mecánicas (estilo, ortografía)	3,38	53,5 %
Usé la IA para estructuras visuales del póster	2,91	39,5 %
Necesidad de verificar la IA más en el Informe	3,86	70,2 %
Necesidad de verificar la IA más en el Póster	3,44	51,8 %

La Tabla 3 recoge los resultados relativos al uso de la IA generativa. Los datos muestran que la herramienta no fue percibida como fundamental para la elaboración de ninguno de los dos formatos de forma mayoritaria: solo el 47,4 % consideró la IA fundamental para el informe y el 50,0 % para el póster. En cambio, la IA fue empleada de forma más generalizada para tareas de carácter mecánico, como la corrección de estilo y ortografía (53,5 %, M = 3,38), mientras que su uso para sugerir estructuras visuales o distribución del espacio en el póster fue el aspecto menos valorado (39,5 %, M = 2,91).

Un hallazgo relevante es la elevada conciencia crítica del alumnado respecto a las limitaciones de la IA generativa. El 70,2 % afirmó que sentía la necesidad de verificar más a fondo las respuestas proporcionadas por la IA cuando las incorporaba al informe técnico (M = 3,86), frente al 51,8 % que experimentó esta misma necesidad en el contexto del póster (M = 3,44). Estos datos sugieren que el rigor técnico exigido por el informe actuó como un filtro natural que incentivó la verificación crítica de los contenidos generados por la IA. La Figura 2 presenta la distribución completa de las respuestas relativas al uso de IA generativa, desglosando los porcentajes de acuerdo, neutralidad y desacuerdo para cada ítem evaluado.

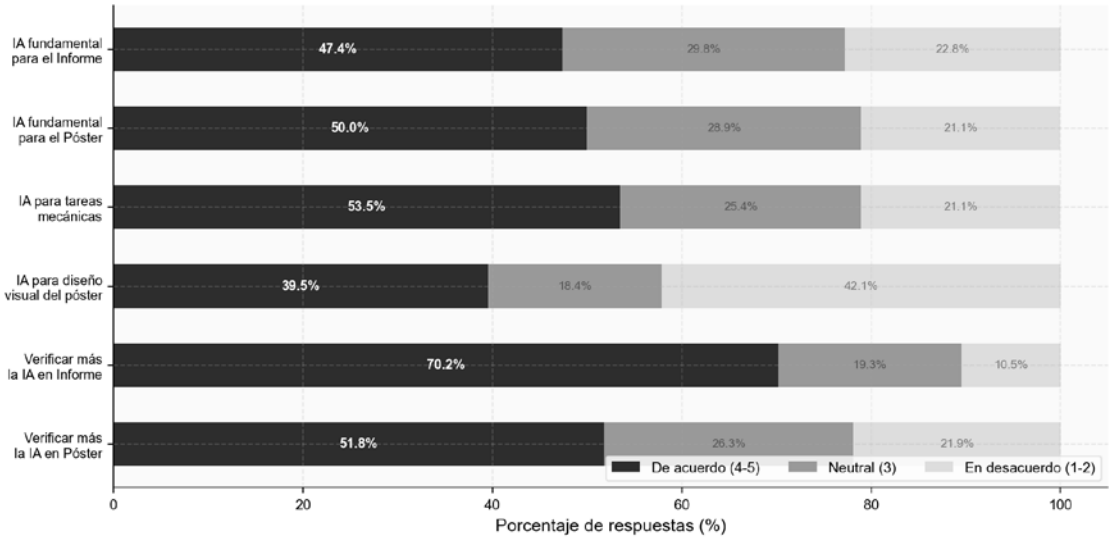


Figura 2. Distribución de respuestas sobre el uso de IA generativa en la elaboración de los formatos (%). Elaboración propia a partir de los datos de la encuesta (N = 114).

Tabla 4. Preferencia y percepción de dificultad por formato (N = 114)

Variable	Informe escrito	Póster científico	Ambos por igual
Formato más desafiante intelectualmente	61 (53,5 %)	32 (28,1 %)	21 (18,4 %)
Formato preferido si solo se pudiera entregar uno	76 (66,7 %)	38 (33,3 %)	—

La Tabla 4 revela una aparente paradoja que ilustra la complementariedad de los formatos. El informe técnico fue considerado la tarea más desafiante por el 53,5 % del alumnado, duplicando la proporción de quienes señalaron el póster (28,1 %). Sin embargo, y a pesar de esta mayor dificultad, el 66,7 % de los estudiantes afirmó que, si tuvieran que elegir un único formato para demostrar formalmente su aprendizaje, elegirían el informe escrito. Este dato sugiere que el alumnado asocia el formato de informe con un mayor rigor académico y una acreditación más sólida de sus conocimientos, aun reconociendo el coste cognitivo que implica.

En cuanto al análisis cualitativo de las preguntas abiertas, la codificación temática de las respuestas al ítem 11 (qué contenido se sacrificó al trasladar el informe al póster) reveló que los elementos más frecuentemente omitidos fueron la descripción detallada de las pruebas realizadas (mencionado en el 11,4 % de las respuestas), los resultados numéricos desglosados (10,5 %), los prompts utilizados (7,0 %) y los detalles metodológicos (6,1 %). De forma consistente, el alumnado describió haber priorizado las conclusiones y los resultados más significativos para garantizar la claridad y el impacto visual del póster. En palabras de un estudiante: "Consideré que era más importante mostrar los resultados obtenidos y hacer un pequeño comentario de las tareas propuestas para evitar el exceso de texto".

El análisis de las respuestas sobre el ahorro de tiempo proporcionado por la IA generativa mostró que la mayoría del alumnado (49,3 %) percibió un mayor ahorro en la redacción del informe escrito, frente al 13,7 % que lo percibió en el diseño del póster y un 5,5 % que lo atribuyó a ambos formatos

por igual. Este resultado es coherente con el carácter más extenso de la tarea de redacción del informe y con el uso de la IA principalmente para tareas de corrección y mejora de estilo.

4. DISCUSIÓN Y CONCLUSIONES

Los resultados obtenidos sugieren que la dicotomía "póster vs informe" no debe interpretarse como una oposición excluyente, sino como una complementariedad pedagógica de alto valor formativo. El informe técnico y el póster científico activan procesos cognitivos diferentes pero sinérgicos: mientras el primero consolida la argumentación lógica, la justificación metodológica y el rigor analítico, el segundo entrena la capacidad de abstracción, la jerarquización de la información y la comunicación visual. Esta dualidad es coherente con los planteamientos de la teoría del aprendizaje experiencial de Kolb (1984), que postula que el aprendizaje profundo requiere la integración de modos de aprehensión y transformación del conocimiento diversos y complementarios.

La comparación con los resultados de investigaciones previas refuerza estos hallazgos. Walkington et al. (2017) documentaron que la participación en congresos estudiantiles mejora significativamente la motivación y la percepción de autoeficacia del alumnado, aunque su estudio no discriminó entre los efectos atribuibles a cada formato de comunicación. Nuestros datos aportan evidencia empírica sobre la contribución específica de cada modalidad, revelando que el informe escrito es percibido como el vehículo más adecuado para el aprendizaje en profundidad, mientras que el póster destaca como catalizador de la capacidad de síntesis y adaptación del discurso a audiencias no especializadas.

En relación con la inteligencia artificial generativa, los resultados matizan algunos discursos que anticipan una dependencia acrítica del alumnado hacia estas herramientas (Cotton et al., 2023). Si bien la IA fue utilizada por una mayoría de estudiantes para tareas mecánicas, una proporción aún mayor manifestó una actitud de verificación rigurosa, especialmente en el contexto del informe técnico.. Este dato sugiere que la exigencia de rigor inherente al formato escrito actúa como un mecanismo natural de control de calidad que fomenta el pensamiento crítico sobre los contenidos generados por la IA. Como señalan Mollick y Mollick (2023), la clave pedagógica no reside en prohibir estas herramientas, sino en diseñar actividades que obliguen al alumnado a ejercer un juicio crítico sobre sus resultados.

La elevada valoración de las rúbricas de evaluación y de la revisión por pares respalda el papel de estos instrumentos como base del aprendizaje. Las rúbricas proporcionan al alumnado criterios explícitos que reducen la incertidumbre y orientan el esfuerzo hacia los objetivos de aprendizaje (Jonsson y Svingby, 2007), mientras que la revisión por pares activa procesos metacognitivos que difícilmente se desarrollan mediante la mera recepción pasiva de retroalimentación del profesorado (Nicol et al., 2014).

Entre las limitaciones del estudio, cabe señalar que los resultados se basan exclusivamente en la percepción del alumnado, sin incluir una evaluación externa del aprendizaje real alcanzado. Futuras investigaciones podrían complementar este enfoque con medidas objetivas de rendimiento, como la comparación de calificaciones entre formatos o el análisis de la calidad de los productos elaborados. Asimismo, la muestra, aunque representativa de la asignatura, procede de un único contexto institucional, lo que aconseja cautela en la generalización de los resultados a otras titulaciones o universidades.

En conclusión, la integración simultánea del informe técnico y el póster científico en el marco de un mini-congreso estudiantil puede constituir una estrategia didáctica sinérgica y replicable para la

formación integral del ingeniero. Dado que la evidencia procede de datos de percepción, estas conclusiones deben interpretarse como indicios del valor formativo percibido y no como medidas del aprendizaje efectivo. El informe consolida el rigor y la exhaustividad técnica, el póster entrena la abstracción visual y la comunicación efectiva, y la combinación de ambos (apoyada por rúbricas claras, revisión por pares y un uso crítico de la IA generativa) conforma un ecosistema de aprendizaje que trasciende la mera adquisición de contenidos para desarrollar competencias profesionales genuinas. La experiencia demuestra que obligar al alumnado a transitar del "texto al píxel" (y viceversa) no solo no es redundante, sino que constituye un ejercicio de traducción cognitiva de alto valor formativo, especialmente pertinente en un contexto profesional donde la capacidad de adaptar el discurso a distintos formatos y audiencias define, cada vez más, la excelencia del ingeniero.

5. REFERENCIAS

Autor. (2025a). De la técnica a la comunicación: Integración de IA generativa y formatos de congreso en la enseñanza de la ingeniería. En *REDES-INNOVAESTIC 2025*. Universidad de Alicante. [Referencia anonimizada para revisión por pares].

Cotton, D. R. E., Cotton, P. A. y Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, *61*(2), 228-239. https://doi.org/10.1080/14703297.2023.2190148

Crawford, J., Cowling, M. y Allen, K. A. (2023). Leadership is needed for ethical ChatGPT: Character, assessment, and learning using artificial intelligence (AI). *Journal of University Teaching and Learning Practice*, *20*(3), Artículo 02. https://doi.org/10.53761/1.20.3.02

García-Peñalvo, F. J., Corell, A., Abella-García, V. y Grande-de-Prado, M. (2020). La evaluación online en la educación superior en tiempos de la COVID-19. *Education in the Knowledge Society*, *21*, Artículo 26. https://doi.org/10.14201/eks.23013

Jonassen, D., Strobel, J. y Lee, C. B. (2006). Everyday problem solving in engineering: Lessons for engineering educators. *Journal of Engineering Education*, *95*(2), 139-151. https://doi.org/10.1002/j.2168-9830.2006.tb00885.x

Jonsson, A. y Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, *2*(2), 130-144. https://doi.org/10.1016/j.edurev.2007.05.002

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Prentice-Hall.

Mollick, E. y Mollick, L. (2023). Using AI to implement effective teaching strategies in classrooms: Five strategies, including prompts. *The Wharton School Research Paper*. https://doi.org/10.2139/ssrn.4391243

Nicol, D., Thomson, A. y Breslin, C. (2014). Rethinking feedback practices in higher education: A peer review perspective. *Assessment & Evaluation in Higher Education*, *39*(1), 102-122. https://doi.org/10.1080/02602938.2013.795518

UNESCO. (2023). *Guidance for generative AI in education and research*. UNESCO Publishing. https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research

Walkington, H., Hill, J. y Kneale, P. E. (2017). Reciprocal elucidation: A student-led pedagogy in multidisciplinary undergraduate research conferences. *Higher Education Research & Development*, *36*(2), 416-429. https://doi.org/10.1080/07294360.2016.1208155

5. Del lenguaje especializado al lenguaje claro: innovación docente y alfabetización en salud mediante el etiquetado alimentario[1]

Isabel Balteiro

Universidad de Alicante

RESUMEN

La alfabetización en salud constituye una competencia fundamental para la toma de decisiones informadas y para la reducción de desigualdades sociales vinculadas al acceso al conocimiento sanitario. En este contexto, el etiquetado alimentario representa un discurso especializado de uso cotidiano cuya comprensión exige competencias lingüísticas, conceptuales y numéricas que no siempre están presentes en públicos no expertos. Este trabajo presenta una experiencia de innovación docente basada en aprendizaje-servicio desarrollada en la Universidad de Alicante en torno al análisis, adaptación y mediación terminológica del etiquetado alimentario. Los objetivos fueron identificar dificultades terminológicas en etiquetas reales, diseñar recursos accesibles en lenguaje claro y valorar el impacto formativo y social de la intervención. La metodología incluyó como instrumentos fundamentales un corpus de etiquetas alimentarias y la aplicación de una plantilla de análisis terminológico, como bases para la elaboración de glosarios y materiales divulgativos, además del uso de pruebas pre/post, rúbricas competenciales y cuestionarios de percepción. Los resultados muestran que el alumnado identifica la tecnicidad, la opacidad, los códigos y la ambigüedad como principales obstáculos de comprensión, y desarrolla competencias en mediación terminológica, comunicación clara, conciencia social y ética profesional. Se concluye que el etiquetado alimentario constituye un espacio privilegiado de intervención educativa y responsabilidad social universitaria.

PALABRAS CLAVE: alfabetización en salud; lenguaje claro; mediación terminológica; etiquetado alimentario; aprendizaje-servicio

1. INTRODUCCIÓN

La capacidad para comprender, interpretar y utilizar información relacionada con la salud es una competencia esencial en las sociedades contemporáneas. Sin embargo, la abundancia de información sanitaria, nutricional y alimentaria no garantiza por sí misma una ciudadanía mejor informada y más saludable (cf. Nutbeam & Lloyd, 2021), ya que el acceso a los textos no implica necesariamente su comprensión ni su aplicación a situaciones concretas de la vida cotidiana. En el ámbito de la alimentación, los usuarios se enfrentan de manera recurrente a la problemática de interpretar etiquetas, tablas nutricionales, declaraciones de propiedades saludables, listas de ingredientes, advertencias de alérgenos y códigos de aditivos que condensan información especializada en formatos breves, normativizados y de alta densidad conceptual.

La alfabetización en salud ha sido definida como un objetivo fundamental de la salud pública, no solo porque permite a las personas acceder a información sanitaria, sino porque las capacita

1 El presente trabajo ha contado con financiación del Vicerrectorado de Igualdad, Inclusión y Responsabilidad Social (Convocatoria de Proyectos de Aprendizaje-Servicio 2025-2026).

para comprenderla, evaluarla y utilizarla en la toma de decisiones (Nutbeam, 2000). Sørensen et al. (2012) proponen una concepción integrada de la alfabetización en salud como la capacidad para acceder, comprender, valorar y aplicar información relacionada con la salud en contextos de prevención, promoción y atención sanitaria. Desde esta perspectiva, comprender una etiqueta alimentaria no es una operación meramente lectora, sino una práctica compleja que exige activar competencias lingüísticas, numéricas, conceptuales y críticas.

El etiquetado alimentario constituye un objeto de estudio especialmente relevante porque se sitúa en la intersección entre discurso especializado, educación para la salud y ciudadanía. Aunque su función principal es informar al consumidor, su comprensión requiere conocimientos que no siempre forman parte de la experiencia cotidiana de la población. Términos como "ácidos grasos saturados", "edulcorantes", "proteínas de alto valor biológico" o "E-250" remiten a conceptos técnicos procedentes de la nutrición, la química, la tecnología de los alimentos y la salud pública. Por ello, el etiquetado alimentario puede considerarse un género discursivo híbrido, a caballo entre el lenguaje usado entre expertos y el lenguaje divulgativo: se dirige al conjunto de la ciudadanía, pero incorpora rasgos propios del discurso especializado.

Desde la teoría terminológica, Cabré (1999) recuerda que los términos no son únicamente denominaciones, sino unidades de conocimiento que permiten organizar y transmitir conceptos especializados. En el caso del etiquetado alimentario, esta función puede convertirse en una barrera cuando el receptor no comparte el mismo marco conceptual que el emisor experto. Gotti (2003) subraya también que los discursos especializados presentan rasgos que los diferencian de los discursos generales, entre ellos la precisión conceptual, la condensación informativa y la presencia de tecnicismos. Estos rasgos, útiles para la comunicación entre especialistas, pueden dificultar la comprensión cuando el destinatario es lego. Por ello, como señala Zethsen (2009), la reformulación dentro de una misma lengua es necesaria cuando existe una distancia entre el lenguaje experto y el lenguaje de los destinatarios. Así, adaptar "lípidos" como "grasas", explicar "valor energético" como "calorías o energía que aporta el alimento" o reformular "edulcorantes artificiales" como "sustancias que endulzan sin azúcar" son operaciones de mediación terminológica que permiten reducir la distancia entre conocimiento experto y comprensión ciudadana.

La literatura sobre uso y comprensión del etiquetado ha mostrado que las dificultades de interpretación no son anecdóticas. Campos et al. (2011), en una revisión sistemática sobre etiquetas nutricionales, señalan que el uso efectivo de la información depende de múltiples factores: nivel educativo, familiaridad con el formato, capacidad numérica y conocimiento nutricional previo. Rothman et al. (2006) demostraron que la alfabetización y la competencia numérica influyen de manera decisiva en la interpretación de la información nutricional. En esta línea, Malloy-Weir y Cooper (2017) muestran que la comprensión y el uso de las etiquetas está condicionada por la alfabetización general, la alfabetización en salud y la competencia numérica, especialmente cuando el consumidor debe comparar productos, interpretar porcentajes y tamaños de las porciones o relacionar los datos de la etiqueta con sus propias necesidades alimentarias. La investigación más reciente ha reforzado esta perspectiva al mostrar que la comprensión y el uso efectivo del etiquetado nutricional no se distribuyen homogéneamente entre la población. En particular, factores socioeconómicos y educativos pueden condicionar la capacidad de los consumidores para interpretar y utilizar la información nutricional disponible en los envases (Shrestha et al., 2023). Asimismo, la evidencia reciente confirma el potencial de sistemas de etiquetado accesibles para facilitar la comprensión y

favorecer elecciones alimentarias informadas (Kelly et al., 2024). Así, Anastasiou et al. (2019) y Pfledderer et al. (2024) relacionaron el uso de la información nutricional con patrones alimentarios más saludables en población adulta y adolescente, respectivamente. Estos trabajos refuerzan, por tanto, la conveniencia de abordar no solo la disponibilidad de la información, sino también las capacidades necesarias para interpretarla y aplicarla.

Desde el punto de vista educativo, estas ideas y las dificultades anteriormente mencionadas abren un espacio de intervención de gran interés. El etiquetado alimentario permite trabajar de manera integrada la comprensión lectora, la alfabetización científica, la competencia matemática, el pensamiento crítico, la educación para la salud y el consumo responsable. Además, ofrece materiales auténticos, cotidianos y fácilmente transferibles a la vida diaria. En otras palabras, como indican Rana et al., (2025), las intervenciones dirigidas a desarrollar la alfabetización en etiquetado pueden contribuir a mejorar tanto la comprensión de la información nutricional como su aplicación a situaciones cotidianas y capacidad de elección alimentaria. Este potencial didáctico se ve reforzado cuando la intervención adopta metodologías activas de aprendizaje-servicio. Bringle y Hatcher (1996) definen el aprendizaje-servicio como una experiencia educativa organizada que combina servicio a la comunidad, reflexión y aprendizaje curricular, por lo que resulta especialmente adecuado para proyectos en los que la universidad actúa como mediadora entre conocimiento experto y ciudadanía. Según Eyler y Giles (1999), el aprendizaje-servicio favorece la conexión entre aprendizaje académico y compromiso social, al permitir que el alumnado aplique conocimientos en contextos reales y contribuya a responder a necesidades comunitarias.

En este marco se sitúa la experiencia descrita en este trabajo, desarrollada en el seno de un proyecto de aprendizaje-servicio (ApS) en la Universidad de Alicante y orientada a mejorar la comprensión del etiquetado alimentario mediante estrategias de mediación terminológica y lenguaje claro. La propuesta se fundamenta en la idea de que la alfabetización en salud no puede limitarse a transmitir información nutricional, sino que debe atender a la forma lingüística en que dicha información se presenta. Si bien la terminología especializada puede facilitar la precisión conceptual en la comunicación entre expertos, también puede generar opacidad, dependencia interpretativa y desigualdad comunicativa en legos. Por ello, la intervención se centra en identificar dificultades terminológicas, adaptar el lenguaje especializado y diseñar recursos divulgativos accesibles para públicos no expertos.

El objetivo general de este estudio es analizar el potencial del etiquetado alimentario como herramienta de alfabetización en salud mediante una experiencia de aprendizaje-servicio basada en la mediación terminológica. De este se derivan cuatro objetivos específicos: 1) identificar dificultades terminológicas presentes en el etiquetado alimentario; 2) diseñar recursos accesibles en lenguaje claro dirigidos a públicos no expertos; 3) evaluar el impacto formativo de la experiencia en el alumnado universitario; y 4) valorar su aportación social en términos de comprensión, autonomía y conciencia crítica. Como hipótesis de trabajo se plantea que una intervención basada en análisis terminológico, reformulación en lenguaje claro y aprendizaje-servicio contribuye al desarrollo de competencias lingüísticas, profesionales y sociales en el alumnado, al tiempo que mejora la accesibilidad de la información nutricional para la población destinataria.

2. MÉTODO

2.1. Descripción del contexto y de los participantes

La experiencia se desarrolló en el marco de un proyecto ApS de innovación docente y responsabilidad social universitaria de la Universidad de Alicante, centrado en la alfabetización en salud a través del etiquetado alimentario. El proyecto se articuló en varias asignaturas universitarias vinculadas a dos grados —Nutrición Humana y Dietética y Educación— y a dos programas de máster relacionados con los lenguajes profesionales y con la innovación educativa, respectivamente, lo que permitió abordar el problema desde una perspectiva interdisciplinar.

El alumnado universitario participó en las distintas fases: análisis de etiquetas reales, identificación de dificultades terminológicas, diseño de glosarios, elaboración de materiales divulgativos y reflexión sobre el impacto social de la comunicación clara.

La preparación para la intervención comunitaria se planteó bajo el título "Etiquetado para todos: Nutrición sin barreras". En ella, el aula se organizó en 10 subgrupos de trabajo, con aproximadamente cuatro estudiantes por grupo. Cada subgrupo desarrolló una actividad diferente en relación con el etiquetado alimentario y con los colectivos destinatarios de la intervención, como veremos.

2.2. Instrumentos

La experiencia utilizó varios instrumentos de recogida y análisis de datos, diseñados para evaluar tanto la dimensión formativa como la dimensión social del proyecto. El primer instrumento es un corpus de etiquetas alimentarias auténticas pertenecientes a distintos grupos de alimentos, lo que permitió trabajar con textos auténticos y exponer al alumnado a la complejidad terminológica, visual y numérica propia del etiquetado alimentario. Entre otros, se encuentran documentados en el corpus unidades y fragmentos terminológicos tales como "jarabe de glucosa-fructosa", "lípidos", "grasas trans", "proteínas de alto valor biológico", códigos de aditivos como "E-250" y "E-330", abreviaturas como "kcal" y expresiones cuantitativas como "por 100 g", "por ración" y "% VRN"; ejemplos que permiten observar que la dificultad no depende únicamente de términos, sino también de códigos, unidades, expresiones cuantitativas y formulaciones aparentemente transparentes que requieren conocimiento especializado.

El segundo instrumento, una plantilla de análisis de dificultades terminológicas, fue diseñada, a partir de dimensiones de dificultad y comprensión identificadas en trabajos recientes sobre la temática que nos ocupa (véase Kloehn et al., 2018; Mukherjee et al., 2018; Silva et al., 2023; Yıldırım & Çakır, 2025), con el fin de identificar y clasificar los elementos que podían dificultar la comprensión a legos: Tras la identificación del producto o tipo de alimento al que se refiere la unidad a estudiar, se analiza su complejidad léxica, baja transparencia semántica, grado de especialización conceptual, ambigüedad, uso de códigos o abreviaturas, complejidad sintáctica y carga numérica, como sigue:

Tabla 1. Plantilla de análisis de dificultad terminológica

Campo	Criterio
Unidad analizada	¿Qué palabra, término, sintagma, abreviatura, código o dato se analiza?
Contexto	¿En qué parte de la etiqueta aparece? Considerar el fragmento complete de la etiqueta en el que aparece la unidad.

Campo	Criterio
Complejidad léxica y/o estructural	Grado en que la forma lingüística puede dificultar el reconocimiento o procesamiento. ¿Es una palabra poco frecuente? ¿Es larga o morfológicamente compleja? ¿Contiene varios elementos especializados? Presencia de formas codificadas que no explicitan directamente el concepto. ¿Qué término, sintagma, código o dato se analiza?
Transparencia semántica	Facilidad con que un lego puede inferir el significado a partir de la forma del término. ¿Puede inferirse el significado a partir de la forma? ¿Su significado resulta evidente sin conocimientos previos?
Especialización conceptual	Cantidad de conocimiento especializado necesario para comprender correctamente el concepto. ¿Pertenece al lenguaje común o requiere conocimientos de nutrición, química, medicina o tecnología alimentaria?
Ambigüedad	Posibilidad de que el término o expresión genere una interpretación incorrecta o excesivamente simplificada. ¿Admite interpretaciones (no) equivalentes o engañosas?
Complejidad sintáctica	Dificultad derivada de la estructura del sintagma o del segmento en que aparece. ¿La comprensión exige procesar estructuras superiores a la palabra?
Carga o complejidad numérica	Necesidad de interpretar números, unidades, porcentajes, proporciones o cantidades. ¿La comprensión exige procesar cantidades?

Esta plantilla se complementó con una propuesta de Índice de Dificultad Terminológica (en adelante, IDT), concebido como una escala exploratoria para estimar el grado de dificultad de un término o sintagma terminológico. Aunque en esta fase el índice se utilizó de manera formativa y no como instrumento estadísticamente validado, su diseño permitió orientar la reflexión del alumnado hacia dimensiones concretas de la dificultad terminológica: familiaridad, transparencia, especialización, ambigüedad, codificación técnica, complejidad sintáctica y carga numérica. Esta aportación metodológica desplaza el análisis desde una intuición general —"esta palabra es difícil"— hacia una evaluación más sistemática del desajuste entre el lenguaje especializado y las competencias del receptor.

El IDT se calculó asignando a siete dimensiones una puntuación de 0 a 2: familiaridad, transparencia semántica, especialización conceptual, ambigüedad, codificación técnica, complejidad sintáctica y carga numérica. En cada dimensión, 0 indicaba ausencia o dificultad mínima, 1 dificultad intermedia y 2 dificultad elevada. La suma generaba una puntuación exploratoria entre 0 y 14 puntos (0–4: dificultad baja; 5–9: media; 10–14: alta).

Tabla 2. Criterios del Índice de Dificultad Terminológica (IDT)

Dimensión	0	1	2
Familiaridad	Uso cotidiano	Reconocible pero poco frecuente	Infrecuente, morfológicamente compleja y/o desconocida para un lego
Transparencia	Claro: significado directa o fácilmente inferible	Parcialmente inferible	Opaco: significado no inferible para un lego
Especialización	Común: concepto cotidiano, léxico general	Semiespecializado; requiere cierto conocimiento previo	Técnico: Requiere conocimiento técnico o especializado

Dimensión	0	1	2
Ambigüedad	Interpretación estable o generalmente inequívoca	Dependiente del contexto; puede generar alguna duda	Alta probabilidad de interpretación errónea; puede inducir interpretaciones divergentes
Codificación técnica	Sin códigos/abreviaturas	Unidad o abreviatura (relativamente) conocida	Código o abreviatura sin explicación
Complejidad sintáctica	Unidad o construcción simple	Sintagma breve y/o moderadamente condensado	Estructura larga o condensada
Carga numérica	No exige procesamiento numérico	Cantidad/unidad que exige operación o relación sencilla	Porcentajes, raciones o comparaciones múltiples que requieren comparación o integración

El tercer instrumento consiste en pruebas pre- y post-intervención dirigidas al alumnado universitario. Estas pruebas evaluaban conocimientos iniciales y finales sobre alfabetización en salud, lenguaje especializado y dificultad del discurso médico-nutricional, así como la capacidad de identificar términos problemáticos y reformularlos en lenguaje claro. Asimismo, incorporaban cuestiones sobre desigualdad comunicativa y responsabilidad profesional en la accesibilidad de la información.

El cuarto instrumento, formado por pruebas pre- y post-intervención dirigidas a la población destinataria de la fase de servicio, y diseñadas en lenguaje claro a partir de situaciones prácticas, valora la comprensión de términos básicos del etiquetado, la interpretación de información nutricional, la identificación de alérgenos y la aplicación de lo aprendido a decisiones cotidianas de compra. También incluía preguntas de autopercepción sobre confianza y facilidad para leer etiquetas.

El quinto instrumento lo constituyen rúbricas competenciales para valorar el desempeño del alumnado y de la población destinataria. En el caso del alumnado, se evaluaron el análisis terminológico, la comprensión conceptual, la mediación terminológica, el uso de lenguaje claro, el diseño de materiales, la aplicación práctica, el trabajo en equipo y la conciencia social. Cada dimensión se valoró atendiendo a cuatro niveles de desempeño: inicial, básico, competente y avanzado. Por ejemplo, en mediación terminológica, el nivel inicial correspondía a una sustitución léxica insuficiente o conceptualmente imprecisa; el básico, a una simplificación comprensible pero incompleta; el competente, a una explicación clara que conservaba el contenido esencial; y el avanzado, a una explicación clara, precisa y adaptada al destinatario, con contextualización o ejemplo funcional. Este mismo principio se aplicó al resto de dimensiones, de modo que la rúbrica no valoraba solo la corrección técnica, sino también la adecuación al receptor. En el caso de la población destinataria, la rúbrica se centró en la comprensión de términos, la interpretación de etiquetas, la identificación de riesgos, la autonomía y la confianza.

Finalmente, se aplicó al alumnado un cuestionario de percepción con ítems cerrados y preguntas abiertas (véase Tabla 3). Este instrumento permitió recoger evidencias cualitativas sobre la evolución de sus concepciones acerca de la alfabetización en salud, la dificultad terminológica, la mediación en lenguaje claro, la responsabilidad comunicativa y la contribución del proyecto a los ODS. La Tabla 3 recoge la formulación operativa de los principales tipos de ítems empleados, sin reproducir aquí la totalidad de los formularios.

Tabla 3. Ejemplos de ítems y tareas

Dimensión	Ejemplo de ítem o tarea	Tipo de respuesta/evidencia
Identificación de dificultad	Valora "jarabe de glucosa-fructosa": ¿resulta claro, complejo o inadecuado para un público no experto? Justifica brevemente.	Elección + justificación
Tecnicidad	Valora "lípidos" atendiendo a su grado de tecnicidad y familiaridad para la población general.	Categoría + comentario
Mediación terminológica	Explica "grasas saturadas" a una persona sin formación nutricional utilizando lenguaje claro.	Respuesta abierta
Reformulación	Reescribe "contenido energético por ración" de manera comprensible sin eliminar la información esencial.	Respuesta abierta
Conciencia comunicativa	Indica qué características de una etiqueta pueden dificultar su comprensión y explica por qué.	Respuesta abierta
Desarrollo competencial	¿Qué competencias consideras que has desarrollado al analizar y adaptar etiquetas reales?	Respuesta abierta
Responsabilidad social	¿De qué manera puede el lenguaje de las etiquetas generar o reducir desigualdades de acceso a la información?	Respuesta abierta
ODS	¿Con qué Objetivos de Desarrollo Sostenible relacionarías la actividad? Justifica la relación.	Selección + justificación

2.3. Procedimiento

El procedimiento se desarrolló en cinco fases consecutivas. La primera fase consistió en la sensibilización teórica y la contextualización del problema. El profesorado introdujo conceptos básicos relacionados con alfabetización en salud, discurso especializado, terminología, mediación intralingüística y lenguaje claro, destacando que las barreras comunicativas derivan tanto de las competencias del receptor como de las características del texto, entre ellas la densidad del lenguaje especializado y la falta de contextualización. Esta fase permitió situar el etiquetado alimentario como objeto de análisis lingüístico, educativo y social.

La segunda fase se centró en la recopilación y análisis del corpus de etiquetas alimentarias. El alumnado trabajó con envases reales y etiquetas de productos cotidianos. A partir de la plantilla de análisis terminológico descrita anteriormente, identificó unidades potencialmente problemáticas, clasificó el tipo de dificultad y justificó su interpretación. El análisis atendió a la lista de ingredientes, la tabla nutricional, las declaraciones nutricionales y de salud, los alérgenos, los códigos de aditivos y las expresiones promocionales. Esta fase mostró que la dificultad del etiquetado no se limita a palabras desconocidas, sino que también incluye formatos numéricos, ambigüedad discursiva, términos aparentemente transparentes pero conceptualmente complejos y estrategias de marketing que pueden orientar la interpretación del consumidor.

La tercera fase consistió en la adaptación terminológica y el diseño de recursos de mediación. A partir de los términos identificados, el alumnado, dividido en 10 subgrupos, elaboró glosarios monolingües

y bilingües con definiciones en lenguaje claro, así como materiales adaptados al público destinatario. Estos materiales incluyeron folletos e infografías sobre cómo leer una etiqueta, pósteres de referencia para centros comunitarios, glosarios de tecnicismos reformulados mediante analogías cotidianas, vídeos explicativos de tres o cuatro minutos con envases reales y actividades educativas como, entre otras, "El semáforo del súper", "Detectives del azúcar" o "Bingo de ingredientes", diseñadas como tareas breves y manipulativas. En "El semáforo del súper", los participantes examinaban fragmentos de etiquetas y los clasificaban en verde, amarillo o rojo según su facilidad de comprensión, justificando qué elemento lingüístico o numérico provocaba la dificultad y proponiendo una reformulación cuando era necesario. En "Detectives del azúcar", comparaban etiquetas de productos de una misma categoría para localizar azúcares con denominaciones diferentes, distinguir la lista de ingredientes de la información nutricional y explicar en lenguaje claro qué información permitía (y cuál no) concluir la etiqueta. En "Bingo de ingredientes", cada participante disponía de una cuadrícula con términos, códigos o expresiones frecuentes; debía localizarlos en envases reales, marcar la casilla correspondiente y explicar el significado o la función del elemento identificado. El objetivo era convertir la lectura de etiquetas en una práctica guiada de reconocimiento, interpretación y mediación.

La cuarta fase correspondió a la implementación de la intervención comunitaria, realizada por el profesorado y una selección de alumnado voluntario. Los materiales preparados en las aulas universitarias se aplicaron en talleres dirigidos a colectivos vulnerables y personas mayores. La dinámica combinó explicación, manipulación de envases, comparación de productos, identificación de ingredientes y resolución de dudas. Para favorecer la participación activa, se evitó una comunicación unidireccional y se propusieron tareas aplicadas, como comparar dos productos similares, identificar azúcares añadidos, reconocer alérgenos o valorar si un producto presentaba un contenido elevado en grasas saturadas.

La quinta fase fue la evaluación. El alumnado completó pruebas pre/post, rúbricas y cuestionarios de percepción, mientras que la población destinataria realizó pruebas adaptadas antes y después de la intervención para medir comprensión, interpretación funcional y confianza. Los datos se analizaron mediante recuentos descriptivos, categorización cualitativa de respuestas abiertas y análisis interpretativo de evidencias competenciales. La triangulación de instrumentos permitió combinar tendencias generales, procedentes de las respuestas cerradas, con ejemplos cualitativos de mediación, conciencia crítica y desarrollo profesional. Para asegurar la trazabilidad entre actividad, instrumento y resultado, las respuestas se organizaron según las mismas categorías de la plantilla (tecnicidad, transparencia, codificación, ambigüedad y carga numérica) y según las dimensiones competenciales de la rúbrica. De este modo, tanto los porcentajes como los ejemplos cualitativos que se presentan a continuación pueden vincularse con tareas e indicadores concretos.

3. RESULTADOS

Los resultados se presentan en cuatro bloques relativos al aprendizaje del alumnado en relación a identificación de dificultades terminológicas, producción de reformulaciones en lenguaje claro, así como desarrollo competencial del alumnado e impacto social percibido. En primer lugar, los datos muestran que el alumnado fue capaz de identificar de manera consistente el carácter complejo de determinados términos del etiquetado alimentario (véase Tabla 4). En relación con el término "jarabe de glucosa-fructosa", 21 (72,4%) de los 29 participantes lo clasificaron como complejo, frente a 6 (20,7%) que lo consideraron claro, 1 que lo marcó simultáneamente como claro y complejo y 1 (3,44%) que lo con-

sideró incorrecto. Este resultado evidencia que la mayoría del alumnado reconoció la opacidad de un término que, aunque frecuente en etiquetas, no resulta transparente para el público general. Su dificultad no procede solo de la longitud del sintagma, sino de la distancia entre la denominación técnica y una categoría cotidiana más comprensible, como "azúcar añadido" o "azúcar concentrado".

En la valoración del término "lípidos" se observó un patrón similar. En este caso, 27 (93%) de los 29 participantes señalaron que el término podía ser problemático al ser de carácter técnico. Solo un participante lo consideró muy común y otro marcó simultáneamente "técnico" e "incorrecto". Este resultado confirma que el alumnado percibe la tecnicidad como uno de los principales factores de dificultad terminológica. Aunque "lípidos" pueda ser un término habitual en contextos académicos o sanitarios, es importante que el alumnado sea consciente (como parece ser el caso) de que no necesariamente forma parte del repertorio léxico cotidiano de públicos no especializados.

Tabla 4. Ejemplo de valoración de la opacidad terminológica por el alumnado

Término	Valoración	% alumnado
Jarabe de glucosa-fructosa	Complejo	21 (72,4%)
	Claro	6 (20,7%)
	Claro y Complejo	1 (3,44%)
	Incorrecto	1 (3,44%)
Lípidos	Técnico	27 (72,4%)
	Muy común	1 (3,44%)
	Técnico e incorrecto	1 (3,44%)

Las respuestas abiertas permitieron identificar cinco grandes tipos de dificultad terminológica: tecnicidad léxica, baja transparencia semántica, codificación técnica, ambigüedad y carga numérica (véase Tabla 5). La tecnicidad se asoció con palabras científicas o poco habituales, como "emulgente", "proteínas de alto valor biológico" o "grasas trans". La baja transparencia apareció en términos cuya forma no permite inferir fácilmente el significado, como "lípidos" o "valor energético". La codificación técnica se ejemplificó mediante aditivos como "E-250", calificada por algunos como especialmente opacos por su combinación de letra y número sin explicar la función del ingrediente. La ambigüedad se vinculó a términos como "natural", "light" o "sin azúcares añadidos", que pueden generar interpretaciones erróneas. Finalmente, la carga numérica se relacionó con la dificultad de interpretar datos "por 100 g", "por ración", porcentajes o cantidades en gramos y miligramos.

Tabla 5. Tipos de dificultades terminológicas identificadas

Tipo de dificultad	Descripción	Ejemplos
Tecnicidad léxica	Uso de términos científicos o especializados	emulgente, grasas trans
Baja transparencia	La forma del término no permite inferir el significado	valor energético, jarabe de glucose-fructosa
Codificación técnica	Uso de siglas, letras, números o códigos	E-250, E-330, kcal
Ambigüedad discursiva	Expresiones que pueden inducir a interpretaciones imprecisas	light, natural, sin azúcares añadidos
Carga numérica	Necesidad de interpretar cantidades, porcentajes o raciones	Por 100 g, por ración, % VRN

En segundo lugar, los datos muestran alumnado competente para reformular términos especializados en lenguaje claro (véase Tabla 6). En la explicación de "grasas saturadas", la mayoría recurrió a fórmulas como "grasas no saludables", "grasas que conviene limitar" o "grasas que en exceso pueden ser malas para la salud". Aunque algunas reformulaciones simplifican en exceso la distinción nutricional, el patrón general muestra que el alumnado comprendió la necesidad de añadir orientación pragmática y no limitarse a una equivalencia léxica. Es decir, no se trató solo de traducir "grasas saturadas" por otro término, sino de explicar su relevancia para la toma de decisiones.

En el caso del término "valor energético", las respuestas fueron más homogéneas. La mayoría lo explicó como "cantidad de energía que aporta un alimento", "calorías que aporta el alimento" o "energía en kcal que tiene cada alimento". Este resultado sugiere que algunos términos técnicos admiten mediaciones relativamente estables cuando existe un equivalente cotidiano ampliamente disponible, como "calorías". En cambio, otros términos, como "grasas trans" o "proteínas de alto valor biológico", generaron mayor variabilidad, lo que indica una dificultad conceptual superior.

La reescritura de "edulcorantes artificiales", "contenido energético por ración" y "grasas trans" ofrece evidencias claras de mediación intralingüística. Algunas respuestas reformularon "edulcorantes artificiales" como "sustancias que endulzan sin usar azúcar", "sustitutos del azúcar" o "endulzantes no naturales". "Contenido energético por ración" fue reformulado como "calorías por porción", "energía que aporta una parte del alimento" o "cantidad de calorías por ración". "Grasas trans" se explicó como "grasas poco saludables", "grasas industriales dañinas" o "grasas que hay que evitar". Estas respuestas muestran que el alumnado desarrolló estrategias de sustitución léxica, explicitación, simplificación sintáctica y contextualización.

Tabla 6. Ejemplos de mediación terminológica producida por el alumnado

Término especializado	Reformulación en lenguaje claro
Edulcorantes artificiales	Sustancias que endulzan sin azúcar
Contenido energético por ración	Calorías por porción
Grasas trans	Grasas poco saludables
Lípidos	Grasas
Valor energético	Energía o calorías que aporta el alimento
Jarabe de glucosa-fructosa	Azúcar añadido o azúcar concentrado

En tercer lugar, los resultados muestran un desarrollo competencial significativo. En las respuestas a la pregunta sobre competencias desarrolladas, el alumnado mencionó de forma recurrente la interpretación de etiquetas, la identificación de ingredientes y aditivos, la comparación de productos, la toma de decisiones saludables, la comunicación clara de conceptos técnicos y la capacidad de explicar información a personas sin formación nutricional. Varias respuestas subrayaron también el cambio de perspectiva: términos que para el alumnado resultaban familiares podían ser difíciles para la población general. Esta toma de conciencia constituye un resultado formativo importante, ya que desplaza la atención desde el conocimiento experto hacia la adecuación al receptor.

El desarrollo de conciencia social y ética profesional fue otro resultado relevante. El alumnado vinculó la mediación terminológica con la justicia comunicativa, la responsabilidad de informar con

claridad, la necesidad de evitar prejuicios hacia personas con menor formación y el compromiso de hacer accesible la educación nutricional. Algunas respuestas señalaron que la falta de comprensión del etiquetado puede afectar especialmente a personas con menos recursos, menor nivel educativo o menor familiaridad con la terminología sanitaria. Esta reflexión conecta directamente con la dimensión de responsabilidad social universitaria del proyecto.

En cuarto lugar, el alumnado identificó la contribución del proyecto a varios ODS. Las respuestas mencionaron mayoritariamente el ODS 3, Salud y bienestar, al considerar que la comprensión del etiquetado puede favorecer hábitos alimentarios más saludables y prevenir enfermedades. También se mencionó el ODS 4, Educación de calidad, por la dimensión formativa y alfabetizadora del proyecto, y el ODS 12, Producción y consumo responsables, por la promoción de decisiones de compra más conscientes. Algunas respuestas hicieron referencia, además, a la reducción de desigualdades, reforzando la conexión entre alfabetización, inclusión y equidad.

En conjunto, los resultados permiten afirmar que la experiencia favoreció tres tipos de aprendizaje: un aprendizaje conceptual, relacionado con la comprensión del etiquetado, la terminología nutricional y los principios de alfabetización en salud; un aprendizaje procedimental, vinculado a la identificación de dificultades terminológicas, la reformulación en lenguaje claro y el diseño de materiales divulgativos.; y un aprendizaje actitudinal, social y ético, centrado en la responsabilidad profesional, la empatía comunicativa y el reconocimiento del lenguaje como posible factor de desigualdad.

4. DISCUSIÓN Y CONCLUSIONES

Los resultados obtenidos constatan la pertinencia de abordar el etiquetado alimentario como un espacio de intervención educativa desde la lingüística aplicada y la alfabetización en salud. En línea con Nutbeam (2000), la experiencia muestra que la alfabetización en salud no puede reducirse a la transmisión de información, sino que debe incluir habilidades de interpretación, evaluación y aplicación. Además, el alumnado desarrolló la capacidad de analizar cómo se comunica ese conocimiento y qué barreras pueden surgir cuando el receptor no comparte el marco conceptual experto.

La identificación de términos como “jarabe de glucosa-fructosa”, “lípidos”, “E-250” o “proteínas de alto valor biológico” confirma que el etiquetado alimentario presenta una densidad terminológica que puede dificultar su comprensión. Este resultado coincide con la concepción de Cabré (1999), según la cual los términos son unidades de conocimiento especializado, y con la caracterización de Gotti (2003) del discurso especializado como un discurso de alta precisión y condensación conceptual. En contextos de comunicación experto-lego, estos rasgos pueden convertirse en obstáculos si no se acompañan de estrategias de mediación.

Uno de los hallazgos más relevantes es que el alumnado no se limitó a identificar términos difíciles, sino que fue capaz de explicitar los criterios de dificultad. Las respuestas apuntaron a la tecnicidad, la falta de familiaridad, la baja transparencia, la presencia de códigos, la ambigüedad y la dificultad numérica. Esta categorización constituye una aportación metodológica importante, ya que permite transformar una percepción intuitiva de dificultad en un análisis más sistemático. Desde este punto de vista, la plantilla de análisis terminológico y el IDT constituyen una aportación metodológica relevante para futuras investigaciones sobre accesibilidad lingüística en salud.

La experiencia también confirma el valor de la mediación terminológica como forma de traducción intralingüística. La reformulación propuesta por Zethsen (2009), necesaria cuando el texto original no resulta accesible para determinados destinatarios, se concretó en la adaptación de términos técnicos a expresiones más cercanas al lenguaje común. Sin embargo, los resultados muestran que la mediación

no consiste meramente en sustituir palabras, sino en contextualizar y conectar conceptos con decisiones cotidianas. Por ejemplo, "grasas saturadas" no se reformula únicamente como "grasas", sino como "grasas que conviene limitar"; "valor energético" no se traduce solo por "energía", sino por "calorías que aporta el alimento.

Desde el punto de vista docente, los resultados son coherentes con la literatura sobre aprendizaje-servicio. La descripción de Eyler y Giles (1999), según la cual esta metodología favorece aprendizajes más significativos porque vincula los contenidos académicos con problemas sociales reales, se concreta en la forma de abordar la terminología como una herramienta para mejorar la comprensión de personas concretas. La actividad no solo reforzó el conocimiento sobre etiquetado, sino que contribuyó a desarrollar una mirada profesional más sensible a la diversidad de destinatarios.

El proyecto también aporta una dimensión relevante de responsabilidad social universitaria. La universidad no es únicamente un espacio de producción de conocimiento, sino agente mediador entre saber experto y ciudadanía. Esta función resulta especialmente importante en ámbitos como la salud, donde la falta de comprensión puede tener consecuencias prácticas para el bienestar. La experiencia y los resultados muestran que la claridad lingüística no es un elemento accesorio, sino una condición básica de equidad. Si la información alimentaria es formalmente pública pero no comprensible, su repercusión social y capacidad de empoderamiento es escasa.

Esta dimensión se refleja también en las respuestas del alumnado, que interpretó la accesibilidad lingüística como un requisito ético de la comunicación sanitaria. Muchos estudiantes afirmaron que el lenguaje puede generar desigualdad cuando no se adapta al receptor, cuando excluye a quienes no poseen formación especializada o cuando obliga a depender de otras personas para interpretar información relevante. Esta conciencia es fundamental en la formación de futuros profesionales de la salud, la educación y la comunicación especializada. En este sentido, el proyecto contribuye a formar profesionales capaces de reconocer que la precisión técnica debe ir acompañada de responsabilidad comunicativa y que el etiquetado alimentario puede funcionar como un eje interdisciplinar para integrar salud, educación, lenguaje y sostenibilidad.

No obstante, el estudio presenta algunas limitaciones. En primer lugar, el número de participantes en la fase de cuestionario fue reducido, por lo que los resultados deben interpretarse como exploratorios. En segundo lugar, aunque se diseñaron pruebas pre- y post- para población destinataria, sería necesario ampliar la muestra comunitaria y aplicar análisis estadísticos más robustos para medir con mayor precisión el alcance de la intervención. En tercer lugar, el IDT se encuentra todavía en una fase inicial y requiere validación, análisis de consistencia interna y contraste con resultados de comprensión.

Estas limitaciones, sin embargo, abren líneas futuras de investigación. En primer lugar, sería conveniente ampliar el corpus de etiquetas y aplicar la plantilla de análisis a distintos grupos de alimentos especialmente dirigidos a infancia, personas mayores o población con necesidades dietéticas específicas. En segundo lugar, se podría validar el IDT mediante codificadores expertos y usuarios legos. En tercer lugar, sería recomendable desarrollar estudios longitudinales para comprobar si la mejora en comprensión se traduce en cambios sostenidos en hábitos de lectura de etiquetas y toma de decisiones alimentarias.

En conclusión, el etiquetado alimentario constituye un espacio privilegiado de alfabetización en salud y de intervención educativa. La experiencia descrita demuestra que el aprendizaje-servicio permite integrar formación académica, compromiso social y transferencia del conocimiento. Asimismo, confirma que la lingüística aplicada puede desempeñar un papel central en la promoción de la salud, al contribuir a que la información especializada sea no solo correcta, sino también comprensible, útil y socialmente inclusiva.

5. REFERENCIAS

Anastasiou, K., Miller, M., y Dickinson, K. (2019) The relationship between food label use and dietary intake in adults: A systematic review. *Appetite, 138*, 280-291. https://doi.org/10.1016/j.appet.2019.03.025

Bringle, R. G. y Hatcher, J. A. (1996). Implementing service learning in higher education. *Journal of Higher Education*, *67*(2), 221-239. https://doi.org/10.1080/00221546.1996.11780257

Cabré, M. T. (1999). *Terminology: Theory, methods, and applications*. John Benjamins Publishing.

Campos, S., Doxey, J. y Hammond, D. (2011). Nutrition labels on pre-packaged foods: A systematic review. *Public Health Nutrition, 14*(8), 1496-1506. https://doi.org/10.1017/S1368980010003290

Eyler, J., y Giles, D. E. (1999). Where's the learning in service-learning? Jossey-Bass.

Gotti, M. (2003). *Specialized discourse: Linguistic features and changing conventions*. Peter Lang.

Kelly, B., Ng, S. H., Carrad, A., y Pettigrew, S. (2024). The potential effectiveness of front-of-pack nutrition labeling for improving population diets. *Annual Review of Nutrition, 44*, 405–440 y 441-470. https://doi.org/10.1146/annurev-nutr-011224-030917

Kloehn, N., Leroy, G., Kauchak, D., Gu, Y., Colina, S., Yuan, N. P. y Revere, D. (2018). Improving consumer understanding of medical text: Development and validation of a new SubSimplify algorithm to automatically generate term explanations in English and Spanish. *Journal of Medical Internet Research, 20*(8), e10779. https://doi.org/10.2196/10779

Malloy-Weir, L., y Cooper, M. (2017). Health literacy, literacy, numeracy and nutrition label understanding and use: A scoping review of the literature. *Journal of Human Nutrition and Dietetics, 30*(3), 309–325. https://doi.org/10.1111/jhn.12428

Mukherjee, P., Leroy, G., Kauchak, D., Navarrete, B. A., Diaz, D. Y., y Colina, S. (2018). The Role of Surface, Semantic and Grammatical Features on Simplification of Spanish Medical Texts: A User Study. *AMIA . Annual Symposium Proceedings. AMIA Symposium*, *2017*, 1322-1331.

Nutbeam, D. (2000). Health literacy as a public health goal: A challenge for contemporary health education and communication strategies into the 21st century. *Health Promotion International, 15*(3), 259–267-. https://doi.org/10.1093/heapro/15.3.259

Nutbeam, D. y Lloyd, J. E. (2021). Understanding and responding to health literacy as a social determinant of health. *Annual Review of Public Health, 42*, 159-173. https://doi.org/10.1146/annurev-publhealth-090419-102529

Organización de las Naciones Unidas. (2015). Transforming our world: the 2030 Agenda for Sustainable Development | Department of Economic and Social Affairs

Pfledderer, C. D., Ranjit, N., Pérez, A., Malkani, R. I., Ferreira De Moraes, A. C., Hunt, E. T., Smith, C. L. y Hoelscher, D. M. (2024). Using the Nutrition Facts Label to Make Food Choices Is Associated with Healthier Eating among 8th and 11th-Grade Students: An Analysis of Statewide Representative Data from the 2019–2020 Texas School Physical Activity and Nutrition Survey. Nutrients, 16(2), 311. https://doi.org/10.3390/nu16020311

Rana, T., Konapur, A., Mishra, P., Kour, M., Rao, A., Arora, M., y Gavaravarapu, S. M. (2025). The impact of educational interventions on food label comprehension and food choices among adolescents: A scoping review. *BMC Public Health, 25*, Article 2682. https://doi.org/10.1186/s12889-025-23940-5

Rothman, R. L., Housam, R., Weiss, H., Davis, D., Gregory, R., Gebretsadik, T., Shintani, A. y Elasy, T. A. (2006). Patient understanding of food labels: The role of literacy and numeracy. *American Journal of Preventive Medicine, 31*(5), 391-398. https://doi.org/10.1016/j.amepre.2006.07.025

Shrestha, A., Cullerton, K., White, K. M., Mays, J. y Sendall, M. (2023). Impact of front-of-pack nutrition labelling in consumer understanding and use across socio-economic status: A systematic review. *Appetite, 187*, 106587. https://doi.org/10.1016/j.appet.2023.106587

Silva, P., Araújo, R., Lopes, F. y Ray, S. (2023). Nutrition and food literacy: Framing the challenges to health communication. *Nutrients, 15*(22), Article 4708. https://doi.org/10.3390/nu15224708

Sørensen, K., Van den Broucke, S., Fullam, J., Doyle, G., Pelikan, J., Slonska, Z. y Brand, H. (2012). Health literacy and public health: A systematic review and integration of definitions and models. *BMC Public Health, 12,* Article 80. https://doi.org/10.1186/1471-2458-12-80

Yildirim, G. y Çakır, M. A. (2025). Assessing food label literacy: Development and validation of a psychometric scale for adults. *Nutrition & Food Science, 55*(8), 1280–1295. https://doi.org/10.1108/NFS-02-2025-0039

Zethsen, K. K. (2009). Intralingual translation: An attempt at description. *Meta, 54*(4), 795–812. https://doi.org/10.7202/038904ar

6. From detection to traceability: A segmented dossier architecture for assessment under generative AI in a History of America course[1]

Antonio Carrasco-Rodríguez

University of Alicante, Spain

ABSTRACT

Generative AI has weakened the evidentiary sufficiency of final written products in unsupervised assessment contexts. This article examines a segmented dossier designed for a fourth-year undergraduate History course on America from colonization to the present. Using a bounded case study supported by document analysis, the study analyzed 26 final dossiers structured through six mandatory blocks: problem delimitation, historical synthesis, verified bibliography, search record, process memory, and AI declaration. The analysis focused on cross-block coherence, process traceability, and the transparency of declared AI use. Results indicate that the architecture redistributed evaluative weight away from the final text alone and toward a linked evidentiary chain. The contribution is threefold: it documents this redistribution in a real course setting; operationalizes it through cross-block coherence, process traceability, and structured AI transparency; and shows that AI declaration becomes most analytically useful when interpreted together with the documentary and reflective components of the task. The case does not establish a universal model, but it indicates that a segmented architecture can make assessment more auditable by shifting evaluative attention toward process, coherence, and structured transparency.

KEYWORDS: generative AI; assessment design; traceability; higher education; history teaching.

1. INTRODUCTION

Generative artificial intelligence (AI) has complicated the evidentiary interpretation of written assessment in unsupervised or weakly supervised settings. In such contexts, the final text may no longer function as a self-sufficient indicator of learning, because its apparent quality is not always a reliable proxy for the student's own cognitive work. This does not invalidate all product-based assessment, but it does weaken models that rely almost exclusively on the submitted text as the main or sole basis for judgment. Recent guidance has accordingly shifted the discussion away from prohibition or detection alone and toward assessment design, with greater emphasis on multiple, contextualized, and better-triangulated forms of evidence rather than exclusive dependence on terminal written output (Bearman et al., 2024; Kizilcec et al., 2024, pp. 1–2; Lodge et al., 2023, pp. 2–3). A related policy-oriented line has framed the educational response to generative AI in terms of regulation, instructional design, and human-centered governance rather than purely reactive control (Chan, 2023, pp. 20–21; Miao & Holmes, 2023, pp. 2, 18).

This is no longer a marginal issue. A review of guidance issued by highly ranked universities showed that generative AI was already being treated as an assessment and academic-integrity issue

1 This work was supported by the University Teaching Research Networks Programme of the Institute of Education Sciences at the University of Alicante, 2025 call. Ref.: 6171.

requiring new or revised policies, recommendations, and communication frameworks (Moorhouse et al., 2023). The rapid growth of AI-related research in higher education points in the same direction: the topic has moved beyond isolated experimentation and become a visible area of pedagogical discussion (Crompton & Burke, 2023; Xia et al., 2024). That broader context matters, but only as far as it helps define the problem more precisely. The issue is not AI in higher education in general, nor the comparative quality of institutional guidelines, nor the performance of automated detectors. The question is narrower and more concrete: what kind of evaluative architecture remains defensible when the final written product no longer provides sufficient evidence on its own?

The case examined here addresses that question through a segmented dossier designed for a fourth-year undergraduate History course on America from colonization to the present. Students were required to produce a thematic dossier organized into six mandatory blocks: problem delimitation, historical argumentative synthesis, verified bibliography, search and verification record, reasoned process memory, and an explicit declaration of generative AI use. The design did not aim to prove authorship in any absolute sense. Its purpose was to distribute evaluative judgment across a broader evidentiary chain by making problem definition, bibliographic decisions, process reconstruction, and AI transparency visible alongside the historical argument itself.

The study therefore examines a bounded teaching case with secondary methodological yield. The central hypothesis is that, under conditions shaped by generative AI, academic integrity may be addressed more defensibly through pedagogical design than through primary reliance on technological surveillance. More specifically, it asks whether a segmented architecture can shift evaluative judgment toward three linked dimensions: process traceability, cross-block coherence, and structured transparency about AI use. The article makes three bounded contributions. It documents how a segmented dossier redistributed evaluative weight away from the final text alone in a real course setting; it operationalizes that redistribution through cross-block coherence, process traceability, and structured AI transparency; and it shows that AI declaration becomes analytically useful only when interpreted together with the documentary and reflective components of the task. It does not seek to validate a universal model, establish causal superiority over alternative formats, or psychometrically validate the rubric used in the case.

2. METHOD

2.1. Context and corpus

The case is drawn from a fourth-year compulsory undergraduate course, *America: History from Colonization to the Present*, taught in the History degree at the University of Alicante, Spain. It is delimited to the practical assignment corresponding to the thematic blocks on pre-Columbian and Hispanic colonial history, in which students were required to produce a problem-oriented dossier based on a thematic axis taken from the course manual and related teaching materials. The assignment was originally designed as pair work, carried an estimated 30 hours of autonomous work, and required a total length of 5,000 to 7,000 words.

Its formal structure was fixed in advance through six mandatory blocks, labeled A to F: A, problem delimitation and selection decisions; B, historical argumentative synthesis; C, verified bibliography; D, search and verification record; E, reasoned process memory; and F, explicit declaration of generative AI use. These components defined the dossier as a segmented written assignment rather than as a continuous essay.

The analytic corpus consisted of 26 final dossiers, treated as individual cases. The unit of analysis was one dossier per case. All dossiers contained the six required blocks. Although the assignment had been designed primarily for pair work, the final corpus included 23 paired dossiers and 3 individual submissions. This difference is reported descriptively only and is not used for comparative analysis between submission formats.

The study is therefore neither a survey of AI use in higher education nor a perception study. It is a bounded teaching case centered on a segmented written assignment in a specific disciplinary context. The aim is not statistical representativeness, but analytic understanding of how a concrete evaluative architecture operated in practice and what kinds of visible evidence it generated across the corpus. Following Yin (2018, p. 11), the study is best understood as a case study of a contemporary phenomenon in context. Any broader value of the findings remains a matter of analytic rather than statistical generalization (Yin, 2018, p. 18).

2.2. Instruments

The analysis relied on three linked documentary instruments. The first was the assignment guide for the practical task, which defined the dossier's formal structure, fixed the relative weight of each block and specified expected lengths, required contents, and permitted and non-permitted uses of generative AI. The second was the analytic rubric. Its function was not limited to scoring: it translated the dossier structure into observable criteria, thresholds, and penalties, making the evaluative logic of the design explicit and comparable across cases. For example, it specified the conditions under which a block became non-evaluable, the minimum requirements for verified references in Block C, the need to identify individual contribution in pair work in Block E, and the requirement to declare both phases of AI use and non-delegated tasks in Block F. The rubric was not treated as a psychometrically validated instrument, but as a structured reading grid. This use is consistent with review literature that treats rubrics as tools for making criteria explicit and supporting formative functions (Panadero & Jönsson, 2013; Reddy & Andrade, 2010).

The third instrument was the corpus of 26 final dossiers itself, read as documentary evidence of how the evaluative design operated in practice. Within this configuration, the assignment guide and rubric defined the intended architecture of the task, while the dossiers documented its realized forms. In this sense, the study follows document analysis as a systematic procedure for reviewing and evaluating documents (Bowen, 2009, p. 27).

2.3. Procedure

The procedure combined functional reading of the dossiers, cross-block analysis, iterative coding, and light descriptive quantification. Each dossier was read as a segmented evidentiary chain rather than as a continuous text. The analysis began with a block-by-block reading of the six mandatory components, treating each as a distinct producer of evidence within the assignment: A as problem delimitation, B as the argumentative core, C as verified bibliographic support, D as the record of search and selection, E as reasoned reconstruction of process, and F as explicit declaration of generative AI use and non-delegated tasks. This first pass established the functional role of each block in relation to the design defined by the assignment guide and rubric.

The analysis then moved from individual blocks to the relations between them. Coherence was examined across the dossier as a linked chain rather than as a sum of components, with particular attention to the relations A–B, B–C, C–D, D–E, and E–F. Traceability was examined at the level of the whole case by asking whether the dossier made the learning process reconstructable through its problem definition, bibliographic selection, search and verification record, process memory, and AI declaration. AI transparency was analyzed through the declared use reported in F and interpreted in relation to E. For reporting transparency, the operational criteria used to translate these readings into the categories reported in the Results are specified in Table 1.

The internal analytic matrix contained one row per dossier and separate fields for the five cross-block relations, global coherence, process traceability, AI transparency, and final typology. Coding was iterative: provisional classifications were checked against block-level notes and against the dossier as a whole. Borderline cases were reread and resolved by the relative density and consistency of the links across A–F rather than by mechanical score thresholds. The matrix was an auxiliary coding device, not an independent source of evidence or a hidden grading instrument.

Table 1. Operational coding framework

Dimension	Evidence examined	Operational category rules
Global coherence	Continuity across A–B, B–C, C–D, D–E, and E–F.	High: the chain is sustained across most or all links and no major contradiction disrupts it. Medium: the dossier remains reconstructable, but one or two links are thin or only partly integrated. Low: several links are weak or disconnected, so the formal presence of A–F does not produce a functional chain.
Process traceability	Visibility of problem continuity, documentary search and selection, reflective revision, and situated AI declaration.	High: the process is reconstructable across these components. Medium: the process is visible overall, but one or two components remain mainly descriptive, organizational, or weakly connected. Low: the blocks are present but do not make the process sufficiently reconstructable.
AI transparency	Block F read together with the process memory in E.	Clear and situated: tool, stages, limits, and non-delegated tasks are specified and consistent with E. Minimal but sufficient: formal disclosure is present with limited contextual detail. Transparent but tensioned: broader uses are reported candidly and create an interpretive boundary issue. Vague or weakly integrated: stages, limits, or relation to E remain unclear.
Final typology	Combined pattern of coherence, traceability, and AI transparency.	Type 1: articulated and transparent. Type 2: strong argument and visible process with uneven integration. Type 3: functionally adequate with stable medium traceability. Type 4: narrated process with incomplete reflective integration. Type 5: fragile articulation of the architecture.

The corpus was finally summarized through limited descriptive quantification. Frequencies were used to report the distribution of levels of global coherence, process traceability, AI transparency regimes, and final typologies. This step was purely descriptive and did not involve sta-

tistical inference, causal testing, or psychometric validation of the rubric. The procedure supports an analytic interpretation of how the evaluative architecture redistributed visible evidence across process, coherence, and transparency within this case, but it does not support representative extrapolation beyond it.

2.4. Ethical and reporting safeguards

The study analyzed course-generated documents at the dossier level. The revised report uses only aggregate counts and composite paraphrases of recurring patterns; it reproduces no student names, marks, demographic attributes, or verbatim passages. The illustrative profiles in Table 3 are analytical composites rather than quotations or case reconstructions attributable to a particular submission. AI declarations are treated as self-reports, and the analysis does not infer undeclared use. These reporting restrictions reduce identification risk in a small cohort and keep the analysis focused on the assessment architecture rather than on individual performance.

3. RESULTS

3.1. Distributed evidentiary functions

The corpus did not operate as a loose aggregation of six formal blocks. Across the 26 dossiers, the segmented architecture generated differentiated evidentiary functions that were not distributed evenly across the task. In most cases, A functioned as the opening delimitation of a problem, B concentrated most of the argumentative weight, C provided bibliographic support with uneven levels of integration, D documented search and selection, E made process and decision-making visible, and F externalized declared AI use and non-delegated tasks. The design therefore worked as a chain of linked forms of evidence rather than as a single terminal text.

A usually functioned as an opening block for the problem rather than as a thematic preface, although in a non-trivial minority it drifted toward general introduction. B was clearly the strongest block in the corpus: most dossiers used it to formulate a thesis, organize historical tensions, and close with some interpretation rather than simple summary. C was more heterogeneous and emerged as the main formal bottleneck, not because it was absent, but because bibliographic sufficiency did not always translate into equally integrated argumentative support. D usually confirmed that a real search process had taken place. E often turned that process into a usable account of decisions and difficulties. F added an empirical layer that a continuous essay would not normally provide: an explicit account of declared AI use.

3.2. Cross-block coherence

Cross-block coherence mattered more than the mere presence of A–F. The strongest relation in the corpus was A–B, that is, between problem delimitation and argumentative synthesis. In most cases, the problem introduced in A remained recognizable in B, even when B widened the historical context or enlarged the argumentative scale. The most unstable link was B–C. Where the dossier worked best, bibliographic notes were visibly tied to specific tensions in the argument. Where it worked less well, C remained numerically acceptable but too generic, too mixed in the quality of

its supports, or too weakly anchored to the thesis of B. The C–D relation was generally positive: D usually explained where sources were sought, what difficulties arose, and why some materials were retained or discarded. D–E was one of the most functional relations in the corpus, especially when E moved beyond chronology and task distribution to reconstruct conceptual difficulties, changes of focus, and revisions of the initial plan. E–F, by contrast, was the most variable relation: some dossiers integrated process memory and AI declaration closely, while others left F as a relatively separate final clause.

At the global level, the corpus did not cluster at the lower end of coherence. Twelve dossiers were coded as high in global coherence, twelve as medium, and two as low. That coding was analytic and governed, not automatic, and it should not be read as a hidden score. Even so, the distribution is clear enough to support a cautious conclusion: the segmented architecture usually produced dossiers that were functionally connected across blocks, though not all with the same density of integration. The main recurrent pressures were also clear. B continued to absorb too much of the dossier's cognitive and textual weight. C remained the weakest link in the chain of support. E–F did not always converge with the same precision.

3.3. Traceability of process

Traceability followed a similar pattern. Here, it did not refer to overall historiographical quality, but to the degree to which the dossier made the learning process reconstructable through a visible problem, cross-block continuity, documented search, metacognitive reflection, and transparent AI declaration. On that basis, 11 cases were coded as high in traceability, 13 as medium, and 2 as low. High-traceability dossiers usually shared four traits: a clear problem sustained from A into B; a documentary chain that made C and D intelligible together; an E block that reconstructed decisions rather than merely listing tasks; and an F block integrated into that sequence rather than appended to it. Medium-traceability cases did not amount to weak cases. They usually made the process visible in general terms, but one or two links remained thinner, most often because the bibliography was less integrated, E remained more organizational than metacognitive, or F declared use correctly without adding much interpretive density. Low-traceability cases preserved the required blocks but did not convert them into a sufficiently strong chain of visible process.

The most functional interaction for traceability did not lie in C or D taken separately, but in the combined relation among C, D, and E. D frequently documented real search activity by naming repositories, databases, keywords, an excess of materials, access problems, or the need for selection. E strengthened that record when it explained how those documentary constraints reshaped the problem, narrowed the topic, reorganized B, or redistributed tasks. The weaker dossiers were not necessarily those with no process, but those in which the process remained only partially visible because E stayed at the level of logistics or because D described repositories or sources consulted without clarifying how they shaped the final bibliographic apparatus. Medium traceability therefore indicates partial visibility of process within the evidentiary logic of the design, not general weakness of the work.

3.4. Transparency of AI use

Block F yielded one of the clearest results in the corpus. AI declaration did not function as a binary yes/no box, but as a stratified layer of evidence. All 26 dossiers included an explicit declaration in F, and none was built around a claim of absolute non-use. The analysis identified four transparency regimes: clear and situated declaration in 8 cases; minimal but sufficient declaration in 6; transparent but tensioned declaration in 10; and vague or weakly integrated declaration in 2. The strongest pattern named the tool, delimited specific stages of use, stated what had not been delegated, and remained consistent with E. The minimal pattern complied formally, but with less density and weaker integration. The most revealing group was the tensioned one. These declarations were often frank and therefore empirically useful, but they described broader uses such as topic suggestion, bibliographic orientation, source screening or summarization, reformulation of substantive passages, or support in writing the process memory. This does not in itself indicate rule-breaking, since the assignment allowed declared AI use provided that key historiographical decisions were not delegated. It does, however, expose a zone of interpretive tension between legitimate support and more expansive forms of dependence. The weakest pattern was not silence, but low precision about stages, limits, and non-delegated tasks, together with a poorer fit between F and E.

The relation between AI transparency and traceability was not mechanical, but it appeared substantively relevant within the corpus. Clear and situated declarations concentrated in the high-traceability segment, or in the most robust part of the corpus. Minimal declarations tended to coexist with medium traceability: they did not damage the architecture, but they did little to strengthen it. Tensioned declarations were more complex. In some cases, they increased process visibility precisely because they were candid about broader uses. In others, they exposed a weaker fit between declared support and the documented sequence of work. Vague declarations tended to weaken traceability because they left F closer to a formal requirement than to an analytic component of the dossier. The key point is not that transparency resolved interpretation, but that it made interpretation possible on empirical grounds that a conventional essay would have concealed.

3.5. Typological synthesis

The typological synthesis confirms that the corpus cannot be reduced to a simple contrast between success and failure. Five operational types were identified from the combination of cross-block coherence, global traceability, and integration of AI transparency: Type 1, articulated and transparent dossier, 8 cases (30.8%); Type 2, strong argument with visible process but uneven integration, 7 cases (26.9%); Type 3, functionally adequate dossier with stable medium traceability, 6 cases (23.1%); Type 4, narrated process with incomplete reflective integration, 3 cases (11.5%); and Type 5, fragile articulation of the architecture, 2 cases (7.7%). Types 1, 2, and 3 account for 21 of the 26 dossiers (80.8%). The characteristic performance of the design therefore lies less in exceptional cases than in a broad central band of functional and relatively traceable dossiers. The typology is interpretive rather than automatic, but it clarifies an important point: the strongest dossiers did not depend only on a good B block. They depended on D, E, and F making the process legible.

Table 2 consolidates the descriptive distributions reported across the four analytical outputs. Percentages are calculated on the full corpus of 26 dossiers and are reported only to make the empirical pattern more visible.

Table 2. Descriptive distribution of analytical categories (N = 26)

Analytical output	Category	n	%
Global coherence	High	12	46.2
Global coherence	Medium	12	46.2
Global coherence	Low	2	7.7
Process traceability	High	11	42.3
Process traceability	Medium	13	50.0
Process traceability	Low	2	7.7
AI transparency	Clear and situated	8	30.8
AI transparency	Minimal but sufficient	6	23.1
AI transparency	Transparent but tensioned	10	38.5
AI transparency	Vague or weakly integrated	2	7.7
Final typology	Type 1	8	30.8
Final typology	Type 2	7	26.9
Final typology	Type 3	6	23.1
Final typology	Type 4	3	11.5

At the upper end of the corpus, the strongest dossiers combined high coherence, high traceability, and a clearly situated AI declaration. Their strength did not rest on the argumentative core alone, but on the visibility of the chain linking problem delimitation, bibliographic support, process reconstruction, and declared AI use. At the weaker end, by contrast, the formal presence of A–F did not always translate into a robust evidentiary sequence. The most fragile dossiers tended to preserve the required structure while leaving thinner links between documentation, reflection, and AI transparency. Between those poles lay a broad middle range of functionally adequate dossiers in which the architecture worked, though with uneven density of integration across blocks.

To make the empirical basis more visible without reproducing identifiable student prose, Table 3 presents composite anonymized profiles assembled from recurring patterns across multiple dossiers. They are not verbatim extracts and should not be read as individual case histories.

Table 3. Composite anonymized examples of recurrent dossier patterns

Composite profile	Recurring cross-block pattern	Interpretive significance
Articulated and transparent	A defines a problem sustained in B; C is tied to specific argumentative tensions; D records repositories, keywords, selection, and discarded options; E explains narrowing or revision; F specifies the tool, stages, limits, and non-delegated tasks in a way consistent with E.	High coherence and high traceability. The final argument remains central, but its production is legible across the dossier.

Composite profile	Recurring cross-block pattern	Interpretive significance
Strong product, uneven integration	B is analytically strong; C is sufficient but comparatively generic; D documents a real search; E remains partly organizational; F complies but adds little contextual detail.	The quality of the final product is not identical to the visibility of process. Traceability remains medium despite a strong argumentative core.
Transparent but tensioned	F candidly reports broader support such as topic suggestion, bibliographic orientation, source screening or summarization, reformulation, or help with the process memory; E integrates those uses only partially.	Disclosure increases visibility but does not by itself resolve the boundary between legitimate support and substantive dependence.
Fragile articulation	A and B are only loosely connected; C is generic; D lists sources or repositories without a clear selection logic; E emphasizes task distribution; F is vague or weakly linked to E.	The six blocks are formally present, but they do not become a robust evidentiary chain.

The overall pattern is clear, though limited to a single case and to governed analytic coding. The segmented architecture did not eliminate the centrality of the main argumentative block, and it did not guarantee uniformly rich reflection in every dossier. It did, however, oblige students to produce additional traces of search, selection, revision, reflection, and declared AI use. Within those limits, the dossier made additional stages of the work visible and redistributed evidentiary attention away from the final product alone.

4. DISCUSSION AND CONCLUSIONS

The main value of the case lies less in detecting AI-supported writing than in redefining what counts as evaluative evidence. In that sense, the dossier belongs to a longer reformist line in higher education assessment that repositions assessment as an integral part of curriculum and pedagogy, rather than merely as a means of ascertaining whether outcomes have been achieved (Boud & Soler, 2016, pp. 407, 411). Under generative AI, that shift has become more pressing. Recent guidance has argued that robust assessment now depends on multiple, inclusive, and contextualized approaches, and on judgments built from more than a final product (Lodge et al., 2023, pp. 2–3). The pattern identified here fits that development. The design did not displace the historical argument from the center of the task, since Block B remained the strongest and heaviest component in most dossiers. It did, however, reduce the evidentiary monopoly of the final prose by requiring visible traces of delimitation, source selection, search history, reflective revision, and declared AI use.

That shift matters because the problem raised by generative AI is not simply whether a text sounds suspicious, but whether the available evidence still supports a defensible judgment about learning. In this case, the segmented architecture produced an environment in which more elements of the learning process became visible within the task itself. The strongest dossiers were not only better argued. Their internal sequence was also more legible: the initial problem remained recognizable in the synthesis, bibliography could be tied to argumentative choices, the search record documented constraints and decisions, and the process memory clarified how the dossier had been built. This did

not replace product with process. The historical synthesis still mattered decisively. What changed was the distribution of validity across a chain of linked evidentiary components. More precisely, the case suggests that, for tasks of this kind under generative AI, evaluative judgment is more defensible when it considers not only the final prose but also the coherence of a documentary chain linking problem definition, source work, process reconstruction, and declared AI use. That is a narrower and more defensible claim than saying that product-based assessment has become useless.

Block F sharpens the point further. The corpus does not support a binary opposition between transparent and non-transparent use, nor does it justify speculation about undeclared uses beyond what students reported. It does show that explicit declaration produces analytically useful variation. Some dossiers stated AI use with clear limits and strong compatibility with the rest of the dossier. Others complied more minimally. Others were transparent but exposed tensions between broader declared support and the documented sequence of work. That pattern connects the case to recent frameworks that seek to clarify permitted and non-permitted uses of generative AI through explicit levels or regimes of disclosure, even though the dossier should not be treated as equivalent to the AIAS model itself (Perkins et al., 2024, pp. 1, 6-9). It also converges with the more applied point made by Furze et al. (2024, pp. 38, 44-47), who show that such frameworks can be implemented as operational assessment guidance rather than left at the level of general principle. In this case, transparency worked best when it was not isolated in F, but aligned with the reflective and documentary evidence provided in D and E.

Declarative traceability has clear limits. An AI declaration is retrospective, self-reported, and potentially strategic; it cannot authenticate authorship or establish the completeness of disclosure. Cross-block coherence can also be reconstructed after the fact, including with AI support. The architecture therefore increases inspectability rather than proof. Its value lies in generating multiple points for evaluative judgment, not in converting declarations into verification. This distinction is especially important in tensioned cases, where candor may reveal extensive use without determining on its own whether that use crossed a pedagogical boundary.

A more limited dialogue with work on authentic assessment is also useful here. Villarroel et al. (2018, pp. 843–845) matter not because the dossier fully exemplifies authentic assessment in any strong or exhaustive sense, but because they frame course design through the linked dimensions of realism, cognitive challenge, and evaluative judgment. The dossier shares part of that logic. Students were not asked only to submit a polished answer. They had to justify a problem, build a historical argument, verify a bibliography, explain search choices, reconstruct their process, and declare the place of AI within that sequence. At the same time, recent work has warned that innovative or authentic designs also depend on clarity, support, and psychological safety (Wake et al., 2024, pp. 511–512, 518). That caution matters here. A demanding architecture can easily become punitive or confusing if expectations remain opaque. In this case, the explicit instructions, differentiated blocks, visible criteria, and declared AI rules appear to have worked as supports rather than as hidden thresholds, although the study includes no perceptual data that would justify a stronger claim about student experience.

Scalability is a second limitation. The assignment required 5,000 to 7,000 words and an estimated 30 hours of student work, but the study did not record marking time or compare workload with a conventional essay. Reading relations across six blocks is likely to add evaluative work even if the fixed structure and rubric make that work more systematic. Transfer to larger cohorts would therefore depend on design choices such as standardized templates, bounded block lengths, staged checkpoints, selective audit of documentary records, calibrated rubric descriptors, and adequate workload alloca-

tion. These are plausible implementation strategies, not effects demonstrated by this case. The model should not be presented as cost-neutral or automatically scalable.

Three bounded conclusions follow. First, the case documents in a real course setting a redistribution of evaluative weight away from the final text alone and toward a broader evidentiary chain. Second, that redistribution becomes analytically workable when it is read through cross-block coherence, process traceability, and structured AI transparency. Third, AI declaration is not a marginal add-on: it becomes most analytically useful when it is interpreted together with the dossier's documentary and reflective components. None of these points demonstrates universal superiority over other assessment formats, and none psychometrically validates the rubric. They support a more limited conclusion: under the conditions documented here, a segmented architecture can distribute evaluative judgment across a broader and more visible set of traces than a continuous text alone.

The case therefore does not offer a general solution to generative AI in assessment. It provides a plausible and documentable response to one specific problem: how to sustain defensible evaluative judgment when the final written product can no longer be treated as self-sufficient evidence of learning. Its value lies in showing that assessment can remain demanding without retreating either to blanket prohibition or to a fiction of textual self-sufficiency. The relevant shift is from detecting the finished text to reading the conditions under which that text was produced. Within the limits of a single course, self-reported AI declarations, composite documentary analysis, and unmeasured marking workload, the segmented dossier made assessment more auditable, made declared AI use more interpretable, and made coherence across stages part of what counted as evidence. Its wider transfer will depend on discipline, task design, cohort size, and local conditions of implementation.

5. REFERENCES

Bearman, M., Tai, J., Dawson, P., Boud, D., & Ajjawi, R. (2024). Developing evaluative judgement for a time of generative artificial intelligence. *Assessment & Evaluation in Higher Education*, *49*(6), 893–905. https://doi.org/10.1080/02602938.2024.2335321

Boud, D., & Soler, R. (2016). Sustainable assessment revisited. *Assessment & Evaluation in Higher Education*, *41*(3), 400–413. https://doi.org/10.1080/02602938.2015.1018133

Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal, 9*(2), 27–40. https://doi.org/10.3316/QRJ0902027

Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education, 20*, Article 38. https://doi.org/10.1186/s41239-023-00408-3

Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. *International Journal of Educational Technology in Higher Education, 20*, 22. https://doi.org/10.1186/s41239-023-00392-8

Furze, L., Perkins, M., Roe, J., & MacVaugh, J. (2024). The AI Assessment Scale (AIAS) in action: A pilot implementation of GenAI-supported assessment. *Australasian Journal of Educational Technology, 40*(4), 38–55. https://doi.org/10.14742/ajet.9434

Kizilcec, R. F., Huber, E., Papanastasiou, E. C., Cram, A., Makridis, C. A., Smolansky, A., Zeivots, S., & Raduescu, C. (2024). Perceived impact of generative AI on assessments: Comparing educator and student perspectives in Australia, Cyprus, and the United States. *Computers and Education: Artificial Intelligence, 7*, 100269. https://doi.org/10.1016/j.caeai.2024.100269

Lodge, J. M., Howard, S., Bearman, M., Dawson, P., & Associates. (2023). *Assessment reform for the age of Artificial Intelligence*. Tertiary Education Quality and Standards Agency. https://www.teqsa.gov.au/sites/default/files/2023-09/assessment-reform-age-artificial-intelligence-discussion-paper.pdf

Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO. https://doi.org/10.54675/EWZM9535

Moorhouse, B. L., Yeo, M. A., & Wan, Y. (2023). Generative AI tools and assessment: Guidelines of the world's top-ranking universities. *Computers and Education Open, 5*, 100151. https://doi.org/10.1016/j.caeo.2023.100151

Panadero, E., & Jönsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review, 9*, 129–144. https://doi.org/10.1016/j.edurev.2013.01.002

Perkins, M., Furze, L., Roe, J., & MacVaugh, J. (2024). The Artificial Intelligence Assessment Scale (AIAS): A framework for ethical integration of generative AI in educational assessment. *Journal of University Teaching and Learning Practice, 21*(6). https://doi.org/10.53761/q3azde36

Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education, 35*(4), 435–448. https://doi.org/10.1080/02602930902862859

Villarroel, V., Bloxham, S., Bruna, D., Bruna, C., & Herrera-Seda, C. (2018). Authentic assessment: Creating a blueprint for course design. *Assessment & Evaluation in Higher Education, 43*(5), 840–854. https://doi.org/10.1080/02602938.2017.1412396

Wake, S., Pownall, M., Harris, R., & Birtill, P. (2024). Balancing pedagogical innovation with psychological safety? Student perceptions of authentic assessment. *Assessment & Evaluation in Higher Education, 49*(4), 511–522. https://doi.org/10.1080/02602938.2023.2275519

Xia, Q., Weng, X., Ouyang, F., Lin, T. J., & Chiu, T. K. F. (2024). A scoping review on how generative artificial intelligence transforms assessment in higher education. *International Journal of Educational Technology in Higher Education, 21*, 40. https://doi.org/10.1186/s41239-024-00468-z

Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). SAGE.

7. El Seminario de lecturas como escenario de formación teórico-práctica y reflexividad crítica en Trabajo Social[1]

Mercedes Cuenca Silvestre[1], Javier Ferrer Aracil[1], Belén Parra Ramajo[2]

[1]*Universidad de Alicante*
[2]*Universidad de Barcelona*

RESUMEN

Atendiendo a la necesidad de promover la reflexividad crítica en estudiantes de Trabajo Social, se lleva a cabo un estudio exploratorio y descriptivo en el marco de un proyecto de innovación docente. Para ello, se implementan tres Seminarios de lecturas con 70 estudiantes del Grado en Trabajo Social de la Universidad de Alicante durante el primer semestre del curso 2025-2026, en la asignatura obligatoria Modelos y métodos del Trabajo Social I. Se realiza una evaluación inicial para analizar la percepción del alumnado, empleando para ello un cuestionario estructurado en dos secciones: datos sociodemográficos y una escala tipo Likert de percepción de logro competencial diseñada ad hoc para este estudio. Los resultados muestran una valoración global positiva y equilibrada, con niveles medio-altos de percepción de logro. La lectura individual es la dimensión más apreciada, seguida de la aplicabilidad práctica de los contenidos, mientras que la reflexión compartida presenta mayor variabilidad, aunque con una percepción favorable. Las correlaciones significativas entre las tres fases revelan que los componentes del Seminario se refuerzan mutuamente, resaltando la relevancia del trabajo grupal para la comprensión aplicada y la conexión entre análisis autónomo y transferencia a situaciones reales. Como conclusión, el Seminario de lecturas se consolida como una estrategia educativa clave para fortalecer la reflexividad crítica en la formación en Trabajo Social.

PALABRAS CLAVE: Trabajo Social, Reflexividad crítica, Seminario de lecturas, Innovación docente.

1. INTRODUCCIÓN

La formación en Trabajo Social debe adecuarse a la complejidad de sociedades atravesadas por profundas desigualdades sociales que afectan a muchas de las personas con las que intervienen los y las profesionales. Ante este escenario, se hace necesario en la educación superior fortalecer competencias como la reflexividad crítica, capaz de favorecer una comprensión más consciente y contextualizada de la práctica profesional. En este marco surge el proyecto de innovación docente "La reflexividad crítica aplicada a la enseñanza del Trabajo Social", orientado a promover el desarrollo de dicha competencia en la formación universitaria.

Los contextos sociales de intervención requieren profesionales capaces de integrar conocimientos técnicos con análisis, deliberación y construcción conjunta de respuestas con las personas, los grupos y las comunidades involucradas, promoviendo valores como la justicia social y el respeto por la diversidad (Federación Internacional de Trabajadores Sociales & Asociación Internacional de Escuelas de Trabajo Social, 2020; Foote, 2013). Para responder a estos retos, la formación

1 El presente trabajo ha contado con una ayuda del Programa de Redes de investigación en docencia universitaria del Instituto de Ciencias de la Educación de la Universidad de Alicante (Convocatoria 2025). Ref.: 6354.

incorpora innovaciones en los procesos de enseñanza-aprendizaje orientadas al desarrollo de la reflexividad crítica, entendida como el proceso que permite identificar y cuestionar supuestos personales, valores y creencias que, a menudo de manera implícita, guían cómo se percibe y se actúa en las intervenciones sociales, de modo que esta conciencia permita transformar tanto la comprensión como la acción profesional (Fook & Gardner, 2007; Sheppard, 1998).

Desde esta perspectiva, identificar estrategias que articulen la teoría con la práctica resulta imprescindible para una formación que fortalezca la comprensión conceptual y la construcción de competencias profesionales críticas y contextuales. La literatura internacional sobre formación en Trabajo Social señala que la mera transmisión de conocimientos teóricos resulta insuficiente si no se relaciona con experiencias de intervención concretas, pues esa articulación facilita aprendizajes significativos y prepara al alumnado para enfrentar situaciones reales de manera reflexiva y fundamentada (Kinni, 2020). En el contexto español, investigaciones recientes han mostrado que la incorporación de metodologías activas y recursos didácticos que conectan explícitamente saberes teóricos con prácticas profesionales contribuye a que el estudiantado comprenda mejor los contenidos curriculares y valore la importancia de los modelos teóricos como guía de intervención (Medina et al., 2024).

En este sentido, los Seminarios de lectura constituyen un ejemplo de innovación pedagógica que integra enfoques de aprendizaje experiencial, colaborativo y crítico. Encuentran su base en la perspectiva sociocultural de Vygotsky (2012), quien sostiene que el pensamiento se desarrolla mediante la interacción social mediada por el lenguaje; la Teoría de la acción comunicativa de Habermas (1987), que resalta la importancia del diálogo igualitario y argumentativo; y la concepción de educación dialógica de Paulo Freire (2012, 2017), que enfatiza la reflexión crítica y transformadora sobre la realidad.

Este enfoque del acto comunicativo conecta directamente con el Trabajo Social, profesión en la que las intervenciones con personas, grupos y comunidades requieren reconocer las perspectivas, experiencias y voces de quienes viven la realidad social que se intenta transformar. Precisamente, una praxis profesional reflexiva en Trabajo Social implica desplazar relaciones basadas en el poder hacia intercambios fundamentados en argumentos válidos y comprensibles, en los que todas las partes se reconocen mutuamente y se respeta el conocimiento que cada quien aporta al proceso de intervención.

De manera similar, los Seminarios de Lectura reproducen este tipo de comunicación dialógica y crítica que posteriormente se empleará durante la práctica profesional. Así, Loza (2004) descompone la construcción del conocimiento de forma progresiva: primero a nivel individual, en el diálogo entre autor/a y lector/a, y luego de forma colectiva, mediante la interacción con otros/as estudiantes que permite enriquecer y profundizar la comprensión de la obra. Hennessy et al. (2016), por su parte, describen un marco metodológico para analizar cómo interactúan las personas cuando dialogan en contextos educativos, identificando aquellas formas de interacción que favorecen el pensamiento crítico y el aprendizaje. Finalmente, experiencias previas muestran cómo la la lectura guiada y reflexionada de forma colaborativa contribuye a un aprendizaje más profundo que puede trasladarse al ejercicio profesional (Afdal et al., 2023; Ament et al., 2025). Combinando estas perspectivas (Figura 1), en la propuesta se reconocen tres fases: una fase individual de acercamiento al texto; una fase interpersonal de diálogo y discusión; y una fase de aplicación a la práctica, donde los aprendizajes se conectan con experiencias concretas y significativas.



Figura 1. Proceso del Seminario de lecturas

Fuente: Elaboración propia a partir de Afdal et al. (2023); Ament et al. (2025); Hennessy et al. (2016) y Loza (2004)

Atendiendo a la necesidad de generar estrategias educativas que permitan desarrollar la reflexividad crítica entre estudiantes de Trabajo Social, se diseña un estudio exploratorio y descriptivo que ofrece información sobre el proceso y los resultados de los Seminarios de lecturas. Se parte de la premisa de que comprender cómo los y las estudiantes perciben y experimentan su aprendizaje permite ajustar las prácticas pedagógicas a sus necesidades, lo que repercute en sus resultados formativos. A partir de este propósito se definen los siguientes objetivos:

(1) Explorar la contribución de los Seminarios de lectura al desarrollo del pensamiento crítico y autónomo en el alumnado del Grado en Trabajo Social.

(2) Analizar la percepción del alumnado sobre cómo el Seminario de lecturas influye en su aprendizaje.

En síntesis, este estudio amplía el conocimiento sobre estrategias formativas centradas en la reflexividad crítica en educación superior y aporta evidencia empírica que puede informar el diseño de futuras propuestas pedagógicas orientadas a fortalecer esta competencia en la formación en Trabajo Social.

2. MÉTODO

El estudio adopta un diseño cuantitativo, de corte transversal y con alcance descriptivo-correlacional. Este tipo de diseño resulta adecuado cuando se pretende caracterizar percepciones, valoraciones o motivaciones en un momento determinado y explorar posibles asociaciones con variables sociodemográficas, sin establecer relaciones de causalidad (Fortune & Reid, 1999; Hernández et al., 2006).

2.1. Descripción del contexto y de los participantes

El Seminario de lecturas se lleva a cabo en el primer semestre del año académico 2025-2026. Más concretamente, con estudiantes de los grupos de mañana de la asignatura obligatoria Modelos y MÉTODOs del Trabajo Social I, impartido en 2º curso del Grado en Trabajo Social de la Universidad de Alicante.

La muestra está integrada por 70 estudiantes que participan en el Seminario y completan el cuestionario de evaluación. En cuanto a su perfil, la mayoría son mujeres (91,4%). La edad media es de 21,89 años (DT=5,56). Respecto a la vía de acceso a la universidad, predomina el acceso desde Bachillerato (67,1%), seguido del Ciclo Formativo de Grado Superior (24,3%). En relación con la situación laboral, algo más de la mitad afirma no trabajar en el momento de la recogida de datos (54,3%), si bien una proporción relevante compatibiliza los estudios con un empleo a tiempo parcial (34,3%). Por su parte, la mayoría señala vivir con su familia (80,0%), mientras que el resto convive principalmente con compañeros o compañeras. Asimismo, el 62,7% del alumnado manifiesta haber sufrido alguna experiencia de discriminación, frente al 37,3% que asegura no haberla experimentado. Respecto al contacto previo con el sistema de protección social, una minoría (14,5%) afirma haber sido persona usuaria de servicios profesionales de Trabajo Social, frente a una amplia mayoría (85,5%) que manifiesta lo contrario. Por último, no se observa variabilidad en cuanto a la realización de Prácticas Externas, ya que la totalidad del alumnado participante no ha cursado estas asignaturas.

2.2. Instrumentos

La recogida de información se lleva a cabo mediante un cuestionario estructurado en dos bloques. El primero recoge variables sociodemográficas y académicas, como la edad, el sexo, la vía de acceso a la universidad, la situación laboral actual, la situación de convivencia, la posible vivencia de experiencias de discriminación, la utilización de servicios profesionales de Trabajo Social y la realización de Prácticas Externas. El segundo es una escala de percepción de logro competencial, diseñada *ad hoc* a partir de la revisión de la literatura y adaptada al contexto del estudio. Su aplicación supone una primera aproximación que permitirá recoger información sobre la percepción del alumnado y analizar aspectos como la claridad, coherencia y adecuación de la escala, con el fin de identificar aspectos susceptibles de mejora en aplicaciones futuras.

La escala de percepción de logro competencial es de tipo Likert de 7 puntos, donde 1 es totalmente en desacuerdo y 7 totalmente de acuerdo. Desde el punto de vista teórico, se estructura en tres fases. La primera hace referencia a las competencias vinculadas con la lectura individual, entendidas como aquellas relacionadas con la comprensión, el análisis y la problematización de los textos. La segunda se refiere a las competencias de reflexión intersubjetiva, esto es, aquellas asociadas al contraste de interpretaciones, al intercambio argumentativo y a la construcción compartida de sentido. La tercera agrupa las competencias relacionadas con la aplicabilidad en la práctica, orientadas a valorar la conexión entre la lectura académica, la interpretación de la realidad social y la fundamentación de la intervención profesional. Cada una de estas fases queda compuesta por 7 ítems, sumando un total de 21 (Anexo I).

Desde el punto de vista psicométrico, el instrumento muestra una consistencia interna muy elevada. Así, las fases de lectura individual y reflexión intersubjetiva presentan, en ambos casos, un α=0,89, alcanzando la de aplicabilidad en la práctica un α=0,91. Por su parte, el conjunto de la escala obtiene un α=0,96.

2.3. Procedimiento

La asignatura, de carácter eminentemente teórico, busca articular los fundamentos de los distintos modelos de intervención con la práctica profesional. Por este motivo el equipo docente ha selecciona-

do, para cada uno de los tres bloques temáticos del Plan docente, artículos de revistas especializadas en Trabajo Social, priorizando aquellos que incorporan enfoques contemporáneos, metodologías innovadoras y evidencia empírica relevante, de manera que el estudiantado pueda establecer conexiones entre teoría y práctica. A partir de aquí, el Seminario de lecturas incluye dos etapas:

(1) cada estudiante escoge una de las tres lecturas propuestas. Tras su lectura redacta un comentario en el que copia literalmente un fragmento que haya elegido y expone por qué ha escogido aquel fragmento en concreto, permitiendo al estudiantado aproximarse de manera individual a los textos;

y (2) el día de la sesión, se subdivide al grupo en función de la lectura escogida. Primero se lee en voz alta (tal como aparece en el texto, respetando la voz del autor o autora), y luego se explican las razones para haberlo elegido. El debate mantenido fomenta el pensamiento crítico, ya que ofrece la oportunidad de compartir diferentes puntos de vista y reconocer las propias suposiciones (Brookfield, 2012).

El equipo docente articula ambas etapas para favorecer tanto la familiarización individual con los textos como la interacción colaborativa y crítica, consolidando un aprendizaje profundo y significativo. Al mismo tiempo, modera el grupo, garantizando el respeto de los turnos de palabra, evitando interrupciones y acotando la duración de las intervenciones, de modo que se promueva una participación amplia, equitativa y enriquecedora.

La administración del cuestionario se realiza a la finalización del Seminario de lecturas. Antes de su cumplimentación, el alumnado recibe información sobre los objetivos del estudio y el carácter voluntario de su participación. Se obtiene el consentimiento informado expreso de cada participante. El tratamiento de la información se realiza bajo condiciones de anonimato y confidencialidad.

El análisis de datos es univariado y bivariado, utilizando el programa Statistical Package for Social Sciences (SPSS), versión 25. En primer término, se procede al análisis descriptivo univariado a través del cálculo de frecuencias absolutas y porcentajes, así como de medidas de tendencia central (media y mediana) y de dispersión (desviación típica y rango). Para obtener las puntuaciones dimensionales, se calculó la media aritmética de los siete ítems correspondientes a cada dimensión Posteriormente, el análisis bivariado se lleva a cabo empleando las pruebas de U de Mann-Whitney y Kruskal-Wallis, así como el coeficiente de correlación de Spearman.

3. RESULTADOS

En términos generales, los resultados evidencian una percepción de logro competencial medio-alta en el conjunto de la experiencia (Tabla 1). Específicamente, la dimensión de lectura individual registra la media más elevada (M=4,85; DT=1,23), seguida de la de aplicabilidad en la práctica (M=4,79; DT=1,28) y la de reflexión intersubjetiva (M=4,72; DT=1,32).

En cuanto a la dimensión de lectura individual, el alumnado percibe un mayor desarrollo en su capacidad para comprender mejor los problemas sociales actuales desde una perspectiva crítica (M=5,11; DT=1,44), impulsar una lectura más analítica y menos superficial (M=5,01; DT=1,68) y reflexionar sobre el propio posicionamiento personal ante los temas tratados (M=4,99; DT=1,56).

Respecto a la dimensión de aplicabilidad en la práctica, destaca la percepción del alumnado sobre su capacidad para relacionar la teoría con situaciones reales de la profesión (M=5,19; DT=1,46). Asimismo, percibe una mejora en su aprendizaje de la asignatura a través del seminario (M=5,00; DT=1,93).

Finalmente, en relación con la dimensión de reflexión intersubjetiva, las puntuaciones más altas corresponden a la capacidad percibida para identificar supuestos implícitos en las lecturas (M=4,87; DT=1,77) y generar preguntas críticas en el intercambio grupal (M=4,86; DT=1,70).

El análisis bivariado no muestra diferencias estadísticamente significativas en ninguna de las fases de logro competencial en virtud del perfil sociodemográfico y académico del alumnado.

Tabla 1. Grado de desarrollo competencial percibido

	Media	**DT**	**Mediana**	**Rango**
1.	4,56	1,68	5,00	6
2.	4,73	1,63	5,00	6
3.	5,01	1,68	5,50	6
4.	4,60	1,48	4,50	6
5.	4,94	1,58	5,00	6
6.	5,11	1,44	5,00	6
7.	4,99	1,56	5,00	6
Lectura individual	4,85	1,23	5,14	5,86
8.	4,87	1,77	5,00	6
9.	4,46	1,59	4,00	6
10.	4,86	1,70	5,00	6
11.	4,40	1,77	5,00	6
12.	4,83	1,69	5,00	6
13.	4,83	1,70	5,00	6
14.	4,81	1,69	5,00	6
Reflexión intersubjetiva	4,72	1,32	5,00	5,86
15.	5,19	1,46	5,00	6
16.	4,69	1,48	5,00	6
17.	4,63	1,33	4,50	6
18.	4,73	1,68	5,00	6
19.	4,49	1,64	4,00	6
20.	4,80	1,61	5,00	6
21.	5,00	1,93	5,00	6
Aplicabilidad en la práctica	4,79	1,28	4,93	5,43

Fuente: Elaboración propia. Las puntuaciones de lectura individual, reflexión intersubjetiva y aplicabilidad en la práctica corresponden a la media aritmética de los ítems integrantes de cada dimensión.

Por otro lado, las tres fases consideradas presentan correlaciones positivas, elevadas y significativas entre sí ($p<0,01$), lo cual sugiere que se comportan como componentes estrechamente relacionados de una misma experiencia educativa. La asociación más intensa se observa entre la reflexión intersubjetiva y la aplicabilidad en la práctica (rho=0,872), seguida de entre la lectura individual y la aplicabilidad en la práctica (rho=0,861) y, por último, entre la lectura individual y la reflexión intersubjetiva (rho=0,813).

4. DISCUSIÓN Y CONCLUSIONES

Los resultados obtenidos tras la aplicación de la escala Likert reflejan una valoración sumamente positiva y, sobre todo, equilibrada por parte del alumnado respecto al Seminario de lecturas. Además, muestran que la percepción sobre el desarrollo de las competencias analizadas no se encuentra condicionado por sus características sociodemográficas o académicas. Esta ausencia de diferencias podría interpretarse como un indicio de la capacidad del Seminario de lectura para generar un entorno de aprendizaje relativamente homogéneo en términos de logro competencial, independientemente de las trayectorias previas o circunstancias personales del estudiantado. Asimismo, el hecho de que una parte relevante del alumnado combine estudios y empleo, o haya experimentado situaciones de discriminación, no parece incidir de manera diferencial en los resultados, lo que invita a reflexionar sobre el potencial inclusivo de la propuesta pedagógica implementada.

Los hallazgos ponen de manifiesto una percepción general de logro competencial situada en niveles medio-altos en el conjunto de la experiencia analizada (Tabla 1). En cuanto a las fases específicas evaluadas, la lectura individual se configura como la más valorada, lo que plantea una apreciación especialmente positiva de los procesos de aprendizaje autónomo. A continuación, se encuentra la dimensión relativa a la aplicabilidad práctica de los contenidos, evidenciando una percepción favorable respecto a la transferencia de los aprendizajes a contextos reales. Finalmente, la reflexión intersubjetiva, aunque se sitúa como la dimensión con menor valoración relativa, sigue mostrando una percepción positiva, aunque con mayor variabilidad en las respuestas.

A nivel de contenidos específicos, el análisis de las fases ordenadas de mayor a menor valoración, permite identificar los aspectos donde existe mayor consenso y éxito metodológico. En primer lugar, la fase individual del Seminario destaca por la capacidad de las lecturas para fomentar una apropiación crítica del texto (Freire, 2012, 2017). Se trata de un momento que permite que el/la estudiante interactúe desde sus propias vivencias con el texto, sin mediadores/as que le dicten qué pensar. En este sentido, los resultados apuntan a que el alumnado siente consolidar procesos de comprensión y percibe avanzar hacia formas de análisis más críticas y profundas de los problemas sociales contemporáneos. Este hallazgo sugiere una apreciación de fortalecimiento del pensamiento crítico, evidenciada tanto en la percepción de capacidad para superar aproximaciones superficiales como en el desarrollo declarado de una mayor conciencia reflexiva sobre el propio posicionamiento ante las temáticas abordadas (Fook & Gardner, 2007; Sheppard, 1998).

En segundo lugar, en la fase de la aplicabilidad en la práctica se observa que el estudiantado otorga un valor significativo a la conexión entre los contenidos teóricos y su proyección en contextos profesionales reales. Los resultados muestran consenso en que la actividad del Seminario es valorada como determinante para conectar la teoría con las situaciones reales de la profesión, reafirmando la percepción de que el conocimiento no se agota en la comprensión conceptual, sino que debe impactar en la acción, reduciendo el individualismo cognitivo, fomentando una cultura democrática y desarrollando habilidades deliberativas transferibles a contextos profesionales (Freire, 2012, 2017). Esta percepción refuerza la relevancia del enfoque metodológico adoptado, en la medida en que favorece la valoración de procesos de aprendizaje significativo y orientado a la práctica profesional, contribuyendo además a la percepción de una mejor asimilación de los contenidos de la asignatura.

Finalmente, en la fase intersubjetiva, los resultados ponen de relieve la importancia atribuida al intercambio grupal como espacio para el desarrollo de habilidades analíticas complejas percibidas. En particular, se reporta una mayor capacidad para formular interrogantes críticos, lo que revela la

percepción de que la interacción con otros/as potencia procesos de co-construcción del conocimiento (Brookfield, 2012). En este sentido, el alumnado valora especialmente cómo el debate grupal facilita la identificación de supuestos implícitos que a menudo pasan desapercibidos en una lectura individual. Se promueve así la construcción colectiva de significados a través del diálogo igualitario, ya que el conocimiento reside en el espacio entre las personas, un espacio en el que todas las interpretaciones son válidas siempre que se basen en argumentos y no en el lugar que se ocupa según el poder del que se disponga (Habermas, 1987; Vygotsky, 2012).

Las correlaciones positivas, elevadas y estadísticamente significativas entre las tres fases analizadas evidencian la existencia de una fuerte interdependencia entre las percepciones de los distintos componentes secuenciados de la actividad. De esta forma, de los resultados se desprende que la lectura individual, la reflexión intersubjetiva y la aplicabilidad en la práctica no son percibidas como habilidades aisladas, sino como elementos integrados que se potencian mutuamente en el marco de una misma experiencia educativa. Especialmente relevante resulta la estrecha asociación entre la valoración de la reflexión compartida y percepción de transferencia a contextos prácticos, lo que apunta al valor del intercambio grupal como facilitador percibido de procesos de comprensión aplicada (Afdal et al., 2023; Ament et al., 2025). Asimismo, la relación entre la lectura individual y la aplicabilidad refuerza la idea de que una base sólida de análisis autónomo contribuye de manera significativa a la movilización del conocimiento en situaciones reales En conjunto, estos hallazgos respaldan una concepción del aprendizaje como proceso holístico, en el que la adquisición, la reflexión y la aplicación del conocimiento se articulan de manera dinámica y recíproca.

La experiencia de los Seminarios de lecturas evidencia que incorporar espacios estructurados de reflexión y análisis en la formación en Trabajo Social es percibido como una estrategia que favorece la innovación docente al combinar lectura individual, reflexión compartida y aplicación práctica de los contenidos. Esta estrategia permite el estudiantado reporte un desarrollo integrado del pensamiento autónomo y de las competencias reflexivas, mostrando que los procesos de aprendizaje no operan de forma aislada, sino que se potencian mutuamente. Para la docencia, esto implica diseñar actividades que promuevan tanto la comprensión autónoma como el intercambio grupal, facilitando la transferencia de conocimientos a contextos reales y mejorando la conexión entre teoría y práctica (Kinni, 2020; Medina et al., 2024). Asimismo, los resultados señalan que la innovación en los MÉTODOs de enseñanza además de incrementar la percepción de logro competencial, ofrece al profesorado herramientas para ajustar y enriquecer los programas formativos, orientándolos a fortalecer la reflexividad crítica de manera sistemática y efectiva en la educación superior en Trabajo Social.

A modo de conclusión, los Seminarios de lecturas constituyen un recurso formativo percibido como especialmente significativo en la preparación de futuros/as profesionales porque generan un espacio que permite a los/as participantes profundizar en contenidos relevantes y promover la valoración de competencias importantes como la escucha activa, el respeto a los/as demás y el aprendizaje colaborativo, todas ellos aspectos fundamentales en su futuro profesional ((Federación Internacional de Trabajadores Sociales & Asociación Internacional de Escuelas de Trabajo Social, 2020).

Los/as participantes en el Seminario de lecturas reportan un aprendizaje significativo: valoran la autonomía al realizar la lectura del texto por su cuenta, reconocen la utilidad práctica de lo aprendido en situaciones reales y consideran enriquecedor el intercambio de ideas y la reflexión conjunta a través de debates y discusiones sin que se identifiquen brechas asociadas a su perfil sociodemográfico del estudiantado.

El estudio presenta varias limitaciones que deben ser consideradas. En primer lugar, la información obtenida refleja la percepción del alumnado sobre la contribución formativa del Seminario, por lo que no permite determinar cambios objetivos en las competencias evaluadas ni atribuirlos causalmente a la actividad. Además, el empleo exclusivo de una escala tipo Likert limita la posibilidad de conocer con mayor profundidad cómo experimentó el alumnado las distintas fases del Seminario y qué procesos concretos favorecieron o dificultaron su aprendizaje. En segundo lugar, la muestra estuvo integrada por 70 estudiantes de una única asignatura, titulación y universidad. Su tamaño y procedencia restringen la generalización de los resultados y pueden haber limitado la capacidad para detectar diferencias asociadas a las características sociodemográficas y académicas del alumnado. Por último, la escala fue diseñada específicamente para este estudio y organizada teóricamente en tres dimensiones. El tamaño de la muestra y el carácter exploratorio de la investigación no permiten determinar con suficiente robustez su estructura interna. El elevado coeficiente alfa de Cronbach del conjunto de la escala y las fuertes correlaciones entre las dimensiones podrían ser compatibles tanto con la existencia de componentes estrechamente relacionados como con un factor general de valoración de la experiencia o con cierto solapamiento entre los contenidos de los ítems.

En coherencia con estas limitaciones, las investigaciones futuras deberían incorporar diseños longitudinales o mediciones pretest y postest, preferentemente acompañadas de indicadores de desempeño, para examinar con mayor precisión los posibles cambios en las competencias evaluadas. Asimismo, resultaría conveniente adoptar diseños mixtos que combinen la escala con entrevistas, grupos de discusión, observación de las sesiones o análisis de las producciones escritas del alumnado. También sería necesario ampliar el tamaño y la diversidad de la muestra mediante la participación de estudiantes de diferentes cursos, asignaturas, titulaciones y universidades. En relación con el instrumento, futuras aplicaciones con muestras más amplias deberán examinar su estructura dimensional y aportar nuevas evidencias sobre la fiabilidad y la validez de sus puntuaciones. Finalmente, podrían evaluarse comparativamente algunas variantes en la implementación del Seminario, como la incorporación de preguntas orientadoras para la lectura individual, la asignación de roles durante la discusión grupal o el uso de otras dinámicas participativas.

5. REFERENCIAS

Afdal, H. W., Spernes, K., & Hoff-Jenssen, R. (2023). Academic reading as a social practice in higher education. *Higher Education, 85*, 1337–1355. https://doi.org/10.1007/s10734-022-00893-x

Ament, J. R., Tort-Cots, I., & Pladevall-Ballester, E. (2025). Implementing academic reading circles in higher education: Exploring perceptions, motivation and outcomes. *Journal of English for Academic Purposes, 75*, 1-12. https://doi.org/10.1016/j.jeap.2025.101489

Brookfield. S. D. (2012). *Teaching for critical thinking. Tools and techniques to help students question their assumptions*. JosseyBass.

Federación Internacional de Trabajadores Sociales & Asociación Internacional de Escuelas de Trabajo Social (2020). *Normas mundiales para la educación y la formación en materia de Trabajo social*. https://www.iassw-aiets.org/wp-content/uploads/2021/05/IASSW-IFSW-Global_Standards_Final-ES.pdf

Fook, J., & Gardner, F. (2007). *Practising critical reflection: A resource handbook*. Maidenhead: Open University Press & McGraw Hill.

Foote, W. (2013). Threshold theory and social work education. *Social Work Education, 32*(4), 424–438. https://doi.org/10.1080/02615479.2012.680436

Fortune, A. E., & Reid, W. J. (1999*). Research in social work (3rd ed.).* Columbia University Press.

Freire, P. (2012). *Pedagogía del oprimido.* Siglo XXI.

Freire, P. (2017). *Pedagogía de la autonomía: saberes necesarios para la práctica educativa.* Siglo XXI.

Habermas, J. (1987). *Teoría de la acción comunicativa (Vols. I y II)*. Taurus.

Hennessy, S., Rojas-Drummond, S., Higham, R., Márquez, A. M., Maine, F., Ríos, R. M., García-Carrión, R., Torreblanca, O. & Barrera, M. J. (2016). Developing a coding scheme for analysing classroom dialogue across educational contexts. *Learning, Culture and Social Interaction, 9,* 16-44. https://doi.org/10.1016/j.lcsi.2015.12.001

Hernández, R., Fernández, C., & Baptista, P. (2006). *Metodología de la investigación (4.ª ed.)*. McGraw-Hill Interamericana.

Kinni, R.-L. (2020). *Integration of theory and practice in social work education: Analysis of Finnish social work students' field reports. Social Work Education, 40(*7), 901–914. https://doi.org/10.1080/02615479.2020.1754385

Loza, M. (2004). Tertulias Literarias. *Cuadernos de Pedagogía 341*, 66–69.

Medina, M. V., Álvarez, G., & García, F. (2024). *La integración de conocimientos teóricoprácticos desde la experiencia de profesionales del Trabajo Social en la formación del estudiantado a través de metodologías activas y participativas. Research in Education and Learning Innovation Archives*, 32, 91–104. https://hdl.handle.net/10481/89616

Sheppard, M. (1998). Practice validity, reflexivity and knowledge for social work. *The British Journal of Social Work, 28*, 763–781. https://doi.org/10.1093/oxfordjournals.bjsw.a011390

Vygotsky, L. S. (2012). *Pensamiento y lenguaje.* Paidós.

6. ANEXOS

Anexo I. Escala Likert organizada según fases

	Ítems
Lectura individual	1. El seminario de lecturas me ayudó a identificar con claridad la tesis central de los textos. 2. Pude reconocer los argumentos principales y secundarios con mayor precisión gracias al seminario. 3. Considero que el seminario impulsó una lectura más analítica y menos superficial. 4. Gracias al seminario, pude detectar contradicciones, tensiones o vacíos en los argumentos de los textos. 5. El seminario favoreció que cuestionara las ideas presentadas en los textos. 6. Las lecturas me permitieron comprender mejor los problemas sociales actuales desde una perspectiva crítica. 7. El seminario me llevó a reflexionar sobre mi posicionamiento personal ante los temas estudiados.
Reflexión intersubjetiva	8. Gracias al debate en el seminario, soy capaz de identificar supuestos implícitos en las lecturas. 9. Me permitió confrontar mis propias creencias y analizar su origen. 10.El intercambio con mis compañeros/as estimuló mi capacidad para generar preguntas críticas. 11. Percibo cambios en mi forma de interpretar los fenómenos sociales después de participar en el seminario. 12. El seminario mejoró mi capacidad de argumentar de forma fundamentada. 13. He desarrollado la capacidad de integrar las aportaciones de los demás para enriquecer la calidad de mis argumentos. 14. El seminario fortaleció mi habilidad para expresar desacuerdos de manera respetuosa y fundamentada.
Aplicabilidad práctica	15. Las discusiones del seminario me ayudaron a relacionar la teoría con situaciones reales de la profesión social. 16. El seminario contribuyó a que analizara las implicaciones éticas de las intervenciones profesionales. 17. Siento que ahora puedo establecer conexiones más fuertes entre los textos y el contexto de intervención social. 18. La experiencia del seminario potenció mi capacidad de autoevaluar mis decisiones o perspectivas. 19. El proceso me ayudó a tomar conciencia de mis sesgos y marcos de referencia. 20. Me siento más capaz de justificar mis ideas con base en teorías y evidencias. 21. Me siento más capaz de integrar los aprendizajes de la asignatura gracias a mi participación en el seminario.

Fuente: Elaboración propia

8. Visible Only to Those Who Know Where to Look: Multimodal Literacy, Visual Grammar, and Critical Reading in the L2 Classroom[1]

Pilar Escabias Lloret

Departamento de Filología Inglesa, Universidad de Alicante

ABSTRACT

Film posters are dense multimodal texts that simultaneously encode narrative, ideology, and promotional strategy within a single visual space, yet for most language learners, only the surface is visible. This paper reports on the first phase of a pedagogical redesign project implemented in a B2-level Spanish as an Additional Language (SAL) course for North American university students. Drawing on Halliday's (2004) Systemic Functional Linguistics and Kress and van Leeuwen's (1996, 2001) Visual Grammar, the project uses a qualitative pre/post design to document shifts in students' readings of two versions of the poster for Guillermo del Toro's *El laberinto del fauno* (2006). Before instruction, seven students produced readings limited to denotative nouns and affective adjectives; no student attended to any verbal element of either poster beyond the title. After one theoretical session and film viewing, the majority deployed semiotic metalanguage, recognised verbal poster elements as analytical objects, and began to articulate ideological and gender dimensions. The study argues that film posters, approached through an explicit theoretical framework, are effective pedagogical objects for developing critical multimodal literacy in L2 contexts. The study specifically aims to investigate whether and how brief explicit instruction in multimodal analytical frameworks transforms the way L2 learners read film posters as semiotic texts.

KEYWORDS: multimodal literacy; visual grammar; film posters; L2/LAL teaching; gender critique; SFL

1. INTRODUCTION

Film posters are everywhere. Students in additional language courses encounter them daily: on cinema websites, bus shelter posters, outdoor advertising displays, and social media, yet rarely examine them with the analytical attention reserved for written texts. This gap matters: posters are dense multimodal artefacts that simultaneously construct character, genre, cultural identity, and gender through carefully orchestrated combinations of image, typography, colour, and layout. As Kress and van Leeuwen (2001) argue, visual and verbal modes are not merely parallel channels of communication but distinct yet interacting systems of meaning-making, each carrying ideological baggage that is far from transparent.

In Barthes's (1977) terms, responses remain at the level of denotation, naming what is pictorially present, without reaching connotation or ideology. As Berger (1972) observed, practices of looking are learned and cultural rather than natural, and can be transformed through education. How to equip language learners with the conceptual vocabulary to read images as texts: that is, as social acts encoding choices about what to represent, how, and for whom.

1 El presente trabajo ha contado con una ayuda del Programa de Redes de investigación en docencia universitaria del Instituto de Ciencias de la Educación de la Universidad de Alicante (convocatoria 2024-26). Red 6328: El cine como herramienta didáctica para la alfabetización audiovisual crítica y la didáctica de las lenguas en Educación Primaria y Secundaria. Coordinadora: Natalia Contreras de la Llave.

Film and audio-visual materials have long been recognised as motivating and culturally rich resources in language teaching, and recent research has demonstrated their potential for developing critical literacy and gender awareness in additional language contexts (Contreras de la Llave, 2022; Contreras de la Llave & Pastor Cesteros, 2023). The present study extends this work by focusing on the film poster as a self-contained pedagogical object: dense, visually complex, and culturally embedded, grounded in an explicit multimodal theoretical framework and assessed through a pre/post design.

The study aims to investigate whether and how explicit instruction in a multimodal analytical framework, Halliday's (2004) SFL metafunctions extended visually through Kress and van Leeuwen's (1996, 2001) *Visual Grammar*, produces a qualitative transformation in L2 university learners' capacity to read film posters as semiotic texts encoding representational, interpersonal, and ideological meanings. This objective is operationalised through a pre/post qualitative design measuring shifts in students' metalanguage use, textual awareness, gaze recognition, gender critique, and cross-poster comparison.

2. THEORETICAL FRAMEWORK

2.1. Systemic Functional Linguistics and Visual Grammar

The project draws on two complementary frameworks. Halliday's (2004) Systemic Functional Linguistics (SFL) conceptualises language as a social semiotic resource that simultaneously performs three metafunctions: the *ideational* (representing the world), the *interpersonal* (enacting social relations), and the *textual* (organising meaning into coherent messages).

Kress and van Leeuwen (1996, 2001) extend these metafunctions to the visual domain through their *Visual Grammar*, proposing that images simultaneously perform *representational*, *interactive*, and *compositional* functions. At the representational level, images depict participants and narrative processes. At the interactive level, they construct viewer relationships through resources including *gaze* (demand, where the depicted participant addresses the viewer directly, versus offer, where they do not), *social distance* (encoded through shot size), and *angle*. At the compositional level, meaning is organised through *salience* (which elements attract visual attention), *framing*, and *information value* (the significance attached to positions in the visual field). Kress and van Leeuwen (2002) further argue that colour and typography are semiotic resources subject to the same framework, a particularly productive claim for film poster analysis, where typeface choices, award badges, taglines, and credits are all meaningful acts.

2.2. Multimodal Literacy, Film Posters, and Gender

Research at the intersection of multimodal analysis and film poster studies has grown substantially. Boysan and Kızıltan (2023) demonstrate the semiotic complexity of contemporary film posters across genre and culture; Budiastono and Noverino (2024) show how multimodal discourse analysis reveals ideological constructions in promotional materials; and Khan and Akhtar (2024) document the visual objectification and gender dynamics encoded in poster composition and gaze. Navarro and Adriano (2023) identify recurrent patterns of marginalisation in the representation of women in film poster design, while de Miguel Casado and García Carrizo (2025) document

persistent patterns of visual subordination in posters for Goya Award-nominated films in Spanish cinema, a finding directly relevant to the corpus analysed in this study.

The gender-critical dimension of the project draws on a tradition of feminist visual analysis, from Mulvey's (1975) foundational account of the gaze as a site of gendered power to more recent empirical work on gender representation in film promotional discourse, which provides the conceptual background for reading compositional choices as ideologically gendered. In this study, that background informs the analytical categories used to assess student responses, rather than constituting a direct object of instruction: what students were asked to notice was the differential construction of the female protagonist's agency, visibility, and spatial positioning across the two poster versions. Herrero and Vanderschelden (2019) demonstrate the effectiveness of film-based pedagogy in developing multimodal and critical literacies across languages, while Holguín Vaca and Castaño Arques (2024) document how critical visual literacy can foster interculturality specifically within the Spanish as an additional language classroom, providing a close empirical parallel to the present study.

3. METHODOLOGY

3.1. Description of the Context and Participants

This study is embedded within a Spanish Cinema course forming part of a B2 Spanish as an Additional Language programme for North American students enrolled in international study abroad programmes at the Universidad de Alicante. The course addressed civil war cinema and the post-war period through *El laberinto del fauno* (*Pan's Labyrinth*, Guillermo del Toro, 2006), and other relevant films. With the course instructor's permission, I designed and delivered a three-session multimodal literacy workshop embedded within the existing syllabus. This arrangement had methodological relevance since I held no evaluative authority over the students, substantially reducing the risk of social desirability bias in their responses. The group comprised seven participants (N = 7), of whom six are native English speakers studying Spanish as an additional language, and one is a native Spanish speaker from a Hispanic American country enrolled in the same programme. The seven participants represented the complete intact group enrolled in the course. Participants are identified by letters A-G. The use of student written productions as research data was approved by the course instructor and programme director. No identifying personal data are disclosed.

The inclusion of one native Spanish speaker provided a useful comparison case: her pre-theory responses displayed the same absence of semiotic metalanguage as those of her L2 peers, suggesting that the limitation was conceptual rather than linguistic.

The analytical corpus consists of two versions of the *El laberinto del fauno* poster (see Figures 1 and 2): the Spanish theatrical release and the English DVD edition, *Pan's Labyrinth*, distributed in the Anglophone market. The film formed part of the course syllabus and provided two culturally distinct versions suitable for contrastive multimodal analysis. The posters were selected because they integrate visual, typographic, textual, and compositional elements in culturally and ideologically significant ways.



Figure 1. Spanish theatrical release poster, *El laberinto del fauno* (del Toro, 2006). © Tequila Gang / Estudios Picasso / Telecinco Cinema. Reproduced for scholarly analysis. Reproduction rights requested from Telecinco, Mediaset, in compliance with applicable intellectual property legislation.



Figure 2. English DVD edition poster, *Pan's Labyrinth* (del Toro, 2006). © Warner Bros. Entertainment. Reproduced for scholarly analysis. Reproduction rights requested from Warner Bros. Entertainment España S.L., in compliance with applicable intellectual property legislation.

3.2. Instruments

To make my own interpretive position transparent, a requirement of reflexive qualitative design, a brief analytical reading of the two posters is offered here as a reference point against which student responses can be assessed. This reading is informed by extended engagement with the film and its visual and literary intertexts; it does not represent the expected horizon of student response after a single session of instruction and a single viewing.

The Spanish poster places Ofelia centrally, facing the viewer with a demand gaze (Kress & van Leeuwen, 1996) that directly implicates the audience in her situation. She is almost enclosed by tree roots; a fairy hovers near her face; the Faun watches from the upper right; the eyeless creature looms from the lower left. At the compositional level, thirteen Goya Award nominations appear in the upper right in gold lettering, evoking the visual iconography of prestige cinema awards. Additionally, under the movie title *El laberinto del fauno*, the director's credit "*Una película de Guillermo del Toro*" appears in red typography. The dominant colour palette of ochres, browns, and earth tones anchors the image in a world of decay and natural matter. Crucially, the absence of promotional quotations from external critics positions the poster as requiring no external mediation.

The English poster recomposes the visual field almost entirely. Blue and black tones replace the warm ochres of the Spanish version, shifting the emotional register towards a gothic atmosphere. Ofelia appears from behind, her figure small against an enormous ancient tree whose gnarled form suggests a monster about to swallow her. An offer gaze is constructed: we are positioned as external observers of her situation rather than participants she addresses. At the compositional level, a press quotation at the top: *"Beautiful and Exhilarating… A Fairy Tale for Grown-Ups"* is attributed to Jonathan Ross, a British television personality. The tagline *"Innocence Has a Power Evil Cannot Imagine"* interpellates a viewer who may not yet know the film's moral stakes. Both posters are calibrated for different audiences through distinct visual and promotional strategies.

3.3. Procedure

The workshop was structured across three sessions of approximately two hours each. Table 1 provides an overview of the full design; each session is described in detail below.

Table 1. Workshop design: sessions, phases, and instruments.

Session	Phase	Instrument	Description
Session 1 *c. 2 hours, in class*	Part A Pre-theory	*Activities 1–3*	Students observe the English and Spanish poster versions without any theoretical instruction. Activity 1: free enumeration of what they see. Activity 2: guided questions on atmosphere, central figure, and genre. Activity 3: cross-poster comparison. No correct or incorrect answers.
	Part B Theoretical instruction	*Summary handout + worked example*	Introduction of SFL metafunctions and Visual Grammar concepts (representational, interactive, compositional; gaze, social distance, salience, framing, information value). Worked analytical table applying the three metafunctions to the verbal elements of the *Ágora* poster (Amenábar, 2009) as a modelled example (see Figure 3). Students explicitly instructed that verbal poster elements are semiotic resources.

Session	Phase	Instrument	Description
	Part C Post-theory	*Task 1 + Task 2*	Task 1: structured analysis applying the three metafunctions to both poster versions. Task 2: metacognitive reflection on the change in perception between the first and second reading, explicitly attributing any change to the theoretical frameworks introduced.
Session 2 *c. 2 hours, in class*	Film viewing	*In-class screening*	In-class viewing of *El laberinto del fauno* (del Toro, 2006) in full (with English subtitles), with no analytical tasks during the viewing.
Session 3 *c. 2 hours independent study*	Post-film re-reading	*Task 3*	Re-reading of both poster versions in light of the film, asking students to consider whether and how the narrative and ideological knowledge acquired through viewing modified their analytical reading. Completed independently at home.

Session 1 began with the pre-theory phase (Part A). Students completed Activities 1-3 before any theoretical instruction, producing the spontaneous, pre-theoretical readings that constitute the baseline data for this study. Students were explicitly told there were no correct or incorrect answers.

Part B introduced the theoretical framework. The framework was presented to students in the form of a summary handout and a worked analytical table applying the three metafunctions to the verbal elements of the poster for *Ágora* (Amenábar, 2009), a film not included in this course but used here as a pedagogical illustration (its poster also features a central female figure addressing the viewer directly with a demand gaze, making it visually relevant to the analytical questions raised by the *El laberinto del fauno* corpus). Students were explicitly told that verbal elements: title, director credit, award nominations, tagline, and critical quotations, are semiotic resources as analytically significant as the visual image itself, and were expected to account for them in their post-theory analyses.

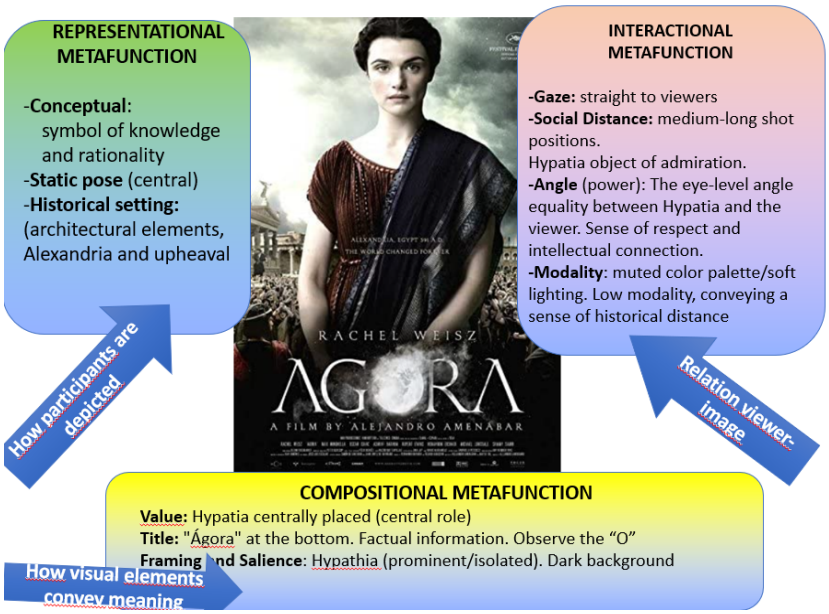


Figure 3. Classroom worked example: multimodal analysis of the Ágora poster (Amenábar, 2009) with researcher annotations, as presented to students in Session 1, Part B. © Telecinco Cinema / Himenóptero / Mod Producciones. Reproduced for scholarly analysis.

Part C comprised Tasks 1 and 2, completed in class immediately after the theoretical instruction, while the framework was fresh and the instructor was available to clarify terminology. Task 1 was a structured analysis applying the three metafunctions to both poster versions. Task 2 was a metacognitive reflection asking students to describe how their perception of the poster had changed after theoretical instruction and why.

Session 2 involved viewing *El laberinto del fauno* in full.

Session 3 (Task 3), completed at home as independent study, asked students to re-read both poster versions in light of the film and consider how their analytical reading had changed.

Student productions across all phases were analysed through close reading focused on four dimensions: (a) *metalanguage use*: the presence and accuracy of semiotic terminology; (b) *textual awareness*: whether verbal poster elements were recognised as semiotic resources; (c) *gender and ideological critique*: whether power dynamics or gender constructions were identified; and (d) *cross-poster comparison*: the depth of comparison between the two versions. Each dimension was assessed as evidenced, partially evidenced, or absent at each measurement point. All student productions are reproduced verbatim throughout, including non-normative forms characteristic of L2 writing at B2 level; translations are normative and render the intended meaning.

4. RESULTS

4.1. Pre-Theory Phase: The Vocabulary of the Surface

Activities 1-3 (Session 1, Part A) present a consistent profile across all seven students. Activity 1 (free enumeration) produced almost exclusively denotative nouns: *árbol* [tree], *tronco* [trunk], *niña* [girl], *luna* [moon], *arco* [arch], *viñas* [vines], *bosque* [forest], *cara* [face], *vestido* [dress], naming what is literally depicted without interpreting what these elements do. Activity 2 (atmosphere) produced affective adjectives: *oscura* [dark], *misterioso* [mysterious], *asustante* [frightening], *lúgubre* [gloomy], *ominoso* [ominous], *peligroso* [dangerous], *frío* [cold].

On textual awareness: Student D named *el título* [the title] as their first visual focus; Student G referenced *las palabras* [the words] as a spatial landmark. Zero students mentioned any other verbal element of either poster: no director credit, no Goya nominations, no press quotation, no tagline. Table 2 maps the three registers of response documented across the study, illustrating the progression from pre-theory to post-theory reading.

Table 2. Three registers of response across the study: from denotation to semiotic analysis (after Barthes, 1977).

Register	Examples from the data	Analytical function
Denotative nouns *(pre-theory)*	*árbol* [tree], *tronco* [trunk], *niña* [girl], *luna* [moon], *arco* [arch], *viñas* [vines], *bosque* [forest], *cara* [face], *vestido* [dress]	Name what is literally depicted without interpreting what elements do or what relationships they construct. Operate at the level of denotation (Barthes, 1977): pictorially present, without connotation or ideological reading.
Affective adjectives *(pre-theory)*	*oscura* [dark], *misterioso* [mysterious], *asustante* [frightening], *lúgubre* [gloomy], *ominoso* [ominous], *peligroso* [dangerous], *frío* [cold], *eerie*	Express the viewer's emotional response. Accurate as impressions, but describe the effect of semiotic choices without naming or analysing those choices.

Register	Examples from the data	Analytical function
Semiotic meta-language *(post-theory only)*	*mirada de demanda* [demand gaze], *saliencia* [salience], *distancia social* [social distance], *jerarquía visual* [visual hierarchy]	Analyse how visual and verbal elements construct meaning — how they position the viewer, encode ideology, and make representational choices. Entirely absent in pre-theory responses.

Note: The table maps the progression documented in this study. Pre-theory responses were confined to the first two registers; the third emerged only after theoretical instruction (Session 1, Part B). The progression corresponds to Barthes's (1977) distinction between denotation and connotation, and to Kress and van Leeuwen's (1996) argument that untrained readers perceive what is in an image but not how it works.

The consistency of this profile across all seven students, including the native Spanish speaker, is illustrated by three representative responses spanning the full range of the group. At one end of the spectrum, Student A's Activity 1 is a pure enumeration of denotative nouns.

> *"Una niña mirando para cima. Unas arboles curvadas. Un cielo gris y oscurro. Muchas viñas. Una arco de piedra. Una cara en el arco. Una luna muy grande."*
>
> [*A girl looking upwards. Some curved trees. A grey and dark sky. Many vines. A stone arch. A face in the arch. A very large moon.*] *Student A, Session 1, Part A, Activity 1 (pre-theory, English poster)*

This response names seven elements correctly but interprets none of them. Each element is present; none is analysed. At the other end of the spectrum, Student E produced the most elaborate pre-theory response:

> *"Hay una niña en el centro de una laguna de piedra, árboles y otras cosas naturales durante la noche… La sensación de isolación del mundo real y con seguridad es presente y un mundo estérico y ficcional reemplaza."*
>
> [*There is a girl at the centre of a stone lagoon, trees and other natural things at night… The sensation of isolation [isolación: non-normative form of aislamiento] from the real world and from safety is present and a strange, otherworldly [estérico: likely intended esoteric] and fictional world replaces it.*] *Student E, Session 1, Part A, Activity 1 (pre-theory, English poster)*

This response is more syntactically complex and emotionally perceptive than Student A's, and yet it too operates entirely within the registers of denotation and affect. The student perceives that the poster constructs a sense of isolation and unreality but cannot yet account for *how*, through what specific visual and compositional choices, this effect is produced. This is precisely what the framework makes possible. The contrast between these pre-theory responses and the same students' post-theory productions is one of the main pieces of evidence this study offers.

Within this consistent pre-theory profile, several responses reveal the latent potential of unaided perception. Without any theoretical prompt, Student E described the protagonist *"como una víctima con diferentes niveles de autonomía de la situación"* [as a victim with different levels of autonomy of the situation], using *autonomía* unprompted and pre-theory. Student F annotated the paired poster image directly with the binary *Small, Scared* (English version) versus *Larger, empowered* (Spanish version). These near-misses demonstrate that the framework does not manufacture perception, it gives language and analytical structure to what students could already feel but could not systematically articulate.

4.2. Post-Theory Phase: Metalanguage, Textual Awareness, and Gender Critique

Metalanguage use. Six of seven students (6/7) incorporated semiotic terminology correctly in Tasks 1 and 2 (Session 1, Part C): *saliencia* [salience], *mirada* [gaze], *distancia social* [social distance], *función representacional* [representational function], *función composicional* [compositional function], *jerarquía visual* [visual hierarchy], *mirada de demanda* [demand gaze]. Student D's response is the single exception, rendered largely illegible by severe orthographic disruption, a methodological case illustrating the double cognitive load when theoretical metalanguage must be expressed in the L2.

Task 2 produced the clearest evidence of students' awareness of their own development. Student B wrote:

> *"Antes solo describía lo que veía en el cartel. Puedo ver cómo la posición de los elementos, la mirada y la luz crean significado. Mi análisis es más profundo porque entiendo cómo funciona la imagen como texto visual."*
>
> [*Before, I only described what I saw in the poster. I can now see how the position of the elements, the gaze, and the light create meaning. My analysis is deeper because I understand how the image functions as a visual text.*] *Student B, Session 1, Part C, Task 2*

Student A produced the most developed post-theory analysis in Task 1, combining all three metafunctions consistently. On the interactive function:

> *"La mirada de la chica está en la audiencia, como ella quiere que todos vean su historia y que presten la atención. […] También el texto sobre las nominaciones a premios, que sirven para convencer a la gente porque deben mirarlo."*
>
> [*The girl's gaze is directed at the audience, as if she wants everyone to see her story and pay attention. […] Also the text about the award nominations, which serve to persuade people that they should watch it.*] *Student A, Session 1, Part C, Task 1*

The demand gaze and the Goya nominations are both correctly analysed as persuasive compositional elements, absent in the pre-theory phase.

Cross-poster comparison. Four of seven students (4/7) articulated differences between the two versions in terms of semiotic choices with cultural or market implications. Student A's Task 1 analysis of the English poster:

> *"Muy diferente al otro cartel, no hay mirada de ningún personaje, entonces la audiencia se siente más como forasteros que no saben nada del mundo aquí. […] La posición de la niña aquí es mucho más vulnerable, y parece que el arco y el árbol van a consumirla."*
>
> [*Very different from the other poster: there is no gaze from any character, so the audience feels more like outsiders who know nothing about this world. […] The position of the girl here is much more vulnerable, and it seems as though the arch and the tree are going to consume her.*] *Student A, Session 1, Part C, Task 1*

The offer gaze and Ofelia's spatial vulnerability are both read as ideologically meaningful compositional choices.

4.3. Post-Film Phase: Theory and Narrative as Complementary Resources

Task 3 (Session 3) produced the most analytically complex responses among students who had already developed solid post-theory readings in Tasks 1 and 2, confirming that theoretical instruction alone does not activate the deepest levels of gender critique: it requires the combination of semiotic

framework and narrative knowledge. Student E produced the richest post-film analysis, with three analytically significant moves. First, an ideologically inflected revision of their cross-poster preference:

> *"Antes de ver la película me gusta el cartel americano más del español… Después de verla pensé que el cartel español muestra la resistencia y la rebeldía de Ofelia."*
>
> [*Before watching the film I liked the American poster more than the Spanish one… After watching it, I thought that the Spanish poster shows Ofelia's resistance and defiance.*]
>
> *Student E, Session 3, Task 3*

Second, an unprompted gender reading of the tree structure: *"el tronco del árbol… es representante del cuerpo de una mujer"* [the tree trunk represents the female body], connecting the poster's visual syntax to corporeality and gender symbolism. Third, a critical discourse observation about promotional omission:

> *"Los dos carteles no tienen información o pistas sobre la tema de la guerra y la monstruosidad de los humanos en la película."*
>
> [*Neither poster contains information or clues about the theme of war and the monstrousness of the humans in the film.*] *Student E, Session 3, Task 3*

The student is not analysing what the poster contains but what it has chosen to exclude, a critical discourse move representing the most advanced level of multimodal literacy documented in the study. Student C produced a parallel gender reading in Task 3: *"los troncos curvados parecen ovarios que reflejan los problemas de la madre"* [the curved trunks look like ovaries reflecting the mother's problems], connecting the curved trunks noted denotatively in Activity 1 to an unprompted gender-critical reading after theory and film, the clearest single illustration of what the combination of theoretical framework and film experience can unlock (see Figure 4 for the visual correspondence).



Figure 4. Still from *El laberinto del fauno* (del Toro, 2006) showing Ofelia approaching the tree, alongside a schematic representation of the female reproductive system. The visual correspondence between the tree structure in the English poster and the anatomical form underpins the gender reading produced independently by Students C and E after watching the film. © Warner Bros. Entertainment; Telecinco-Mediaset Screenshot reproduced for scholarly analysis.

4.4. Cross-Student Data Summary

Table 3 maps each student's performance across the five analytical dimensions at both measurement points. Each cell shows the pre-theory state on the left and the post-theory state on the right, separated

by an arrow (○/◑/● → ○/◑/●). Pre-theory symbols appear in grey; post-theory symbols appear in bold black, making the direction and extent of each student's development visible at a glance without requiring the reader to process text. Shading alternates by row for readability. The table makes immediately visible one of the study's most consistent findings: with the exception of Student F's pre-theory annotation of the paired poster images, every student begins from a baseline of absence across all five dimensions: the left-hand side of every cell is ○. After a single two-hour session of theoretical instruction, six of seven students moved from absence to evidenced metalanguage use; five moved from absence to evidenced or partial textual awareness; five showed evidenced or partial gaze recognition; and four produced evidenced or partial gender or ideological critique. The movement from ○ to ● or ◑ in a single session is the most economical summary of what explicit multimodal instruction makes possible.

Table 3. Cross-student developmental map (N = 7). Each cell: pre-theory → post-theory. ○ = absent · ◑ = partial · ● = evidenced · ●● = strongly evidenced. Pre-theory symbols in grey; post-theory symbols in bold. Rows shaded alternately for readability. Student D excluded from aggregate counts.

Std.	Metalanguage	Textual awareness	Gaze	Gender / ideological	Cross-poster	
	Pre → Post	**Pre → Post**	**Pre → Post**	**Pre → Post**	**Pre → Post**	**Notable arc**
A	○ → ●	○ → ●	○ → ●	○ → ◑	○ → ●	Demand gaze → full metafunctional analysis
B	○ → ●	○ → ◑	○ → ●	○ → ○	○ → ◑	'La imagen como texto visual'
C	○ → ◑	○ → ◑	○ → ◑	○ → ●	○ → ○	Curved trunk → female body (independent)
D	○ → ○	○ → ○	○ → ○	○ → ○	○ → ○	Language barrier — methodological case
E	○ → ●	○ → ●	○ → ●	○ → ●●	○ → ●	autonomía pre-theory → war omission post-film
F	○ → ○	○ → ○	○ → ○	◑ → ○	○ → ○	Small/Scared pre-theory; narrative drift post-film
G	○ → ●	○ → ◑	○ → ●	○ → ◑	○ → ●	las palabras pre-theory → colour as meaning
Pre-theory	0/7 metalanguage · 0/7 textual awareness (title excepted) · 0/7 gaze identified · 0/7 gender/ideological reading · 0/7 cross-poster semiotic comparison					
Post-theory	6/7 metalanguage · 5/7 textual awareness · 5/7 gaze identified · 4/7 gender/ideological reading · 4/7 cross-poster semiotic comparison					

Note: Student F's Gender/ideological cell shows ◑ → ○, reflecting a pre-theory near-miss (the annotation Small, Scared / Larger, empowered on the paired poster image) that did not develop into semiotic analysis post-theory. Student D is included in the table for completeness but excluded from aggregate counts (see Section 4.2).

5. DISCUSSION AND CONCLUSIONS

The study set out to investigate whether explicit multimodal instruction transforms L2 learners' capacity to read film posters as semiotic texts. The results confirm that explicit theoretical scaffolding, like Halliday's SFL metafunctions extended visually by Kress and van Leeuwen's Visual Grammar, produces a qualitative transformation in students' capacity to read film posters as multimodal texts. This transformation is not merely terminological: it represents a reconceptualisation of the image as a social semiotic object that actively constructs meaning, consistent with recent findings in multimodal literacy pedagogy (Boysan & Kızıltan, 2023; Budiastono & Noverino, 2024). Particularly telling in this regard is the pre-theory profile of the group's native Spanish speaker: her responses were indistinguishable from those of her L2 peers: denotative nouns, affective adjectives, zero semiotic metalanguage. Full linguistic competence in the language of instruction does not, in itself, generate multimodal analytical capacity. This finding isolates the contribution of the framework with unusual clarity: what changed between the pre-theory and post-theory phases was not linguistic access but conceptual equipment.

The most significant finding from a pedagogical design perspective is that theory alone does not activate the deepest gender and ideological readings: it is the combination of semiotic framework and narrative knowledge, theory plus film, that makes these readings possible. Students who produced the most sophisticated analyses did so in Task 3, after watching the film. This suggests a dialogic rather than linear instructional sequence, resonating with Contreras de la Llave and Pastor Cesteros's (2023) argument for the integration of structured audio-visual analysis within L2 curricula as a tool for developing critical and intercultural competence. The gender readings that emerged, the protagonist's agency encoded differently across the two poster versions; the tree as a representation of the female body; the promotional omission of the film's war narrative, connect the corpus to the broader dynamics of gender representation in Spanish cinema documented by de Miguel Casado and García Carrizo (2025), and confirm that critical gender literacy requires explicit scaffolding rather than mere encouragement.

Several limitations must be acknowledged. The sample is small (N = 7) and the corpus limited to one film and two poster versions, which constrains the generalisability of the findings. Student D's case raises the question of double cognitive load when metalanguage must be expressed in the L2. Two further methodological limitations warrant acknowledgement. First, the absence of a control group means that causal claims about the specific contribution of the theoretical intervention versus the film viewing experience cannot be fully separated, a constraint that future designs with larger samples should address (Unsworth, 2024). Second, while this study focused on semiotic metalanguage as its primary analytical lens, the development of linguistic competence, lexical richness, syntactic complexity, and the use of argumentative connectors in L2 Spanish, represents a productive avenue for future research, one that would allow multimodal and linguistic dimensions of L2 development to be studied in tandem.

The pedagogical implications of these findings extend beyond Spanish as an Additional Language. Film poster analysis is a low-cost, high-yield activity, requiring no specialist technology, adaptable to any film in the syllabus, that simultaneously develops visual literacy, critical reading, and intercultural competence. The three-session workshop format tested here is designed to be portable and replicable. Future phases of the project will expand the corpus to increase the sample size, and explore transferability to other languages and educational levels, with greater attention to the scaffolding of gender readings.

A final consideration concerns the choice of corpus. *El laberinto del fauno* celebrates its twentieth anniversary in 2026, and while its canonical status within Spanish cinema curricula makes it a rich object of analysis, with two culturally distinct poster versions, dense intertextual connections to European fairy-tale traditions, and an extended transmedia life encompassing novelisations, storyboard volumes, and cross-genre references (Herrero, 2019), its familiarity to educators should not be mistaken for exhaustibility. The film's layering of a dark fairy-tale narrative over the Spanish Civil War and its aftermath resonates with a broader tradition of using art, story, and visual culture to foster critical dialogue about historical violence and memory (Martínez-Carratalá & Miras, 2023). Future iterations of this project will extend the corpus to posters by directors such as Icíar Bollaín, Isabel Coixet, Pedro Almodóvar, Alejandro Amenábar, Arantxa Echevarría, Javier Ambrossi and Javier Calvo, amongst others, examining whether the semiotic transformations documented here hold across different visual registers, promotional cultures, and gender encodings (Holguín Vaca & Castaño Arques, 2024). The framework is designed to be portable: any film poster of sufficient semiotic density can serve as an analytical object.

This study began with an image: a girl, a tree, a monster, a darkness, and asked what it would mean to look more carefully. The film's closing voiceover offers a resonant answer: *"She was loved by her people and she left behind small traces of her time on earth, visible only to those who know where to look"* (del Toro, 2006). The final shot shows the dead tree flowering (see Figure 5). It is a precise metaphor for what this project asks of students: not that they see everything, but that they learn to look differently, and in doing so, discover that what they once took for a surface conceals a depth. Like Ofelia, they are being invited to step through a threshold they had not previously perceived.

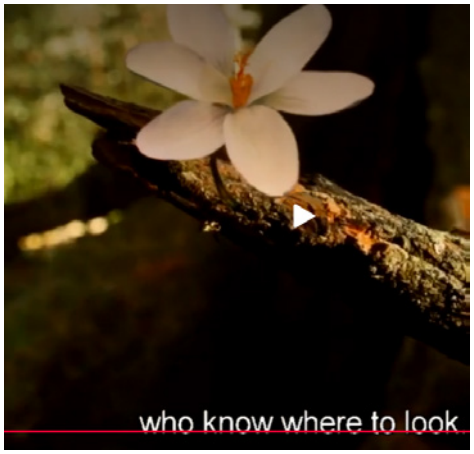


Figure 5. Final shot of El laberinto del fauno (del Toro, 2006): the branch of the dead tree flowers. A metaphor for the analytical transformation this project seeks to produce. © Warner Bros. Entertainment. Screenshot reproduced for scholarly analysis.

REFERENCES

Amenábar, A. (Director). (2009). *Ágora* [Film]. Himenóptero; Mod Producciones; Telecinco Cinema.

Barthes, R. (1977). *Image, music, text* (S. Heath, Trans.). Fontana Press. (Original works published 1961–1971)

Berger, J. (1972). *Ways of seeing*. Penguin / BBC.

Boysan, H. C., & Kızıltan, N. (2023). A semiotic landscape analysis of the film posters: Parasite, Mother and Whiplash. Literacy Trek, 9(3), 85–114. https://doi.org/10.47216/literacytrek.1390732

Budiastono, H., & Noverino, R. (2024). A multimodal critical discourse analysis of The Marvels movie poster. International Journal of English and Applied Linguistics, 4(1), 106–115. https://doi.org/10.47709/ijeal.v4i1.3779

Contreras de la Llave, N. (2019). En defensa del cine en el aula: Por una pedagogía mediática y crítica enmarcada en el currículo educativo. Cuadernos de Pedagogía, 150–153.

Contreras de la Llave, N. (2022). Cine y perspectiva de género como herramientas de transformación social en la didáctica del español como lengua extranjera [Doctoral dissertation, Universidad de Alicante]. RUA Repositorio Institucional. https://rua.ua.es/dspace/handle/10045/138858

Contreras de la Llave, N., & Pastor Cesteros, S. (2023). Alfabetización audiovisual en español como L2 a través del cine para alumnado adolescente en inmersión. Porta Linguarum, VI, 53–68. https://doi.org/10.30827/portalin.viVI.28686

de Miguel Casado, J., & García Carrizo, J. (2025). Representación gráfica de la mujer en el cine español: Análisis gráfico de pósteres de películas nominadas a los Goya desde 2018 a 2023. Grafica, 13(25), 59–69. https://doi.org/10.5565/rev/grafica.341

del Toro, G. (Director). (2006). El laberinto del fauno [Film]. Tequila Gang; Estudios Picasso; Telecinco Cinema.

Halliday, M. A. K. (2004). *An introduction to functional grammar* (3rd ed.). Edward Arnold.

Herrero, C., & Vanderschelden, I. (Eds.). (2019). *Using Film and Media in the Language Classroom: Reflections on Research-Led Teaching*. Multilingual Matters. https://doi.org/10.21832/9781788924498

Herrero, C. (2019). From new literacies to transmedia literacies: The new approaches to transmedia and language pedagogy project. In R. Biasini, N. Becerra, H. Magedera-Hofhansl, & A. Reimão (Eds.), *Innovative Language Teaching and Learning at University: A Look at New Trends* (pp. 19–26). Research-publishing.net.

Holguín Vaca, D. P., & Castaño Arques, A. (2024). "Esta imagen me hizo sentir, pero también pensar": Literacidad visual crítica e interculturalidad en el aula de español como lengua adicional. Íkala, Revista de Lenguaje y Cultura, 29(2). https://doi.org/10.17533/udea.ikala.355938

Jewitt, C., Bezemer, J., & O'Halloran, K. (2016). Introducing multimodality. Routledge.

Khan, M., & Akhtar, R. (2024). Framing femininity: Analyzing visual objectification and gender dynamics in Hollywood movie posters through semiotic and sociocultural lenses. Educational Administration: Theory and Practice, 30(11), 1164–1174. https://doi.org/10.53555/kuey.v30i11.9150

Kress, G., & van Leeuwen, T. (1996). *Reading images: The grammar of visual design*. Routledge.

Kress, G., & van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*. Arnold.

Kress, G., & van Leeuwen, T. (2002). Colour as a semiotic mode: Notes for a grammar of colour. Visual Communication, 1(3), 343–368. https://doi.org/10.1177/147035720200100306

Martínez-Carratalá, F. A., & Miras, S. (2023). Poetry, Art and Historical Memory: Fostering Dialogue about War. Bookbird: Journal of International Children's Literature, 61(3), 36-44. https://doi.org/10.1353/BKB.2023.A903438

Mulvey, L. (2013). Visual pleasure and narrative cinema. In C. Penley (Ed.) *Feminism and film theory* (pp. 57–68). Routledge. (Original work published 1975)

Navarro, A. L. N., & Adriano, T. Q. (2023). A multimodal discourse analysis on female representations in film posters. Journal Corner of Education, Linguistics, and Literature, 4(1). https://doi.org/10.54012/jcell.v4i1.305

Unsworth, L. (2024). Multimodal literacy in a new era of educational technology: Comparing points of view in animations of children's and adult literature. *ECNU Review of Educati*on, 7(2), 384–405. https://doi.org/10.1177/20965311231179738

9. Aprender italiano con la Inteligencia Artificial Generativa: autoaprendizaje y autocorrección

Maria Angelica Giordano Paredes[1], María Soledad Villarrubia Zúñiga[2]

[1]*Universidad Nacional de Educación a Distancia*
[2]*Universidad de Alicante*

RESUMEN

La presente investigación se centra en el autoaprendizaje y la autocorrección dentro de un proyecto de innovación docente diseñado en el Área de Filología Italiana de la UNED, en el marco institucional del Vicerrectorado de Innovación Educativa y el Instituto Universitario de Educación a Distancia (2025-2026). Los objetivos se centran en el estudio de la autonomía en el aprendizaje del italiano, a través de las posibilidades que ofrece la inteligencia artificial generativa (IAGen, en adelante). El trabajo se llevó a cabo en el primer cuatrimestre del curso 2025-2026. El resultado obtenido, una muestra del 46,35% del total del estudiantado matriculado en la asignatura, evidencia la efectividad de la aplicación de la IAGen en el ámbito lingüístico. Se utiliza una práctica basada en el microaprendizaje con el diseño de microactividades dirigidas a la autocorrección en la relación alumnado-plataformas IAGen y se trabaja la autoevaluación a través de una reflexión final del aprendiente sobre el proceso de adquisición de competencias digitales y lingüísticas que concluyen con el autoaprendizaje de la lengua italiana, nivel A2. Los resultados alimentan las hipótesis sobre las expectativas de esta tecnología para potenciar y mejorar el aprendizaje de una lengua extranjera. Se trata solo del inicio de un largo proceso de adquisición lingüística en el que el docente, el discente y la IAGen pueden conformar un excelente equipo de trabajo.

PALABRAS CLAVE: inteligencia artificial generativa, aprendizaje de lenguas, lengua italiana, autocorrección, autoaprendizaje.

1. INTRODUCCIÓN

En cuanto a la cuestión específica objeto de este estudio, el trabajo se centra en el auge de la IAGen en el contexto educativo, abriendo nuevas expectativas metodológicas en el aprendizaje autónomo y a distancia, especialmente, en el marco de la educación superior y en el aprendizaje de lenguas. Las plataformas como *ChatGPT*, *Gemini* o *Microsoft Copilot*, entre otras, permiten generar una retroalimentación inmediata y personalizada en entornos de enseñanza atractivos y accesibles. Por eso, se convierten en recursos muy útiles en el aprendizaje de lenguas para el desarrollo de las competencias lingüísticas y metacognitivas. Estos entornos de apoyo basados en IAGen ofrecen oportunidades para la práctica individual, la corrección de errores, la reformulación textual, las simulaciones y la autorreflexión sobre el proceso de adquisición de lenguas.

La enseñanza universitaria a distancia y semipresencial, como es el caso de la UNED, se enfrenta a desafíos particulares relacionados con la autonomía del alumnado, la autorregulación y la necesidad de retroalimentación constante, necesarios para compensar la ausencia del contacto personal de la educación presencial. El alumnado *online* desarrolla gran parte de su aprendizaje de manera autónoma e independiente. Por esto, resulta relevante explorar las metodologías que pueden reforzar dos aspectos especialmente sensibles: el autoaprendizaje y la autoevaluación. En este marco descrito, es donde la IAGen puede actuar como mediadora entre el estudiante y los contenidos y objetivos, para

facilitar los procesos de autocorrección y autorregulación. La IAGen puede ofrecer apoyo inmediato y constante, además de orientaciones personalizadas que complementan al docente sin sustituirlo.

Los estudios más recientes documentan resultados muy prometedores sobre la integración de la IAGen en el aprendizaje de lenguas. Algunos señalan los beneficios relacionados con la personalización del aprendizaje (Cisneros et al., 2024; Holmes et al., 2023; López et al., 2023); otros se centran en el incremento de la motivación (Villamar et al., 2024) o en la retroalimentación inmediata y el fortalecimiento de la autonomía del estudiante (Montenegro et al., 2025; Rodríguez, 2025); algunos estudian la personalización del aprendizaje (Arévalo et al. 2025; Armijos et al., 2025) y el impacto cognitivo (Qu et al., 2025). Así mismo, hay investigaciones orientadas específicamente en la autocorrección y la autoevaluación mediante IAGen (De la Torre, 2025). Estas últimas muestran que estas herramientas pueden favorecer el proceso de revisión textual y reflexión metalingüística, siempre que exista una mediación pedagógica adecuada y una correcta alfabetización respecto de los *prompts*.

No obstante, todavía existen ciertas dudas sobre el grado de fiabilidad lingüística de la IAGen, el riesgo de la dependencia tecnológica del estudiantado y la necesidad del acompañamiento docente. Por esta razón, hay estudios que advierten sobre la valoración positiva de la retroalimentación automática del alumnado, pero también, de la necesidad de orientación del profesorado para validar los contenidos, interpretar las correcciones y consolidar el aprendizaje (Camino, 2024; Mora, 2025).

Partiendo del marco descrito, la investigación que se presenta analiza la implementación de microactividades lingüísticas mediadas por IAGen para la asignatura *Lengua Moderna I. Lengua Extranjera: Italiano* (A1-A2) impartida en el Grado de Lengua y Literatura Españolas de la Facultad de Filología de la UNED. El trabajo se centra concretamente en dos dimensiones que ya se han comentado: la autocorrección y el autoaprendizaje entendidos como elementos básicos para promover la autonomía del estudiantado universitario en una universidad a distancia. Los objetivos de la investigación son varios: i) analizar el potencial de la IAGen educativa como instrumento de autocorrección en el aprendizaje inicial del italiano como lengua extranjera; ii) explorar la contribución de la IAGen en el desarrollo del autoaprendizaje y la autoevaluación; iii) identificar la percepción del alumnado sobre la utilidad, las limitaciones y el grado de aplicabilidad en su proceso de aprendizaje de la lengua italiana a través de estas herramientas.

Como hipótesis inicial, se plantea la incorporación estructurada de plataformas de IAGen en actividades de microaprendizaje. Se tiene en cuenta que estas pequeñas píldoras de aprendizaje favorecen los procesos de aprendizaje con una autoevaluación y autocorrección consciente que mejora la percepción de la autonomía del alumnado. Esto tendrá, finalmente, un impacto positivo en el aprendizaje lingüístico en contextos universitarios de educación a distancia, contribuyendo también a la motivación.

2. MÉTODO

El presente estudio muestra los resultados específicos centrados, exclusivamente, en el autoaprendizaje y la autocorrección dentro de un proyecto de innovación docente diseñado por el Área de Filología Italiana de la UNED, en el marco institucional del Vicerrectorado de Innovación Educativa y el Instituto Universitario de Educación a Distancia (2025-2026). Para su desarrollo se adopta un MÉTODO mixto, entendido en el sentido propuesto por autores como Creswell (2014) y Creswell y Plano (2017), referente a la integración de enfoques cuantitativos y cualitativos para conseguir una comprensión completa del elemento estudiado. La parte (Yucar y Bernedo, 2020) se orienta hacia la medición y el análisis de cuestionarios al inicio y final de primer cuatrimestre del curso académico,

periodo en el que se implementa el proyecto. La fase cualitativa, siguiendo la perspectiva interpretativa presentada por Flick (2015), se compone de las respuestas abiertas de los estudiantes sobre el procedimiento y los resultados de la experiencia académica, además de las evidencias y reflexiones sobre las actividades propuestas.

El proyecto se implementa en el primer cuatrimestre del curso académico 2025-2026 y se evalúa en la convocatoria ordinaria de febrero dentro de las actividades de evaluación continua. Las microactividades del proyecto representan el 10% de la evaluación total de la asignatura.

A continuación, se detalla el contexto y el perfil de los participantes, los instrumentos, así como el procedimiento para la recogida de datos.

2.1. Descripción del contexto y de los participantes

El proyecto se desarrolla en la plataforma Ágora de la UNED, específicamente en el curso virtual de la asignatura *Lengua Moderna I. Lengua Extranjera: Italiano (Multigrados)* del Grado en Lengua y Literatura Españolas de la Facultad de Filología. La participación del estudiantado es voluntaria y se realiza a distancia, sin grupo de control debido a las características propias de la UNED. Participan 70 estudiantes de los 151 matriculados en la asignatura (46.35%). El alumnado pertenece a tres facultades de Humanidades: Grado de Filología, Grado en Geografía e Historia y del Grado en Filosofía, lo que caracteriza la asignatura como educación para multigrados.

El perfil es muy heterogéneo en cuanto a edad, intereses y formación académica: hay estudiantes que cursan este grado como segunda carrera; hay jóvenes que eligen la UNED como primera opción. Además, se destacan algunos estudiantes con una fuerte vocación sobre esta asignatura. Esta diversidad determina un rango de edad amplio (desde los veinte hasta los setenta años). Las diferencias generacionales constituyen una de las principales características de la asignatura y, en general, de las titulaciones de la UNED, una institución que asegura el acceso a la formación a lo largo de la vida. Los niveles lingüísticos, siguiendo lo establecido en el *Marco Común Europeo de Referencia* (*MCER*), también son variados y responden a los perfiles descritos, aunque el nivel determinado en la asignatura corresponde al A2 al finalizar el cuatrimestre.

2.2. Instrumentos

El diseño de la propuesta de innovación docente se elabora a partir de las herramientas de IAGen creadas por el Vicerrectorado de Innovación Educativa de la UNED y puestas a disposición del profesorado. Son los *Edubots de Inteligencia Artificial en Educación* o *chatbots* de IAGen para la enseñanza y el aprendizaje, integrados en el *HUB de innovación educativa* (https://hubinnovacion.uned.es/inicio).

Para este proyecto se utilizó un *chatbot* que permitía la interacción con los estudiantes en lengua italiana. Este se compartió con el estudiantado del proyecto a través de un enlace en el curso virtual. Además, los participantes pudieron optar entre otras aplicaciones de IAGen, entre ellas *ChatGPT*, *Gemini* o *Copilot* de Microsoft. No obstante, la más utilizada fue el *ChatGPT* en sus diferentes versiones.

Se emplearon dos cuestionarios tipo Likert, de elaboración propia, uno inicial y otro final, diseñados para conocer, respectivamente, los conocimientos y expectativas previas del alumnado y su valoración de la experiencia tras la aplicación de la IAGen. En el cuestionario inicial se incluyeron *ítems* sobre el conocimiento, uso previo de la inteligencia artificial, utilidad percibida para el aprendizaje de lenguas, las expectativas ante el proyecto y sus posibles aplicaciones en el aprendizaje del italiano, así

como preguntas abiertas para recoger las percepciones del alumnado sobre posibilidades y limitaciones de la IAGen en las diferentes competencias lingüísticas. El cuestionario inicial tiene nueve *ítems*, pero para este estudio se analizaron solo los *ítems* 2, 3, 4, 6, 7 y 9. El cuestionario final, sin embargo, se centró en la valoración de la experiencia de aprendizaje y, principalmente, en los procesos de autocorrección, autoevaluación y autonomía con escalas de valoración y preguntas sobre la satisfacción con el proyecto, la utilidad de las correcciones de la IAGen, la interacción con la herramienta, el logro de los objetivos personales, el nivel final de aprendizaje y la autoevaluación. El cuestionario final consta de diez *ítems*, de los cuales solamente se analizaron tres para esta investigación: 2, 4 y 9. Ambos cuestionarios se administraron telemáticamente.

2.3. Procedimiento

El proyecto se implementó a partir de una propuesta de diez microactividades para desarrollar la competencia digital, fomentar la autoevaluación y la autocorrección así como el autoaprendizaje de la lengua italiana en los niveles A1-A2 (*MCER*). Cada microactividad se relacionaba con una unidad del manual básico de la asignatura, hasta un máximo de diez unidades y microactividades integradas en las Pruebas de Evaluación Continua (PEC). Esta propuesta se desarrolló en el curso virtual de la asignatura de italiano, como ya se ha mencionado.

En la Tabla 1 se muestra la correlación entre las microactividades propuestas y las unidades del manual de la asignatura. La finalidad fue facilitar la práctica de los contenidos estudiados e impulsar la autocorrección mediante el uso de las aplicaciones de IAGen.

Tabla 1. Propuesta de Microactividades

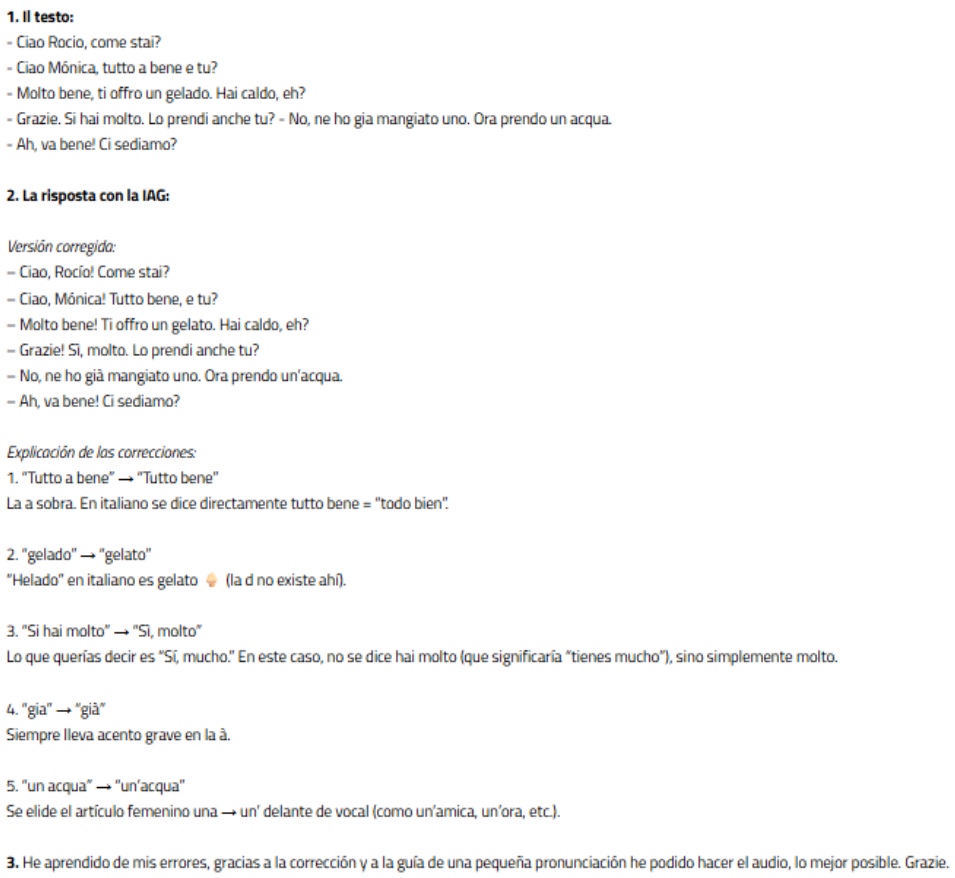
Nº	Título de la microactividad	Lección correspondiente	Modalidad y observaciones metodológicas
1	Mi presento.	Ciao, mi chiamo…	Actividad escrita de presentación.
2	Microdialogo al bar.	Tu che cosa prendi?	Actividad oral que parte de un texto escrito y autocorregido.
3	Com'è la tua giornata tipo?	Com'è la tua giornata?	Actividad escrita de descripción de acciones habituales.
4	La mia città	Che bella città!	Actividad escrita de descripción del entorno urbano.
5	Che cosa avete fatto l'estate scorsa?	Generazione Erasmus	Actividad escrita de acciones pasadas.
6	Cosa metti in valigia?	Acquisti per ogni occasione	Actividad oral que parte de un texto escrito y autocorregido.
7	Come sarebbe la casa dei tuoi sogni?	Coinquilino cercasi	Actividad escrita de descripciones hipotéticas y deseos.
8	Ricordi e descrizioni della famiglia	Non c'era una volta	Actividad escrita sobre recuerdos y descripción de la familia.
9	Raccontare i vostri progetti di futuro	Che cosa farai e dove andrai?	Actividad oral que parte de un texto escrito autocorregido.
10	Consigli per una vita sana	Facciamo un po' di sport?	Actividad oral sobre consejos y recomendaciones.

Fuente: Elaboración propia.

El proceso de autocorrección se llevó a cabo individualmente en una primera fase, para luego, posteriormente, compartirse en el foro del curso virtual donde los estudiantes contactaban para generar así una dinámica colaborativa. Esta se desarrollaba siguiendo el siguiente procedimiento: a) Elaboración de un microtexto escrito con los conocimientos adquiridos de la lengua italiana; b) Corrección del texto con diferentes aplicaciones de IAGen; c) Reelaboración del microtexto con las sugerencias de la IAGen y el empleo de la autocorrección según los conocimientos adquiridos; d) Reflexión de resultados de aprendizaje y puesta en práctica del autoaprendizaje con análisis crítico y argumentativo del *feedback* aportado por la IAGen.

La Figura 1 muestra el proceso de autocorrección que siguió el estudiantado en cada microactividad. En primer lugar, cada estudiante elaboraba el microtexto con sus propios conocimientos de la lengua italiana para luego corregirlo a través de la interacción con una aplicación de IAGen. A continuación, la herramienta proporcionaba los resultados de la corrección que servían de *feedback* para la reflexión sobre el autoaprendizaje. Finalmente, el proceso se concluía con una reflexión crítica y personalizada sobre los resultados de autoaprendizaje obtenidos en la autocorrección.

El ejemplo siguiente (Figura 1) revela una corrección pormenorizada de las estructuras y del léxico que aportaban conocimientos y orientaban al alumnado para que pudiera seguir avanzando y mejorando en el proceso de adquisición del italiano.



1. Il testo:
- Ciao Rocio, come stai?
- Ciao Mónica, tutto a bene e tu?
- Molto bene, ti offro un gelado. Hai caldo, eh?
- Grazie. Si hai molto. Lo prendi anche tu? - No, ne ho gia mangiato uno. Ora prendo un acqua.
- Ah, va bene! Ci sediamo?

2. La risposta con la IAG:

Versión corregida:
– Ciao, Rocío! Come stai?
– Ciao, Mónica! Tutto bene, e tu?
– Molto bene! Ti offro un gelato. Hai caldo, eh?
– Grazie! Sì, molto. Lo prendi anche tu?
– No, ne ho già mangiato uno. Ora prendo un'acqua.
– Ah, va bene! Ci sediamo?

Explicación de las correcciones:
1. "Tutto a bene" → "Tutto bene"
La a sobra. En italiano se dice directamente tutto bene = "todo bien".

2. "gelado" → "gelato"
"Helado" en italiano es gelato 🍦 (la d no existe ahí).

3. "Si hai molto" → "Sì, molto"
Lo que querías decir es "Sí, mucho." En este caso, no se dice hai molto (que significaría "tienes mucho"), sino simplemente molto.

4. "gia" → "già"
Siempre lleva acento grave en la à.

5. "un acqua" → "un'acqua"
Se elide el artículo femenino una → un' delante de vocal (como un'amica, un'ora, etc.).

3. He aprendido de mis errores, gracias a la corrección y a la guía de una pequeña pronunciación he podido hacer el audio, lo mejor posible. Grazie.

Figura 1. Ejemplo del proceso de autocorrección de una microactividad

Fuente: Curso virtual de Ágora: https://agora.uned.es/mod/forum/discuss.php?d=412721#p1240090.

Lo mismo se observa en la Figura 2. El *ChatGPT* hace la corrección de la microactividad y el estudiante reflexiona sobre esta. Así se evidencia la adquisición del aprendizaje a través de la identificación y análisis de los errores. Este ejemplo muestra los pasos 2 y 3 en el proceso de elaboración de las microactividades.

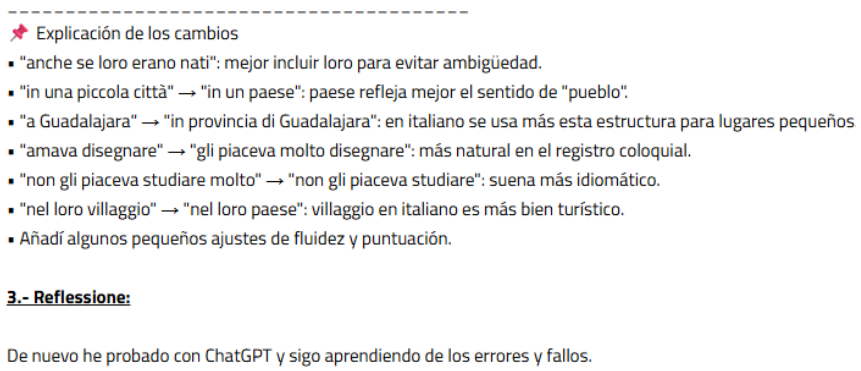


Figura 2. Ejemplo de corrección y reflexión crítica
Fuente: Curso virtual de Ágora: https://agora.uned.es/mod/forum/discuss.php?d=412969#p1338865.

3. RESULTADOS

El análisis de los cuestionarios inicial y final arrojan información detallada sobre los conocimientos previos del estudiantado en relación con el uso educativo de las aplicaciones de IAGen y las competencias desarrolladas tras la experiencia de inmersión en el proyecto.

3.1. Resultados del cuestionario inicial

En el cuestionario inicial [*ítem* 2. ¿Cómo valoraría el uso educativo de la IAG?] se refleja el desconocimiento tanto de la herramienta de IAGen como de sus posibles aplicaciones didácticas. Sin embargo, las respuestas son moderadamente optimistas: el estudiantado evita unas posturas extremas, manteniéndose en una línea equilibrada y prudente. Así, el 13.3% afirma que su uso es 'aceptable'; el 56.7% sostiene que es 'bueno'; el 23.3% manifiesta que es 'muy bueno'. Solo el 6.7% opina que es 'excelente'. Es interesante observar que ninguno lo califica de 'malo'. No obstante, en la siguiente pregunta [*ítem* 3. ¿Considera que la IAG es útil para aprender lenguas?], las respuestas se polarizan ligeramente: el 64% opina que sí y el 36% considera que no.

Estos resultados descubren una tendencia general hacia una actitud positiva, aunque sin un conocimiento profundo ni una firme convicción. Esta idea se confirma en las respuestas del *ítem* 4 [¿Para qué ha utilizado la Inteligencia Artificial?]. El 52% del alumnado confirma usar la IAGen para buscar información y el 9% para otros usos no especificados y fuera del ámbito lingüístico. En cuanto al uso educativo y lingüístico, el 26% afirma que la emplea para mejorar textos escritos y el 13% para traducir textos. El análisis conjunto de estas respuestas evidencia que una mayoría (el 61%) utiliza la IAGen con fines no educativos ni de aprendizaje lingüístico, mientras que solo el 39% reconoce haber vinculado su uso con la adquisición de idiomas, concretamente, en tareas de redacción y traducción.

A pesar de no tener experiencia con el uso de la IAGen para el aprendizaje de lenguas, como se demuestra en los primeros *ítems* analizados, el alumnado mantiene una expectativa muy alta respecto al proyecto. Así lo manifiesta en el *ítem* 6. [¿Considera que este proyecto podría mejorar el aprendizaje de la lengua italiana?]. El 73% de los informantes responde afirmativamente frente al 27% que lo hace negativamente. Estos datos, de nuevo, muestran que el estudiantado tiene una visión muy positiva del proyecto y lo acoge con interés y motivación. Esta actitud es, sin duda, un factor esencial para el éxito de cualquier iniciativa de innovación docente como esta.

Sobre el *ítem* 9 [Cree que el uso de la IAGen podrá…] se ofrece una respuesta basada en cuatro opciones: a) Aprender más vocabulario; b) Aprender la lengua con más facilidad; c) Escribir mejor en italiano; d) Mejorar la pronunciación. En el Gráfico 1 se representan los resultados obtenidos: el 26% identifica las posibilidades de aprender más léxico y el 22% cree que podría mejorar la pronunciación. Resulta especialmente llamativo que el 39% opine que la IAGen podría facilitarle el aprendizaje de la lengua italiana, y que solo el 13% esperaría mejorar la escritura. Estos resultados son alentadores si se tiene en cuenta que estos estudiantes no habían tenido ninguna experiencia lingüística previa con herramientas de IAGen.

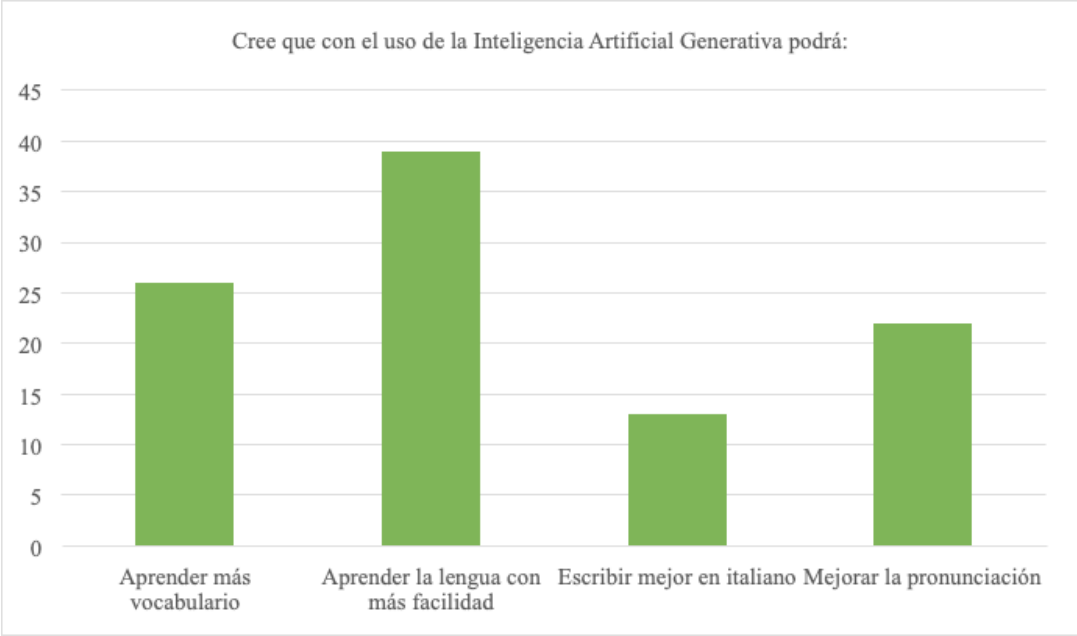


Gráfico 1. *Ítem* 9 del cuestionario inicial
Fuente: Elaboración propia a partir del cuestionario inicial.

Del estudio cualitativo del cuestionario inicial, y en relación con el *ítem* 7 [¿Considera que la inteligencia artificial generativa podría mejorar todas las competencias?], se seleccionaron diez respuestas, entre las más representativas por su variedad y relevancia. E1: 'No lo sé; seguramente sí, pero nunca lo he aplicado a todas. No sé cómo me puede ayudar en el oral'. E2: 'Sí, sería una herramienta de gran utilidad'. E3: 'Considero que no, porque en aspectos más filológicos no creo que sea óptimo el uso de la IAG'. E4: 'Estoy convencido de que esta tecnología es positiva'. E5: 'No sé, nunca la he usado'. E6: 'Sí, pero solo si se conocen sus limitaciones'. E7: 'Sí, es un medio de aprender tanto cuestiones gramaticales como de pronunciación'. E8: 'Si se refieren a las compe-

tencias lingüísticas, creo que puede aportar al aprendizaje, pero mejorar todas las competencias de una persona parece poco realista'. E9: 'Sí, aunque creo que la parte oral tendría más dificultades a la hora de mantener una conversación'. E10: 'Veo especiales limitaciones para la producción oral'.

De las respuestas se deduce que el estudiantado percibe el potencial lingüístico de la IAGen, aunque mantiene reservas respecto a su eficacia en la producción oral. Esta cautela puede deberse a que, en la actualidad, esta es una de las áreas tecnológicas menos desarrolladas. En conjunto, las respuestas muestran una buena predisposición e incluso cierto entusiasmo, lo que constituye un elemento favorable para el éxito en el autoaprendizaje. La aceptación inicial y la actitud positiva sugieren que el uso de la IAGen puede contribuir de manera significativa a la mejora de su proceso de aprendizaje autónomo.

3.2. Resultados del cuestionario final

A continuación, se analizarán los datos cuantitativos del cuestionario final una vez finalizado el proceso de inmersión lingüística. Se han tenido en cuenta solamente los *ítems* relacionados con el autoaprendizaje y los resultados después del proceso de autocorrección y autoevaluación de las microactividades propuestas, ya que es esta la finalidad de la investigación.

Así, en el *ítem* 2 [¿Cómo valoraría su satisfacción general con el proyecto?] los estudiantes respondieron en función de una escala de valoración del 1 al 5, donde el 5 representa el nivel más alto (excelente) y el 1 el más bajo (deficiente/muy malo). Respecto a los resultados, la opción 1 obtiene un 4%; la opción 2 registró un 0%; el 3 alcanzó un 4%; mientras que las categorías 4 y 5 concentraron las mayores puntuaciones, con un 50% y un 42% respectivamente. Por tanto, la mayoría de los discentes dio una puntuación alta del proyecto otorgándole las mayores calificaciones: un 92% del estudiantado se sitúa entre los valores más altos [4 y 5], correspondiente a las categorías de 'muy buena' y 'excelente', lo que revela un elevado nivel de satisfacción con el proyecto y la eficacia del aprendizaje.

En el *ítem* 4, se valoran puntos importantes relacionados con los objetivos de la investigación: Opción 1 ['La inteligencia artificial generativa ofrece una amplia gama de posibilidades para el aprendizaje efectivo de la lengua italiana']. El 69.2% manifiesta estar de acuerdo; el 15.4% está totalmente de acuerdo; mientras que el 15.4% no está ni de acuerdo ni en desacuerdo. La mayoría (84.6%) manifiesta una vez más su satisfacción con los resultados del proyecto, mientras que una minoría no termina de estar seguro o tener una idea clara de los objetivos de este. En la Opción 2 [El aprendizaje con las correcciones de la Inteligencia Artificial Generativa ha sido mucho más provechoso]. En este caso, el 53.8% está de acuerdo; el 23.1% está totalmente de acuerdo; el 19.2% no está ni de acuerdo ni en desacuerdo y, finalmente, el 3.8% manifiesta su total desacuerdo. De nuevo, en la contestación obtenida de los estudiantes se aprecia un alto grado de satisfacción con un porcentaje total del 76.9%.

En la Tabla 2, se presenta el alto grado de satisfacción de los estudiantes con los resultados de aprendizaje. Se valora positivamente la mayor parte de los descriptores y se deja un margen muy reducido a la insatisfacción. Tanto la autocorrección como la autoevaluación resultan aspectos muy favorecidos. Las categorías 'muy satisfecho' y 'totalmente satisfecho' superan el 50% en todas las valoraciones. De nuevo, esto confirma una tendencia mayoritaria hacia la satisfacción respecto del proceso y los resultados del proyecto.

Tabla 2. Grado de satisfacción con los resultados de aprendizaje

	Nada satisfecho	Algo satisfecho	Muy satisfecho	Totalmente satisfecho
La IA en el aprendizaje de lenguas	3.8	15.4	50	30.8
Las correcciones de la IA	3.8	11.5	53.8	30.8
La interacción con la IA	3.8	15.4	61.5	19.2
Objetivos personales de aprendizaje	-	7.7	69.2	23.1
Nivel final de aprendizaje del italiano	-	11.5	65.4	23.1
Autoevaluación del propio aprendizaje	3.8	7.7	61.5	26.9

Fuente: Elaboración propia

En cuanto al análisis cualitativo, el cuestionario final refleja las opiniones del alumnado en el *ítem* 9 [¿Cree que la Inteligencia Artificial Generativa es una herramienta efectiva para el correcto aprendizaje de una lengua extranjera y que contribuye al autoaprendizaje y a la autoevaluación del proceso, contribuyendo a la autonomía del estudiante?]. Se hace una selección de diez respuestas, entre las más relevantes en la misma línea de los objetivos de la investigación, como se muestran a continuación. E1: 'En mi opinión, la Inteligencia Artificial Generativa está bien para aprender, pero mediante correcciones, como si fuese un profesor corrigiendo, pero a veces falla si no le especificas muy bien el prompt. Así que, hay que saber cuándo usarla y cómo usarla'. E2: 'Sí, puedes obtener correcciones en cualquier momento de una calidad aceptable'. E3: 'No del todo. No me parece una herramienta fundamental para el aprendizaje del idioma'. E4: 'Sí, muy útil para dudas concretas, ayuda a ser resolutivo de manera autónoma, pero en mi caso sigo necesitando un profesor que me guíe'. E5: 'Es difícil evaluar cómo de fiel es una Inteligencia Artificial a la lengua objetivo cuando se está aprendiendo, por lo que no me atrevería a hacer una valoración completa en este caso. No obstante, por lo que parece sí que ayuda más de lo que esperaba en un principio y puede ser una buena herramienta para explorar'. E6: 'Sí, en mi experiencia ha resultado muy enriquecedora'. E7: 'Sí, si se saben hacer los *prompts* adecuados'. E8: 'Sí, ayuda bastante pero también hay que tener nociones de lo que se pregunta porque también tiene errores'. E9: 'Sí, con seguimiento de profesor'. E10: 'Creo que es un total acierto'.

Como se observa, el análisis cualitativo está en consonancia con el estudio cuantitativo en lo que se refiere al entusiasmo mostrado por los discentes sobre la IAGen y al rendimiento en el autoaprendizaje de lenguas. Por tanto, su implementación resulta ser un apoyo valioso para los procesos de corrección y aprendizaje. Estos resultados, sin embargo, no invalidan al profesor que sigue siendo una figura importante en el aprendizaje lingüístico y, además, confirman que la confianza del estudiantado sobre la IAGen todavía está lejos de ser plena.

Los cuestionarios evidencian dos momentos del proyecto: mientras el inicial refleja un sondeo inicial sobre el conocimiento didáctico de la IAGen por el estudiantado, el final, corrobora la efectividad de la aplicación en los resultados de aprendizaje.

4. DISCUSIÓN Y CONCLUSIONES

Son muchos los estudios desarrollados en los últimos años que analizan la aplicación de la IAGen en el aprendizaje de lenguas para la innovación metodológica y mejora de los procesos formativos, superando

progresivamente los enfoques más tradicionales. Arévalo et al. (2025) llevaron a cabo un trabajo con estudiantes de inglés como lengua extranjera en el que utilizaron diferentes aplicaciones de IAGen, entre ellas el *ChatGPT*, para el desarrollo de las competencias escritas y orales. Los resultados del trabajo demostraron, en la línea del estudio aquí presentado que "la implementación de la IA en la enseñanza del lenguaje fortalece profundamente los procesos de personalización y los resultados académicos" (p. 371).

Además, como en esta investigación, centrado en el estudio del autoaprendizaje y la autocorrección, De la Torre (2025) desarrolla una experiencia lingüística con estudiantes de español como lengua adicional en la que analiza las posibilidades del *ChatGPT 3.5* para promover la autonomía del alumnado mediante procesos de revisión para la mejora textual. Se trata de un estudio especialmente relevante para esta investigación, ya que comparte el foco de la autocorrección asistida por IA como estrategia de aprendizaje autónomo. Bounif (2026), en un trabajo de revisión sistemática sobre los beneficios de la IAGen en el aprendizaje de lenguas, identifica cinco grandes categorías: cognitivo-lingüística, pedagógica, motivación-afectiva, accesibilidad y social-colaborativa. Estas dimensiones se relacionan con el desarrollo de habilidades lingüístico-cognitivas, la adaptación a necesidades individuales de aprendizaje, la reducción de barreras formativas y el fortalecimiento de la interacción colaborativa. Estas categorías también se observan en el presente estudio, en concreto, en la relación con la personalización del aprendizaje, la retroalimentación inmediata y el refuerzo de dinámicas autónomas y colaborativas en el aprendizaje de una lengua extranjera, el italiano. Por tanto, se confirma la línea de estudio y los resultados.

Asimismo, en un estudio cuasiexperimental con estudiantes universitarios de nivel A2 de lengua inglesa, León et al. (2025) confirman la percepción positiva de los estudiantes con la integración de la IAGen en las tareas escrita y orales. Entre los beneficios se destacan: la retroalimentación, la mejora en la organización de ideas y el incremento en la autonomía del alumnado. También identifican limitaciones significativas, como la necesidad de orientación docente y el riesgo de dependencia tecnológica, al igual que en los resultados de esta investigación.

Como limitaciones del estudio, hay que señalar que los resultados obtenidos deben interpretarse en el contexto específico en el que se ha desarrollado la investigación, lo que limita su generalización a otros entornos de enseñanza de lenguas. Además, la constante evolución de las herramientas de inteligencia artificial generativa (IAGen) exige investigar su impacto mediante estudios con muestras más amplias y diversos contextos educativos.

En conclusión, la revisión de estos estudios permite situar la presente propuesta en una línea de investigación emergente que concibe la IAGen no como sustituta del profesorado, sino como herramienta de apoyo al aprendizaje autónomo, que permite la autocorrección y la autorregulación en contextos de enseñanza de lenguas adicionales y en particular, en cursos a distancia donde los discentes tienen mayor autonomía.

5. REFERENCIAS

Arévalo-Barros, J. F., Trujillo-Cumbal, P. M., Ayovi-Ayovi, Y. E., y Jiménez-Pesantez, M. J. (2025). Aplicación de la inteligencia artificial en la personalización del aprendizaje en lenguaje. *Vitalia. Revista Científica y Académica, 6*(2), 354–381. https://doi.org/10.61368/r.s.d.h.v6i2.607

Armijos-Saca, R. A., Gutiérrez-Vargas, M. M., Crespo-Gordillo, B. Z., Espinoza-Agila, L. E., y Púa-Pilay, J. E. (2025). La inteligencia artificial (IA) en la personalización del aprendizaje en el aula. *Ciencia Latina Revista Científica Multidisciplinar, 9*(1), 7590–7601. https://doi.org/10.37811/cl_rcm.v9i1.16427

Bounif, H. (2026). Beneficios de la inteligencia artificial en el aprendizaje de lenguas extranjeras: revisión sistemática. *Horizontes. Revista de Investigación en Ciencias de la Educación, 10*(41), 1–19. https://doi.org/10.33996/revistahorizontes.v10i41.1206

Camino-Herrera, C. A. (2024). Acompañamiento pedagógico mediado por inteligencia artificial: impacto en la retroalimentación docente en tiempo real. *Innovarium International Journal, 2*(3), 1–12. https://revinde.org/index.php/innovarium/article/view/25

Cisneros-Vásquez, E. D., Nevárez-Loza, R. F., Farez-Cherrez, A. M., y Torres-Montes, R. E. (2024). Uso de la inteligencia artificial en la personalización del aprendizaje. *Conocimiento Global, 9*(1), 75–83. https://doi.org/10.70165/cglobal.v9i1.339

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approach* (4th ed.). SAGE.

Creswell, J. W., y Plano Clark, V. L. (2017). *Designing and conducting mixed methods research* (3rd ed.). SAGE.

De la Torre-Sánchez, Á. (2025). Autocorrección en el aula de ELE a través de *ChatGPT* 3.5. *Revista Internacional de Lenguas Extranjeras, 23*, 27–49. https://doi.org/10.17345/rile23.4090

Flick, U. (2015). *INTRODUCCIÓN a la investigación cualitativa* (3.ª ed.). Morata.

Holmes, W., Bialik, M., y Fadel, C. (2023). *Artificial intelligence in education: Promises and implications for teaching and learning* (2nd ed.). Center for Curriculum Redesign. https://curriculumredesign.org/wp-content/uploads/AIED-Book-Excerpt-CCR.pdf

León-Moral, A. P., Feijoo-Rojas, K. J., Rodríguez-Torres, M. M., y Ramírez-Romero, E. V. (2025). Innovación pedagógica con inteligencia artificial generativa para fortalecer las habilidades productivas en inglés en estudiantes universitarios. *Journal of Science and Research, 10*, 1–25. https://doi.org/10.5281/zenodo.17857927

López-López, H. L., Rivera-Escalera, A., y Cruz-García, C. R. (2023). Personalización del aprendizaje con inteligencia artificial en la educación superior. *Revista Digital de Tecnologías Informáticas y Sistemas, 7*(1), 123–128. https://doi.org/10.61530/redtis.vol7.n1.2023.165.123-128

Mora-Zambrano, E. R. (2025). Evaluación automatizada mediante IA: impacto en la objetividad y eficiencia docente. *Revista Ingenio Global, 4*(1), 263–275.

Montenegro-Montenegro, D. E., Saá-Montalvo, J. F., Yance-Tutiven, L. R., y Jerez-Villacrés, P. (2025). Chatbots como herramientas de retroalimentación inmediata para estudiantes en entornos virtuales. *Technology Rain Journal, 4*(1). https://doi.org/10.55204/trj.v4i1.e79

Qu, X., Sherwood, J., Liu, P., y Aleisa, N. (2025). Generative AI tools in higher education: A meta-analysis of cognitive impact. En *Extended abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25)* (Article 302, pp. 1–9). Association for Computing Machinery. https://doi.org/10.1145/3706599.3719841

Rodríguez-Mireles, M. Á. (2025). Aplicación de la inteligencia artificial en la evaluación formativa: un enfoque para la retroalimentación inmediata y mejora del rendimiento académico. *Innovarium International Journal, 3*(1), 1–13. https://revinde.org/index.php/innovarium/article/view/31

Villamar-Ponce, J. L., Ponce-Merino, S. R., Tumbaco-Figueroa, G. P., y Pisco-Rodríguez, L. V. (2024). Inteligencia artificial como catalizador en la motivación y el compromiso académico de estudiantes universitarios. *Serie Científica de la Universidad de las Ciencias Informáticas, 17*(8), 70–85.

Yucra-Quispe, T. y Bernedo-Villalta, L. Z. (2020). Epistemología e investigación cuantitativa. *Igobernanza, 3*(12), 107–120. https://doi.org/10.47865/igob.vol3.2020.88

10. Innovación docente a través de la historia familiar como parte del aprendizaje significativo en el desarrollo personal y la elección profesional[1]

Asunción Lillo-Beneyto, José Emiliano Ramírez-García

Universidad de Alicante

RESUMEN

El artículo que se presenta está enfocado en los procesos formativos, priorizando la actividad del discente como estrategia de autoformación permanente a través de un modelo de formación curricular orientado al desarrollo de objetivos, competencias y resultados para gestionar su propio aprendizaje a lo largo de la vida. Gestado por los docentes de dos asignaturas del Grado de Trabajo Social en la Universidad de Alicante: Trabajo Social con personas y familias y Habilidades Profesionales II. Se pretende identificar y exponer elementos para elaborar las trayectorias vitales de los discentes que generen aprendizaje significativo y les ayude a la comprensión de su desarrollo personal y elección profesional. Para ello se hará uso del genograma, del relato narrativo, de la línea de vida, de la autoevaluación y la coevaluación grupal, y de la exposición . Los resultados confirman que el alumnado se reconoce dentro de su sistema familiar y se autopercibe con los roles adjudicados dentro de su realidad. Destacan experiencias vividas en la infancia como detonante de la elección de la profesión. También han influido las experiencias relacionadas con los cuidados y con los acontecimientos vitales no normativos. Se concluye que la práctica reflexiva colaborativa aporta información novedosa con relación a la resiliencia como ejemplo del ciclo constante de aprendizaje entre vivencias y reconstrucción del conocimiento.

PALABRAS CLAVE: Innovación docente, aprendizaje autorregulado, práctica colaborativa reflexiva, elección profesional.

1. INTRODUCCIÓN

La formación de los discentes (estudiantes) y de los profesores en la educación superior está experimentando continuos cambios. El sistema de aprendizaje por competencias implementado en el Grado de Trabajo Social tras el consenso del Libro Blanco de la ANECA (2004-2005), la reforma del Proceso de Bolonia (RD 1393/2007) y su posterior consolidación en la Universidad de Alicante (curso 2009-2010) establecen un marco regulador que, en la asignatura "Trabajo Social con Personas y Familias", articula tres ejes esenciales: las competencias básicas para el análisis ético y la transmisión de conclusiones, las competencias técnicas para el uso aplicado de herramientas de diagnóstico e intervención familiar (como genogramas o entrevistas), y las competencias transversales orientadas al desarrollo de la empatía relacional y el trabajo interprofesional ante situaciones de crisis. La asignatura de Habilidades Profesionales II se introduce con el nuevo Plan de Estudios en el curso académico 2010-2011 y pretende desarrollar capacidades prácticas para diseñar planes de intervención, aplicar técnicas profesionales a entornos simulados y mantener una conducta ética a través de sus competencias instrumentales, cognitivas y actitudinales, operando de manera complementaria la teoría y la

1 El presente trabajo ha contado con una ayuda del Programa de Redes de investigación en docencia universitaria del Instituto de Ciencias de la Educación de la Universidad de Alicante (convocatoria 2025-2027). Ref.: (6489)

práctica en estas dos asignaturas.

A estos cambios se adiciona la incorporación de plataformas virtuales, inteligencia artificial o recursos multimedia que exigen un desarrollo de nuevas competencias digitales para enseñar, aprender y evaluar. La digitalización del aprendizaje invierte menos recursos en los procesos formativos que priorizan la actividad del discente aprendiendo de su experiencia como estrategia de autoformación permanente (Domingo, 2021; Beroíza-Valenzuela, 2022). Las evaluaciones que se obtenían con la realización de exámenes han sido sustituidas o complementadas por proyectos, portafolios, evidencias empíricas del proceso seguido en el aprendizaje, pero son pocos los trabajos donde la innovación docente parte de la premisa de convertir en expertos de sus propias vidas al estudiantado (Anderson & Gehart, 2007) y de esta manera acercarse a comprender la vida de los otros. En educación, esta idea se traduce en que el estudiantado no es un receptor pasivo de contenidos, sino alguien que posee saberes, experiencias, valores y perspectivas que deben formar parte del proceso de aprendizaje. La dimensión humana es esencial en la profesión que escoge el alumnado en las titulaciones relacionadas con la ayuda y para comprender a las personas hay que explorar en sus vivencias y relaciones. La elección de la profesión, en este caso trabajo social, está unida según Fombuena Valero (2011) a los acontecimientos no normativos que los marcan profundamente en su ciclo vital. Es por ello que consideramos que relacionar los conocimientos nuevos con conocimientos previos permite dar significado a lo aprendido y poder utilizarlo en otras situaciones de la vida (Ausubel, 2002). La estrategia de autoformación permanente que presentamos a través de la elaboración y construcción por parte de cada uno de los discentes del genograma familiar, línea de vida y relato narrativo familiar, es el eje articulador del proceso educativo que constituye este modelo de formación curricular estructurado sistemáticamente y orientado al desarrollo de objetivos, competencias y resultados para gestionar su propio aprendizaje a lo largo de la vida. La estructura curricular teórica del modelo se sustenta en la epistemología circular recurrente, el modelo estructural de Minuchin (1980) y en los principios, conceptos y técnicas del modelo milanés (Selvini et al., 1991). Esta autoformación no es independiente, puesto que en todo momento del proceso el alumnado está acompañado y supervisado por el docente en el aula. Esta experiencia se lleva a cabo desde las dos asignaturas del Grado en Trabajo Social de la Universidad de Alicante: Trabajo Social con Personas y Familias y Habilidades Profesionales II, mencionadas anteriormente, y está conectada con el proyecto "Formación Reflexiva en Trabajo Social: Aprendizaje Participativo e Innovación con familias desde el Modelo sistémico" (Convocatoria 2025.Ref.: 6489).

Las Ciencias del Aprendizaje sostienen que la instrucción efectiva requiere una participación cognitiva activa por parte del estudiante (Willingham, 2021) y coinciden en que la construcción de conocimientos duraderos exige esfuerzo cognitivo, procesamiento profundo de la información, práctica deliberada y la resolución activa de problemas. El aprendizaje depende de una práctica intencionada, deliberada, que está orientada a superar limitaciones y mejorar el desempeño a través del esfuerzo de hacer y no del consumo de un producto acabado (Ericsson et al., 1993). El aprendizaje profundo requiere que el discente participe activamente en la construcción de significados y en la organización del conocimiento (Ausubel, 2002). La educación superior ya no se limita a una etapa específica sino a un aprendizaje a lo largo de la vida que se concibe como un proceso continuo de actualización profesional y personal (Illeris, 2014).

El modelo que se propone se construye desde las necesidades educativas identificadas en la práctica y se fundamenta en la formación basada en competencias y desarrollo profesional (Perrenoud, 2004). El aprendizaje no se limita a la adquisición de objetivos disciplinares, sino que el currículo

se concibe como un espacio de reflexión crítica en el que los discentes interpretan, resignifican y articulan sus experiencias de vida a través de la elaboración de prácticas narrativas (White & Epston, 1993), que les han de servir para comprenderse y explorar sus vivencias y relaciones. En este sentido, las investigaciones recientes ponen de manifiesto la relevancia de la dimensión emocional en los procesos de aprendizaje en Educación Superior, mostrando la relación entre la inteligencia emocional, las emociones de logro y el uso de estrategias de aprendizaje profundo (Nieto et al., 2024), así como la influencia de determinadas metodologías de aprendizaje sobre la inteligencia emocional y las estrategias de aprendizaje de los estudiantes universitarios (Redondo-Rodríguez et al., 2023). Las prácticas reflexivas (Martín, 2014) sirven como elemento articulador de teoría y práctica (Schön, 1998) y llevan a los participantes a ser sujetos activos de su formación, convirtiendo la práctica docente en praxis aristotélica. Por tanto, el objetivo último de esta publicación es identificar y exponer elementos para elaborar las trayectorias vitales de los discentes que generen aprendizaje significativo y les ayuden a la comprensión de su desarrollo personal y elección profesional a través del pensamiento crítico y relacional.

2. MÉTODO

El enfoque del estudio es fenomenológico y pretende comprender cuál es la vivencia individual de los discentes ante una misma actividad docente, así como los significados que atribuyen a dicha experiencia (Castillo-López et al., 2022). El diseño se articula mediante el estudio de caso, lo que permite profundizar en la experiencia de los participantes y en las dinámicas familiares que surgen de sus relatos. Los resultados se analizaron mediante un análisis inductivo que aborda diferentes temas desde un enfoque reflexivo, que más adelante expondremos, orientado a identificar e interpretar patrones presentes en la información cualitativa. El análisis se centra en la información contenida en el genograma y, especialmente, en los significados construidos a partir del relato narrativo familiar. Esto nos permite, desde un enfoque constructivista y sistémico, obtener la experiencia relacional de intercambio a través del proceso de formación (Honoré, 1980).

Se elaboraron dos guías de análisis, la primera entregada al realizar el genograma, antes de elaborar el relato familiar y la segunda al finalizar el proceso; ambas integradas por preguntas orientadoras que permitieron examinar de manera sistemática la información relacional contenida en los genogramas familiares y en los relatos de vida, facilitando la identificación de temas, códigos y categorías durante el análisis temático inductivo.

2.1. Descripción del contexto y de los participantes

El modelo pedagógico se centró por un lado en la importancia de poder ofrecer al estudiante oportunidades para explorar sus emociones, sentimientos, creencias, roles, funciones que desempeñan dentro del sistema familiar y, por otro lado, en aplicar un modelo formativo curricular para generar aprendizaje significativo, identificar patrones relacionales y fortalecer el desarrollo personal y la elección profesional. Esta investigación sistematiza los datos de cinco cursos académicos seguidos con un total de participación de 385 estudiantes que siguieron el desarrollo metodológico del modelo que se construyó como guía para la reflexión, aprendizaje y producción de conocimiento (Lay & McGuire, 2010; Schön,1998) en relación con los siguientes temas: la experiencia personal como vía de autoconocimiento y comprensión relacional. Las experiencias vitales que transforman la identidad y las

relaciones familiares. La reconstrucción de significados. La muestra fue por conveniencia y participación voluntaria del alumnado hasta el curso 2024-2025, ya que no formaba parte de las pruebas de evaluación. En todo momento se preservó el anonimato del genograma.

Tabla 1. Participación desagregada por cursos y sexo

CURSO	HOMBRE	MUJER	TOTAL
2021-2022	10	71	81
2022-2023	1	26	27
2023-2024	9	61	70
2024-2025	10	92	102
2025-2026	10	95	105

2.2. Instrumentos

Los instrumentos utilizados en las distintas fases del desarrollo los presentamos teniendo en cuenta el momento en el que se desarrolla según la experiencia de innovación.

El primer instrumento es el genograma según McGoldrick y Gerson (1985). El ejercicio consiste en la elaboración gráfica de la estructura de la familia del alumnado en tres generaciones. La representación de estos datos permite a los docentes hipotetizar significados relacionales que en el proceso de narración autobiográfica serán validados o refutados por los discentes (Sorrentino, 2019). A continuación, se trazan líneas de colores que sitúan las relaciones con distintos significados: rojo, relaciones intensas; amarillo, relaciones débiles que te acompañan a lo largo de la vida; azul, relaciones conflictivas y verdes, relaciones vividas que han marcado la vida.

El segundo instrumento es el relato narrativo sobre su historia familiar que incluya los acontecimientos que han identificado en la primera fase del modelo. Para Bruner (1991) el relato lo concibe como una estructura mediante la cual las personas organizan y dotan de significado sus experiencias. Aplicando esto a la familia, el relato narrativo familiar recoge cómo sus miembros, en este caso nuestro alumnado, explican, interpretan y transmiten su historia compartida. Siguiendo a Guerra Reyes (2021), este instrumento es útil para evaluar el significado que se le da en el momento actual a ese momento trascendental que sucedió en tu vida.

El tercer instrumento es la línea de la vida con los cinco acontecimientos que se consideran relevantes en su trayectoria vital. El objetivo trata de identificar, por una parte, las experiencias significativas en su sistema familiar de origen que han influido en su desarrollo personal y su elección e identidad profesional y, por la otra, desarrollar las competencias personales y reflexivas.

Asimismo, se utilizaron dos guías estructuradas mediante preguntas orientadoras para apoyar la reconstrucción del relato familiar y la autoevaluación y reflexión de la experiencia. Estas preguntas no eran de obligatoria respuesta y permitieron explorar y relacionar de manera sistemática los datos objetivos y relacionales representados en el genograma con los significados expresados en los relatos.

Tabla 2. Guía de reconstrucción del relato familiar por pasos

Pasos	Preguntas
Reconstrucción descriptiva a partir de los subsistemas que conforman el genograma familiar.	¿Qué generaciones están representadas? ¿Qué fechas, edades, profesiones, parentescos o circunstancias están explícitas en el genograma? ¿Señala todos los subsistemas que consideras importantes y el tipo de relación que hay entre ellos?
Identificar roles, funciones, límites y vínculos de los distintos subsistemas.	¿Cuáles son los roles y funciones que desempeñan los miembros de cada subsistema que has elegido? ¿Entre los subsistemas cómo son los límites? ¿Qué tipo de vínculo, crees que tienen? ¿Qué relaciones son cercanas, distantes, conflictivas, ambivalentes o interrumpidas?
Comunicación y tipo de reglas.	Explica si la comunicación que más prevalece en los subsistemas elegidos es analógica, digital, complementaria o simétrica. Detecta las diadas, las coaliciones o triangulaciones existentes y explica su significado. Señala el tipo de reglas explícitas, implícitas, meta reglas (secretas) que crees que se dan entre los subsistemas.
Identificación de patrones de comportamiento, lealtades, creencias y mitos transgeneracionales	¿Se observan patrones de proximidad, conflicto, cuidado, abandono, pérdidas, migraciones, separaciones? ¿Qué creencias con carga emocional parecen transmitirse entre generaciones? ¿Qué roles o responsabilidades se repiten? ¿Qué formas de relacionarse aparecen en diferentes generaciones? ¿Qué patrones se mantienen y cuáles se transforman? ¿Qué convicciones compartidas por todos los miembros de la familia, han creado la identidad de la misma? ¿Puedes detectar lealtades visibles e invisibles dentro del sistema familiar? ¿Qué dirías de ellas?
Conexión dato-relato	¿Qué acontecimientos vitales han marcado tu vida? ¿Qué fechas se relacionan con los acontecimientos vitales seleccionados? ¿Cómo crees que puedes relacionar lo representado gráficamente con lo narrado? ¿Qué acontecimientos parecen explicar determinadas relaciones? ¿Qué elementos del genograma adquieren significado cuando se contrastan con el relato?
Construcción de significados relacionales	¿Qué significado relacional atribuyes a cada uno de los acontecimientos elegidos? ¿Qué relación temporal y emocional crees que existe entre los acontecimientos y los vínculos que estaban creados y los cambios que pudieron ocasionar? ¿Qué consideras importante, problemático, doloroso o protector?
Tensiones y contradicciones	¿Existen discrepancias entre el genograma y el relato? ¿Qué información aparece en el genograma, pero no crees que es relevante para el relato, y viceversa? ¿crees que describes un vínculo de manera diferente a como podría deducirse del gráfico? ¿Hay relaciones simultáneamente protectoras y conflictivas? ¿Qué ambivalencias aparecen? ¿Qué aspectos permanecen sin explicar?¿Crees que son importantes para el futuro?

Tabla 3. Instrumento de autoevaluación y reflexión sobre la experiencia

Recuerdos del inicio del proceso.	¿Qué esperabas encontrar al realizar el trabajo que se proponía? ¿Qué significado pensaste que tenía el reconstruir tu historia familiar a partir de elaborar el gráfico del genograma?
Durante el proceso de elaboración del genograma	¿Qué experimentaste mientras reconstruías tu genograma? ¿Qué fue lo que más llamó tu atención? ¿Crees que hubo algún momento que te resultara significativo? ¿Puedes compartirlo? ¿Aparecieron momentos donde los recuerdos, emociones o pensamientos te invadieron? ¿Puedes señalarlos? ¿Hubo algo que pudieras observar en el genograma que no habías percibido antes?
Significados construidos	¿Qué significado tiene para ti lo que has descubierto durante este proceso? ¿Cambió tu manera de comprender alguna relación familiar? ¿Cómo? ¿Hay alguna experiencia familiar que ahora interpretas de manera diferente?¿Qué relaciones o acontecimientos adquirieron un significado diferente al reconstruirlos?
Aprendizajes y cambios	¿Qué te llevas de esta experiencia? ¿Qué has aprendido sobre tu familia o sobre las relaciones familiares? ¿Hay algo que ahora puedas mirar de una manera diferente? ¿Encontraste conexiones entre acontecimientos familiares y determinadas relaciones que anteriormente no habías considerado? ¿Hubo algo que te sorprendiera al observar tu historia familiar en conjunto? ¿Qué ha significado para ti, hacer este ejercicio?

2.3. Procedimiento

El desarrollo de la experiencia de innovación se organizó en tres fases, que se presentan a continuación.

Primera fase: elaboración del genograma y de la línea de la vida. El ejercicio consiste en la elaboración gráfica de la estructura de la familia del alumnado en tres generaciones. A continuación, se trazan líneas de colores que sitúan las relaciones con distintos significados: rojo, relaciones intensas; amarillo, relaciones débiles que te acompañan a lo largo de la vida; azul, relaciones conflictivas y verdes, relaciones vividas que han marcado la vida. Esta fase finaliza escribiendo la línea de la vida con los cinco acontecimientos que se consideran relevantes en su trayectoria vital.

Segunda fase: elaboración de un relato narrativo sobre su historia familiar que incluya los acontecimientos que han identificado en la primera fase del modelo. El objetivo es reconocer los aprendizajes adquiridos a través de vincular patrones relacionales, creencias y roles desempeñados dentro del sistema familiar. Este proceso permitirá aplicar los aprendizajes obtenidos a nuevos contextos en la práctica profesional.

Tercera fase: presentación del trabajo realizado en grupos pequeños, máximo 4 personas, y la supervisión del docente, retroalimentando la reflexión narrativa y la capacidad de adaptación y cambio. El estudiantado de manera voluntaria realiza una exposición narrativa de su biografía señalando en su línea de vida los cinco hitos más relevantes. Este momento de exposición oral no solo está dotado de contenido, sino también de emoción, ya que se incorporan al relato los significados resilientes y fortalezas que el grupo le ha podido devolver, y es la primera vez que cuenta su propia historia desde esa mirada, mucho más ordenada y significativa.

2.4. Procedimiento de análisis de los datos

El procedimiento de análisis de los datos se desarrolló a partir de los dos instrumentos de análisis descritos anteriormente: la guía destinada a la reconstrucción del relato familiar (Tabla 2) y el instrumento orientado a la autoevaluación y reflexión de la experiencia (Tabla 3). Ambos se configuraron como guías de análisis estructuradas mediante preguntas orientadoras, no de obligatoria respuesta, permitiendo explorar y relacionar de manera sistemática los datos objetivos y relacionales representados en el genograma con los significados expresados en los relatos. Estas preguntas facilitaron un proceso progresivo de interpretación, partiendo de la identificación y descripción de los elementos presentes en los datos, para posteriormente explorar las relaciones entre acontecimientos, vínculos, experiencias y significados atribuidos por los discentes. A partir de este proceso comparativo y reflexivo se fueron identificando unidades de significado y códigos iniciales, que posteriormente se agruparon y relacionaron en categorías y temas emergentes. De este modo, las guías constituyeron un recurso para orientar y sistematizar el análisis temático inductivo, manteniendo la construcción de categorías y temas vinculada a los datos y evitando establecer de antemano un sistema categorial cerrado.

3. RESULTADOS

A continuación, se presentan los resultados aportados por los discentes, organizados a través de los temas, códigos y categorías más relevantes de sus relatos, surgidos una vez realizado el análisis inductivo de los genogramas, las historias de vida familiar y autoevaluación.

Tema: La experiencia personal como vía de autoconocimiento y comprensión relacional.
Códigos: Conversación con los padres. Descubrimiento de hechos desconocidos. Reconstrucción de la historia familiar. Cambio en la percepción de la realidad familiar. Comprensión de los acontecimientos familiares. Sorpresa ante una metodología experiencial. Protagonismo de la propia experiencia personal.

Categoría 1: Descubrimiento e interpretación de la historia familiar.

A los participantes les sorprendió que se trabajase con sus experiencias vividas, pero al mismo tiempo sintieron que ser protagonistas del trabajo se convirtió en un reto.

> "Cuando empecé con el genograma, tuve que llamar por teléfono a mi madre y a mi padre para que me diera datos, fechas y me explicaran algunas cosas (…) desde que mis padres se separaron no había conversado así con ellos, me gustó" (ES/12/ 2023-24).

Categoría 2: Impacto formativo del proceso realizado.

Las narrativas muestran que el trabajo con las experiencias propias fue percibido como una experiencia formativa que favoreció el autoconocimiento y la comprensión de las emociones y de las relaciones familiares.

> "Quiero recalcar que este trabajo es duro en sí mismo. Implica revivir recuerdos, enfrentarse a la historia personal y, en muchos casos, volver a conectar con experiencias dolorosas o traumáticas, pero gracias a ello he identificado patrones generacionales que se han repetido en mi genograma y que me han ayudado a tomar conciencia de cómo nuestra historia influye en nuestro y nuestra vocación de acompañar a otros". (ES/76/ 2024-25).

"(…) Algo que me sorprendió gratamente de esta asignatura fue que se trabajase con nuestras propias experiencias vividas. De esta forma, nosotros mismos hemos sido los protagonistas de varias actividades realizadas, y esto ha permitido que nos conozcamos más, sobre todo respecto a las emociones que podemos sentir y como se construyen nuestras relaciones familiares" (ES/17/2022-23).

Tema: Experiencias vitales que transforman la identidad y las relaciones familiares.
Códigos: Aprendizaje a partir de la experiencia. Comprensión de sí mismo a través de reflexionar sobre sí mismo. Impacto de eventos críticos. Acontecimientos vitales no normativos. Asunción de rol de cuidador. Responsabilidad familiar. Influencia del cuidado en la trayectoria de vida.

Categoría 3: Aprendizaje y crecimiento personal

"Pienso que la elaboración del genograma y la posterior historia familiar nos sirve para que veamos cómo afectan las relaciones a la evolución de las personas y esto nos hará mucho más fácil trabajar correctamente con las personas usuarias. Además, no puedo dejar de señalar como la profesora ha sido un pilar fundamental para todo el trabajo realizado" (ES/32/2021-22).

"De otra manera, pienso, que es importante reflexionar sobre cómo nos ha afectado a nosotros las relaciones con los diferentes familiares para entender cómo pueden afectar a otras personas" (ES/3/2023-24).

"Conocer mi historia me ayuda a comprender mejor a las y los demás, a mirar sin juicio y con la convicción de que el dolor, cuando se entiende y se trabaja, puede transformarse en fuerza, conciencia y propósito" (ES/01/2025-26).

Categoría 4: Acontecimientos vitales y rol del cuidado como configuradores de la historia familiar.

Todos los estudiantes que respondieron a las guías afirman que han aprendido y han crecido a través de sus experiencias, aunque resaltan que los cambios más importantes que han marcado su historia han sido acontecimientos vitales no normativos y el rol del cuidado que les adjudicaron.

"Mi madre me contó que no deseaba tener hijos porque estaba cansada de haber cuidado durante toda su infancia y adolescencia a sus hermanos y hermanas. Tenía un rol que no le tocaba, pero luego nos tuvo a nosotras, se separaron y yo fui quien cuidé" (ES/80/2021-22).

"Las relaciones que más me han marcado en mi vida han sido, sin duda, la de mi abuela y la de mi hermana. Mi hermana porque es como una hija para mí y tenemos buena relación." (ES/14/2021-22).

"Mi familia me ha inculcado que debo de responsabilizarme por los cuidados de los demás, siempre he sido hermana y prima mayor, así que me ha tocado cuidar de los pequeños, por ser mayor y por ser mujer." (ES/61/2021-22).

Los discentes construyen un nuevo significado sobre su trayectoria al mirar retrospectivamente su historia familiar, permitiéndoles identificar relaciones, emociones o influencias que anteriormente permanecían implícitas o no habían sido conscientemente relacionadas.

Tema: Reconstrucción de significados a partir de las experiencias familiares, valores, identidad y elección profesional.
Códigos: ausencia previa de conexión entre historia familiar y elección profesional. Reconocimiento de la influencia de las relaciones parentales. Rechazo de normas familiares. Defensa de la justicia so-

cial. Compromiso frente a las injusticias. Reconstrucción retrospectiva del significado de la elección profesional. Normas y relaciones vividas como injustas. Experiencias adversas en la infancia.

Categoría 5: Proyección de la experiencia familiar en la elección profesional

> "Nunca me había planteado que la elección de mi profesión podía llegar a tener una relación con mi historia familiar. Tras elaborar la misma, junto al genograma, he podido comprender que la relación con mis padres ha influido en gran medida a esto. Mi infancia y adolescencia ha estado marcada por unas normas y reglas injustas y patriarcales que no aceptaba. Siempre he tenido el sentimiento de defender lo que es justo y luchar por las injusticias, y creo que de forma inconsciente me he decantado por una profesión desde la que puedo luchar por ello." (ES/37/2025-26).

> "La elección de estudiar y trabajar como profesional del Trabajo Social considero que ha estado influenciada. Por un lado, por vivir en una sociedad en la que querer cambiar ciertas problemáticas y poder contribuir a la defensa de las personas y los derechos humanos. Por otro lado, por la influencia que mi historia familiar ha generado en mí, (…)" (ES/26/2023-24).

Categoría 6: Construcción de una identidad orientada hacia la justicia.

Un número considerable de historias vinculan las experiencias adversas vividas en su infancia con la elección de carrera.

> "Elegir esta profesión está impulsada por muchos motivos. El primero de ellos, cronológicamente hablando, puede estar relacionado con el bullying sufrido en primaria, ya que lo concibo como algo muy injusto y contra lo que hay que luchar, ya que es el principal motivo de suicidio en niños en etapa escolar, por lo que debe erradicarse con mucho trabajo de concienciación". (ES/95/2024-25).

> "Con 8 años, mi madre falleció y todo cambió, viví muchas peleas entre las dos familias y ahora, he llegado a la conclusión de que ejercí de cuidadora de mi padre y yo, me cuidé poco, mi papel de hija se convirtió en madre." (ES/43, 2024–2025)

Se ha detectado que aún no tienen bien integrado el proceso existente entre los datos objetivos y los datos relacionales que les permiten construir significados relacionales. Sin embargo, los discentes que forman parte y/o pertenecen a una familia de padres separados y/o familia reconstituida, son más propensos a elaborar este tipo de significados relacionales y señalan la lealtad como un principio ético que ha marcado sus vidas.

> "Con la separación siempre tenía que estar eligiendo a uno u a otro y no lo entendía, ahora, con el trabajo realizado, creo que me sentía mal porque quería estar bien con los dos y parecía que eso no era posible" (ES/52/2024-25).

> "Asumí el rol que no me correspondía cuando se separaron. No podía soportar que ya no estuvieran juntos. Me convertí en la madre de mi padre cuando estaba con él y era el paño de lágrimas de mi madre, cuando estaba con ella, dejé de hacer lo que me correspondía por edad para volver a unirlos, pero eso no pasó" (ES/15/2025-26).

4. DISCUSIÓN Y CONCLUSIONES

Como primer elemento encontramos una mayor implicación del estudiantado en los ejercicios de entrenamiento realizados en la primera fase ya que les resulta atractivo explorar sobre la vida de su familia, lo que les produce curiosidad e interés, y este recuerdo perdura y lo manifiestan en la eva-

luación del ejercicio. Willingham (2021) sostiene esta afirmación al establecer una relación entre el aprendizaje efectivo y duradero y una participación cognitiva activa por parte del estudiantado.

El segundo elemento a tener en cuenta para elaborar la trayectoria de vida es que este aprendizaje sea reflexivo y se lleve a cabo a través de una práctica intencionada (Ericsson et al. 1993). Los resultados muestran cómo las emociones del alumnado les han facilitado ese aprendizaje profundo al que hace referencia Nieto et al. (2024), a través de metodologías basadas en inteligencia emocional (Redondo-Rodríguez et al., 2023).

El tercer elemento sería utilizar la experiencia propia de cada discente y conseguir así un aprendizaje significativo para toda la vida y que les pueda servir de estrategia de autoformación continua (Domingo, 2021; Beroíza-Valenzuela, 2022).

El cuarto elemento se revela como la relación entre los acontecimientos no normativos y el rol de cuidado adjudicado dentro de su familia para la elección de la carrera, al igual que presentó Fombuena Valero (2011) en su estudio, con la diferencia de que esta elección se presenta de una manera inconsciente. A través de esta investigación, esa elección pasa a un plano consciente y revelador para los discentes, donde descubren las conexiones de los hitos acontecidos en su trayectoria vital y su elección y permanencia en la titulación.

En conclusión, las narrativas de los discentes vienen a demostrar que su participación, desde un plano cognitivo y emocional, les ayuda a reconstruir significados diferentes de su relato de vida a los asumidos por el contexto familiar y a relacionar el conocimiento previo con el que le proporciona el grupo, convirtiéndose en un aprendizaje efectivo y significativo (Bransford, 2000; Ausubel, 2002).

La educación superior debe contribuir al aprendizaje continuo y a lo largo de la vida, alejándonos de la idea de que el aprendizaje se da en una etapa concreta de la vida (Illeris, 2014), por lo que el trabajo desde las experiencias propias convierte a las personas en expertas sobre su propio relato del que siempre se puede profundizar más y del que se puede extraer conocimiento para aplicarlo en la práctica profesional mediante narración, interpretación, resignificación de experiencias y producción de saberes.

La falta de tiempo lectivo y el número elevado de estudiantes no permitió dedicar un mayor espacio individualizado a la realización de los procesos de autoevaluación y coevaluación.

5. REFERENCIAS

Anderson, H., & Gehart, D. (Eds.). (2007). *Collaborative therapy: Relationships and conversations that make a difference*. Routledge.

Ausubel, D. P. (2002). *Adquisición y retención del conocimiento: Una perspectiva cognitiva*. Paidós.

Beroíza-Valenzuela, F. (2022). [Review of the book La revolución reflexiva. Una invitación a crear un futuro de colaboración, by H. Maturana & X. Dávila]. *Revista Saberes Educativos, (9)*, 185–189. https://doi.org/10.5354/2452-5014.2022.67539

Bruner, J. (1991). *Actos de significado: Más allá de la revolución cognitiva*. Alianza.

Castillo-López, M., Romero-Sánchez, E., & Mínguez-Vallejos, R. (2022). El MÉTODO fenomenológico en investigación educativa: Una revisión sistemática. Latinoamericana de Estudios Educativos, 18(2), 241–267. https://doi.org/10.17151/rlee.2023.18.2.11

Colomer, J., Serra, T., Gras, M. E., & Cañabate, D. (2021). Longitudinal self-directed competence development of university students through self-reflection. *Reflective Practice, 22(5)*, 727–740. https://doi.org/10.1080/14623943.2021.1964947

Domingo, À. (2021). La práctica reflexiva: Un modelo transformador de la praxis docente. *Zona Próxima, (34)*, 1–21. https://doi.org/10.14482/zp.34.370.71

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100(3)*, 363–406. https://doi.org/10.1037/0033-295X.100.3.363

Fombuena Valero, J. (2011). Un estudio de las familias de origen de los y las trabajadoras sociales desde el modelo contextual. *Comunitania. Revista Internacional de Trabajo Social y Ciencias Sociales, (2)*, 23–37. https://doi.org/10.5944/comunitania.2.2

Guerra Reyes, F. (2021). La línea de vida: una técnica de recolección de datos cualitativa. *Revista Ecos de la Academia, 5(10)*, 21–29. https://revistasojs.utn.edu.ec/index.php/ecosacademia/article/view/499

Honoré, B. (1980). *Para una teoría de la formación*. Narcea.

Illeris, K. (2014). *Transformative learning and identity*. Routledge.

Lay, K., & McGuire, L. (2010). Building a lens for critical reflection and reflexivity in social work education. *Social Work Education, 29(5)*, 539–550. https://doi.org/10.1080/02615470903362382

Lorenzo Moledo, M. M., Sáez-Gambín, D., Ferraces Otero, M. J., & Varela Portela, C. (2021). Reflection and quality assessment in service-learning projects. When, with whom, and why. *Frontiers in Education, 5*, Article 605099. https://doi.org/10.3389/feduc.2020.605099

Martín, M. (2014). Práctica reflexiva, identidad y trabajo social. *Revista de Trabajo Social Plaza Pública, 7(12)*, 225–244. https://revistaplazapublica.wordpress.com/wp-content/uploads/2015/11/12-12.pdf

McGoldrick, M., & Gerson, R. (1985). *Genogramas en la evaluación familiar*. Gedisa.

Minuchin, S. (1980). *Familias y terapia familiar*. Gedisa.

Nieto, A., Sánchez-Rosas, J., & Gómez-Iñiguez, C. (2024). Identificando el rol de la inteligencia emocional en las emociones de logro y en sus efectos sobre las estrategias de aprendizaje profundo en estudiantes universitarios. *Revista de Psicodidáctica, 29(1)*, 47–56. https://doi.org/10.1016/j.psicod.2023.11.004

Perrenoud, P. (2004). *Diez nuevas competencias para enseñar*. Graó.

Redondo-Rodríguez, C., Becerra-Mejías, J. A., Gil-Fernández, G., & Rodríguez-Velasco, F. J. (2023). Influence of gamification and cooperative work in peer, mixed and interdisciplinary teams on emotional intelligence, learning strategies and life goals that motivate university students to study. *International Journal of Environmental Research and Public Health, 20(1)*, Article 547. https://doi.org/10.3390/ijerph20010547

Schön, D. A. (1998). El profesional reflexivo: Cómo piensan los profesionales cuando actúan. Paidós.

Selvini-Palazzoli, M., Boscolo, L., Cecchin, G., & Prata, G. (1991). *Paradoja y contraparadoja*. Paidós Ibérica.

Sorrentino, A. (2019). El genograma como instrumento gráfico para hipotetizar el funcionamiento mental del paciente. *Vincularte. Revista Clínica y Psicosocial, 4(4)*, 30–48. https://revistavincularte.cl/index.php/ojs/article/view/47

White, M., & Epston, D. (1993). *Medios narrativos para fines terapéuticos*. Paidós.

Willingham, D. T. (2021). *Why don't students like school? A cognitive scientist answers questions about how the mind works and what it means for the classroom* (2nd ed.). Jossey-Bass.

11. Visualizar la comprensión traductora: externalización gráfica del razonamiento en la formación inicial de traductores

Victoria I. Miralles Wieczorek

Universidad de Alicante

RESUMEN

Este trabajo analiza el potencial pedagógico de la externalización visual en la formación inicial de traductores, en la asignatura de Traducción General Directa II (alemán-español). Parte de una dificultad: ante problemas de comprensión, el alumnado tiende a permanecer anclado a la forma lingüística del texto origen y recurrir a estrategias de sustitución literal, lo que genera bloqueos interpretativos y soluciones poco naturales. El objetivo consistió en analizar cómo una microintervención basada en viñetas podía contribuir a hacer observable la construcción del sentido antes de la reformulación verbal, para favorecer una comprensión más reflexiva del texto y del proceso traductor. Asimismo, se exploraron las transformaciones observadas en la participación, la reflexión y la resolución de dificultades interpretativas. La investigación adoptó un estudio de caso cualitativo con diecisiete estudiantes. El análisis se basó en las anotaciones derivadas de la observación docente, las producciones gráficas, la comparación entre versiones iniciales y reformuladas y la reconstrucción de episodios representativos de las discusiones en el aula. Los resultados muestran un desplazamiento desde la literalidad hacia formas de comprensión más orientadas al sentido y menos dependientes de la forma lingüística del texto origen, con mayor regulación entre iguales y una reconfiguración parcial de determinadas dinámicas de aula hacia formas de negociación colectiva del significado. En conclusión, la externalización gráfica favoreció formas más visibles y explícitas de comprensión compartida y de regulación del proceso interpretativo.

PALABRAS CLAVE: externalización visual, competencia traductora, subcompetencia estratégica, aprendizaje visible, innovación docente.

1. INTRODUCCIÓN

En las primeras etapas de la formación universitaria en traducción, es frecuente observar dificultades relacionadas con la comprensión de los textos de partida. Ante segmentos que presentan cierta complejidad, el alumnado tiende a centrarse en la superficie lingüística del texto origen y a resolver la tarea mediante sustituciones literales, incluso cuando estas generan soluciones poco naturales o conceptualmente inconsistentes en la lengua meta. Más que un problema estrictamente lingüístico, esta situación revela una dificultad para construir una representación global del significado. Con el fin de hacer observable este proceso, el presente estudio explora una microintervención basada en la externalización gráfica mediante viñetas, orientada a favorecer una interpretación más reflexiva del texto y del proceso traductor antes de la reformulación verbal. En asignaturas de traducción general situadas en fases iniciales del grado, este fenómeno adquiere especial relevancia, puesto que el estudiantado se encuentra en un momento de transición entre enfoques centrados en el aprendizaje de lenguas y el desarrollo progresivo de la competencia traductora.

El grupo PACTE concibe esta competencia como el sistema subyacente de conocimientos necesarios para traducir, caracterizado por un conocimiento experto predominantemente procedural u operativo (PACTE Group, 2003, 2020). Dentro de este sistema, la subcompetencia estratégica ocupa

una posición central, ya que regula el proceso traductor mediante la planificación de la tarea, la identificación de problemas, la activación de recursos y la evaluación de resultados parciales en función del objetivo final (PACTE Group, 2020). En etapas iniciales de formación, esta dimensión resulta especialmente relevante, pues el alumnado no solo debe aprender a producir traducciones, sino también a supervisar y regular de manera consciente el proceso de toma de decisiones que las sustentan. Esta necesidad puede situarse en el marco del modelo de adquisición de habilidades de Dreyfus y Dreyfus (2005), que describe una progresión desde formas de actuación apoyadas en reglas explícitas hacia modalidades progresivamente más intuitivas y contextualizadas.

En este contexto, uno de los desafíos pedagógicos consiste en hacer accesible un razonamiento que permanece habitualmente oculto en la denominada «caja negra» del traductor (Palumbo, 2009). Su opacidad limita las posibilidades de retroalimentación formativa y dificulta la toma de conciencia sobre los procesos que orientan las decisiones traductoras. Desde el modelo de Aprendizaje Visible, Hattie (2009, 2023) sostiene que el aprendizaje mejora cuando se convierte en una meta identificable y tanto el profesorado como el alumnado pueden identificar qué se está aprendiendo, cómo se está aprendiendo y qué dificultades emergen durante su desarrollo. En este marco, el docente deja de actuar exclusivamente como transmisor o evaluador del producto final y asume un rol de activador del aprendizaje, capaz de valorar el impacto de su enseñanza y ajustar su intervención a partir de evidencias observables.

En paralelo, los enfoques socioculturales del aprendizaje han destacado el papel de la mediación y de la interacción social en la construcción del conocimiento. Desde la perspectiva de Vygotsky (1978), los procesos cognitivos se desarrollan mediante herramientas simbólicas que median la actividad mental y posibilitan la reorganización progresiva de operaciones inicialmente externas hacia formas internalizadas de regulación. En el ámbito de la formación traductora, Kiraly (2013) retoma esta concepción para entender el aprendizaje como un fenómeno emergente, situado y socialmente construido, en el que la competencia se desarrolla mediante la interacción, la reflexión compartida y formas de andamiaje docente orientadas a sostener la participación progresiva del alumnado.

En relación con esta dimensión mediadora, la producción de representaciones visuales por parte del alumnado se ha estudiado en distintos ámbitos educativos como una estrategia de aprendizaje generativo, pues la construcción de una representación pictórica externa a partir de un texto requiere seleccionar e integrar información relevante (Dechamps & Skulmowski, 2025). Desde la perspectiva multimodal, la representación visual no constituye un mero apoyo ilustrativo, sino un modo semiótico con recursos propios que participa activamente en la construcción del significado (Jewitt, 2008; Kress & van Leeuwen, 2006). La transducción entre modos semióticos —en este caso, el paso de una representación verbal a otra visual— supone reconstruir el significado mediante los recursos propios del nuevo modo (Bezemer & Kress, 2008). En la tarea propuesta, este cambio modal exige al aprendiz concretar relaciones que la formulación verbal puede mantener implícitas y permite exteriorizar aspectos de su interpretación. La investigación sobre representaciones externas ha mostrado que la construcción de representaciones por parte del propio alumnado puede favorecer una mejor comprensión de la situación representada (Ainsworth, 2006). El dibujo, por tanto, no se limita a exteriorizar aquello que el estudiante tiene en mente, sino que puede transformar su propia comprensión al favorecer la construcción y reorganización activa del conocimiento (Ainsworth & Scheiter, 2021; Stieff et al., 2026). Como sintetizan las autoras, «Drawing does not simply externalize what learners have in mind; it changes their mind» (Ainsworth & Scheiter, 2021, p. 63).

En este marco, la aportación específica del presente estudio reside en el uso de viñetas producidas por el propio alumnado antes de la reformulación verbal, con el fin de desplazar temporalmente la atención de la forma lingüística del texto origen y favorecer la construcción del sentido. La propuesta se inspira asimismo en la concepción didáctica del cómic desarrollada por Rodríguez Diéguez (1988, pp. 12, 148), quien vincula la producción gráfica del alumnado con la creación y expresión de contenidos propios y plantea la transposición del texto verbal a signos visuales como un procedimiento orientado a objetivar su comprensión. Esta objetivación no implica un acceso directo a los procesos cognitivos del alumnado. Shreve et al. (2011), en el ámbito de la traducción a la vista, plantean que el análisis de la ejecución permite aproximarse a dichos procesos de manera indirecta a partir de sus manifestaciones observables.

En coherencia con estos planteamientos, el presente trabajo se articula en torno a la siguiente pregunta de investigación: ¿en qué medida la externalización visual del sentido contribuye a hacer explícitos aspectos del proceso de comprensión y a reconfigurar determinadas dinámicas de aula en la formación inicial de traductores? Para responder a esta pregunta, se pretende: (1) describir cómo la representación gráfica modifica la interacción en el aula; (2) examinar si favorece la explicitación de los razonamientos interpretativos; y (3) valorar su potencial como estrategia de innovación docente en el ámbito de la enseñanza de la traducción.

2. MÉTODO

La investigación se diseñó como un estudio de caso cualitativo de carácter exploratorio, en coherencia con los marcos metodológicos de los estudios de traducción (Chesterman & Williams, 2002; Saldanha & O'Brien, 2013). Este diseño permite un análisis situado y descriptivo de la intervención pedagógica, que se detalla a continuación mediante la caracterización del contexto, los instrumentos y el procedimiento seguido. De acuerdo con Hansen (2003), el análisis combinó evidencias procedentes de distintas fuentes y las examinó de forma conjunta, con el fin de reforzar la consistencia interpretativa de los hallazgos.

2.1. Descripción del contexto y de los participantes

La experiencia se desarrolló en la asignatura Traducción General Directa II (alemán-español) de segundo curso del Grado en Traducción e Interpretación. El grupo estaba compuesto por 17 estudiantes con asistencia variable a lo largo de las 21 sesiones de dos horas desarrolladas durante un semestre académico. Dado el carácter eminentemente práctico de la asignatura y la importancia de la participación activa en las actividades de traducción y discusión colectiva, la asistencia se registró sistemáticamente en cada sesión como parte del seguimiento docente. La media fue de aproximadamente diez estudiantes por sesión (59 % del grupo) y osciló entre siete y quince asistentes. Como prerrequisito para cursar la asignatura, el alumnado debía acreditar un nivel mínimo B1 de alemán y haber superado la asignatura Traducción General Directa B-A I. No obstante, el grupo contaba todavía con una experiencia limitada en la construcción del sentido en tareas de traducción. El análisis no se orientó al seguimiento individual del alumnado, sino a la identificación de patrones recurrentes de interacción y explicitación del razonamiento observados a lo largo de la intervención.

El desarrollo de la intervención en el contexto ordinario de la asignatura permitió analizar estos procesos en condiciones habituales de enseñanza, lo que favoreció la validez ecológica del estudio

(Saldanha & O'Brien, 2013). La intervención no modificó los objetivos curriculares, los contenidos ni las condiciones naturales del aula.

2.2. Materiales e instrumentos

El corpus de trabajo estuvo integrado por textos de diversa tipología y grado de especialización, entre ellos, recetas, canciones, cómics, textos de literatura infantil, textos turísticos y textos de carácter económico. Los fragmentos analizados en este artículo proceden del álbum de literatura infantil *Der kleine Drache Kokosnuss. Seine ersten Abenteuer* (Siegner, 2002) y del manual de aprendizaje de alemán *Wirtschaftsdeutsch UNI? SICHER!* (Jahr, 2020).

De este corpus, se seleccionaron los *rich points* o segmentos problemáticos, siguiendo la propuesta metodológica del grupo PACTE (2020), entendidos como fragmentos que concentran problemas de traducción prototípicos y requieren procesos de resolución no automatizadoss, que permiten observar manifestaciones relacionadas con la subcompetencia estratégica. Estos segmentos fueron identificados tanto durante la planificación de las actividades como a partir de los bloqueos interpretativos que emergieron en el desarrollo de las sesiones.

Como instrumento de mediación didáctica, se utilizaron las viñetas de externalización visual elaboradas por el alumnado, priorizando la elaboración conceptual y la creación de contenidos propios sobre cualquier criterio de destreza artística (Rodríguez Diéguez, 1988).

2.3. Procedimiento

En coherencia con el enfoque metodológico adoptado, no se persiguió una generalización estadística, sino la comprensión contextualizada de procesos de aprendizaje y dinámicas observables en un contexto formativo situado.

Cuando surgían bloqueos interpretativos o segmentos especialmente problemáticos, se suspendía momentáneamente la dinámica habitual de traducción y se solicitaba al alumnado que representara gráficamente el sentido del fragmento mediante una viñeta esquemática antes de proceder a su reformulación verbal.

La representación se compartía en la pizarra y se discutía colectivamente. La docente orientaba la discusión mediante preguntas abiertas dirigidas a que el alumnado explicitara sus interpretaciones, identificara posibles incoherencias y contrastara distintas hipótesis de comprensión, evitando proporcionar soluciones inmediatas. Las observaciones recogidas se documentaron mediante las anotaciones de la docente-investigadora.

El análisis de los datos se desarrolló mediante una aproximación interpretativa orientada a identificar cambios observables en la explicitación de la comprensión y en la reformulación durante la microintervención. Se articuló en torno a indicadores como la verbalización espontánea de hipótesis interpretativas, el contraste colectivo de alternativas, la reformulación de propuestas y la justificación explícita de decisiones traductoras, que sirvieron como criterios para organizar e interpretar las evidencias recogidas. Para ello, se revisaron las anotaciones de observación de aula elaboradas durante las sesiones, las producciones gráficas realizadas por el alumnado y las diferencias entre versiones iniciales y reformuladas de las traducciones. El análisis se centró en identificar regularidades en la resolución de bloqueos interpretativos, la participación del alumnado, la interacción en torno a la construcción del sentido y la regulación del error durante el proceso de aprendizaje. Con el fin de

reforzar la consistencia interpretativa, se contrastó sistemáticamente la información procedente de las distintas fuentes.

Dada la naturaleza cualitativa del estudio y el doble rol de la investigadora como docente, el análisis se desarrolló desde una postura reflexiva orientada a reforzar la transparencia del proceso analítico (Saldanha & O'Brien, 2013).

3. RESULTADOS

Los resultados que se presentan a continuación ilustran distintos cambios observables durante la microintervención y permiten examinar cómo la externalización visual se relacionó con la explicitación de la comprensión, la reformulación y la interacción en el aula de traducción.

3.1. Externalización visual y resolución de un bloqueo interpretativo en un texto infantil

Uno de los episodios más representativos se produjo durante el trabajo con un fragmento de literatura infantil en el que se explicaba la forma de la Tierra mediante una comparación progresiva: *Nicht so rund wie eine Marzipankartoffel oder eine Rumkugel, sondern eher wie ein Kürbis* (Siegner, 2002, p. 7). La comprensión de esta comparación planteó dificultades derivadas de la interpretación del referente cultural *Marzipankartoffel*.

La primera traducción propuesta por el alumnado —«no tan redonda como una patata de mazapán o una bola de ron, sino más bien como una calabaza»— fue aceptada en una primera fase de trabajo sin objeciones relevantes. No obstante, la incorporación de la fase de externalización visual modificó la secuencia de trabajo. Al solicitar la representación gráfica del fragmento, un estudiante dibujó una «patata» con la forma irregular asociada al tubérculo común. Durante el propio proceso de representación se hizo visible una incongruencia: la comparación perdía coherencia si el primer elemento no respondía a una forma esférica. Ante esta representación, la docente preguntó si la comparación del texto seguía teniendo sentido con esa forma. Esta intervención llevó al alumnado a cuestionar su interpretación inicial y a reconocer que estaba atribuyendo a la «patata de mazapán» la forma de una patata convencional. Con el fin de verificar el referente cultural, el grupo realizó una búsqueda de imágenes de *Marzipankartoffel* y comprobó que se trataba de pequeñas bolas de mazapán recubiertas de cacao, de forma prácticamente esférica. A partir de esta constatación, el propio alumnado comenzó a proponer equivalentes funcionales culturalmente más próximos, como «trufa de chocolate» o «bolitas de anís», que conservaban el rasgo relevante de la comparación.

La representación permitió identificar un problema que no había resultado evidente en la formulación inicial. La dificultad no se situaba únicamente en el plano léxico, sino en la activación de referentes culturales divergentes entre el texto origen y el texto meta. La posterior comprobación del referente desplazó la atención desde la correspondencia léxica hacia el rasgo conceptual relevante para la comparación —la forma esférica— y orientó la búsqueda de soluciones hacia la conservación de ese rasgo en el texto meta.

3.2. Redistribución del control interpretativo

Un segundo episodio significativo se produjo durante el trabajo con un texto de carácter económico en el que se formulaba la idea de que, para alcanzar el «cerebro racional» del cliente, era necesario

convencer previamente al *Türsteher* (Jahr, 2020, p. 52), entendido como una instancia de filtrado emocional. Las primeras propuestas de traducción recurrieron a equivalencias literales como «portero» o «conserje». Aunque formalmente adecuadas, estas opciones resultaban insuficientes para representar con precisión la función conceptual desempeñada por el término dentro de la estructura metafórica del texto.

La introducción de la representación gráfica dio paso a una nueva fase de la discusión interpretativa. Ante las dificultades para comprender el valor metafórico del término, se solicitó a la estudiante responsable del fragmento que lo representara gráficamente en la pizarra. El dibujo mostraba un cerebro dividido en niveles, una puerta simbólica y una figura encargada de regular el acceso a determinadas informaciones. A partir de esta visualización, la docente comenzó a formular preguntas orientadas a comprender la función conceptual del elemento: si se trataba únicamente de alguien que abría una puerta o de una instancia de control que decidía qué información podía pasar al plano racional. La discusión llevó progresivamente al alumnado a abandonar la interpretación estrictamente literal del término y a centrar su atención en la función de vigilancia y filtrado desempeñada por la figura representada.

Como consecuencia de esta discusión, el propio alumnado comenzó a proponer alternativas como «guarda jurado» o incluso «segurata», orientadas menos a reproducir el significado léxico del término original que a preservar la función desempeñada por este dentro del modelo explicativo. De este modo, los criterios de validación de las propuestas se desplazaron progresivamente desde la equivalencia formal hacia la coherencia funcional del esquema interpretativo construido en el aula.

3.3. Transformaciones observadas en el tratamiento del error y en la dinámica de aula

La incorporación de la viñeta también se asoció con modificaciones en la forma en que el error emergía y era tratado durante la actividad de traducción, así como en las dinámicas de participación y construcción colectiva del conocimiento entre el alumnado. En la secuencia habitual de trabajo, las incorrecciones tendían a detectarse principalmente durante la revisión del producto textual final. Con la incorporación de la representación visual, las inconsistencias comenzaron a aparecer en una fase intermedia del proceso, cuando la representación visual hacía más explícitas las hipótesis de comprensión del alumnado.

El carácter provisional y esquemático de las viñetas pareció favorecer un espacio de trabajo menos condicionado por la necesidad de producir respuestas correctas de manera inmediata. Las discrepancias identificadas tendieron a interpretarse como indicios de razonamientos en construcción susceptibles de revisión colectiva, más que como errores concluyentes.

Durante las discusiones de aula se observó una mayor disposición a mostrar dudas, plantear preguntas, sopesar distintas lecturas y revisar soluciones iniciales. Aunque la dinámica habitual de las clases ya promovía la participación mediante preguntas abiertas dirigidas a todo el grupo, la introducción de la viñeta modificó la naturaleza de esa interacción. Mientras que la revisión exclusivamente verbal tendía a concentrar la discusión en el estudiante responsable de la traducción, el soporte visual favorecía que el resto del alumnado interviniera con mayor espontaneidad, planteara hipótesis, señalara posibles incoherencias en las representaciones y propusiera alternativas de manera colaborativa. Las intervenciones del alumnado tendieron a orientarse progresivamente hacia la justificación de interpretaciones y la negociación de alternativas, más que hacia la búsqueda inmediata de equivalencias

formales. En este contexto, el error pasó a constituir un recurso informativo para revisar y ajustar la comprensión compartida del texto durante el propio proceso de traducción.

3.4. Transformaciones observadas en la mediación docente

La modificación de la secuencia didáctica implicó también cambios observables en la mediación docente durante la actividad traductora. En lugar de intervenir principalmente en la validación final del producto textual, la actuación docente pasó a concentrarse con mayor frecuencia en fases intermedias del proceso.

Las preguntas se orientaron a explorar las hipótesis interpretativas que sustentaban las propuestas de traducción y a comprobar su coherencia con la representación gráfica, sin ofrecer una solución traductora inmediata. De este modo, la intervención docente se desplazó de la valoración del producto hacia el examen de los razonamientos que precedían a su formulación definitiva.

En este contexto, la viñeta funcionó como un recurso mediador que proporcionó un soporte compartido para identificar posibles dificultades interpretativas antes de su estabilización en una solución traductora. La retroalimentación pudo integrarse así en el propio desarrollo de la tarea y orientarse hacia el proceso de comprensión y reformulación.

3.5. Síntesis de patrones observados

A lo largo de las sesiones, el análisis permitió identificar cuatro patrones interrelacionados: (1) una explicitación progresiva del razonamiento traductor mediante representaciones visuales compartidas; (2) una resolución más reflexiva de dificultades asociadas a la comprensión del texto; (3) una mayor verbalización y revisión colectiva de hipótesis de comprensión durante el proceso de traducción; y (4) una transformación progresiva del tratamiento del error y de las dinámicas de mediación e interacción en el aula. Considerados de forma conjunta, estos hallazgos sugieren dinámicas compatibles con manifestaciones observables de la subcompetencia estratégica descrita por el grupo PACTE (2020), particularmente en relación con la identificación y resolución de problemas y la toma de decisiones durante el proceso de reformulación. La Tabla 1 sintetiza estos patrones.

Tabla 1. Patrones observados durante la intervención didáctica

Categoría emergente	Manifestación observada	Evidencia empírica
Explicitación del razonamiento traductor	Verbalización de hipótesis interpretativas	Discusión de *Marzipankartoffel* y *Türsteher*
Resolución de dificultades interpretativas	Revisión y reformulación de propuestas tras la representación gráfica	Revisión de incongruencias conceptuales
Transformaciones en la participación y regulación del error	Verbalización, contraste y revisión colectiva de hipótesis	Negociación de alternativas y reformulación compartida
Transformaciones en la mediación docente	Exploración de las hipótesis interpretativas mediante preguntas abiertas	Intervención docente durante fases intermedias del proceso y contraste entre representación gráfica e interpretación verbal

4. DISCUSIÓN Y CONCLUSIONES

Partiendo de una dificultad frecuente en la formación inicial de traductores —la tendencia a resolver problemas de comprensión mediante sustituciones formales sin una reconstrucción conceptual suficiente—, la microintervención diseñada no modificó los contenidos ni los objetivos de la asignatura, sino la secuencia habitual de trabajo, al introducir una fase de externalización visual entre la comprensión inicial y la reformulación verbal.

Los resultados sugieren una reconfiguración parcial de determinadas dinámicas de aula. Al desplazar temporalmente la reformulación lingüística e introducir una fase de explicitación gráfica, el trabajo sobre los segmentos problemáticos dejó de organizarse exclusivamente desde la equivalencia inmediata y pasó a estructurarse como un proceso de análisis compartido del sentido. En este contexto, la externalización visual pareció favorecer una explicitación más temprana de dificultades interpretativas que inicialmente permanecían implícitas durante la reformulación verbal, así como un desplazamiento progresivo desde formas de validación centradas en la equivalencia formal hacia criterios más vinculados a la coherencia conceptual, en línea con enfoques socioconstructivistas de la formación traductora que destacan el papel de la mediación y de la interacción en el aprendizaje (Kiraly, 2013).

En relación con la pregunta de investigación planteada, los indicios observados permiten sostener que la externalización visual contribuyó, en este contexto específico, a hacer observables aspectos de la comprensión antes de la reformulación verbal y a reorientar parcialmente determinadas dinámicas de aula. Los cambios observados se concretaron en una mayor verbalización, contraste y revisión colectiva de las hipótesis interpretativas, acompañadas de un desplazamiento de la mediación docente hacia la exploración de dichas hipótesis en fases intermedias del proceso.

La principal aportación del trabajo reside en mostrar el potencial de una microintervención didáctica para favorecer la explicitación y revisión de los razonamientos interpretativos, así como la construcción social del conocimiento en la formación inicial de traductores. Más que proponer la externalización visual como una técnica aplicable de forma aislada, la experiencia invita a reconsiderar la fase de comprensión en el aula de traducción como un momento susceptible de intervención pedagógica compartida. En el contexto analizado, esta propuesta se concretó en una secuencia didáctica de cuatro pasos: (1) identificación del bloqueo interpretativo en un *rich point*, (2) representación gráfica en la pizarra, (3) discusión colectiva de las distintas hipótesis mediante preguntas abiertas y (4) reformulación verbal orientada a la coherencia funcional. Por su carácter flexible, esta secuencia ofrece una estructura susceptible de adaptación a otros contextos de formación inicial en los que la comprensión de determinados segmentos constituya el principal obstáculo para la reformulación traductora.

No obstante, los hallazgos deben interpretarse como indicios contextualizados derivados de un estudio de caso cualitativo. La ausencia de un grupo de comparación y el carácter interpretativo del análisis limitan tanto la posibilidad de establecer relaciones causales como la transferencia de los hallazgos a otros contextos educativos. Por ello, futuras líneas de trabajo podrían explorar el impacto de este tipo de microintervenciones en grupos más amplios, con otras combinaciones lingüísticas o mediante el seguimiento longitudinal del alumnado.

En definitiva, los resultados sugieren el potencial de la externalización visual como vía de intervención pedagógica sobre la fase de comprensión en la formación inicial de traductores.

REFERENCIAS

Ainsworth, S. E. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, *16*(3), 183-198. https://doi.org/10.1016/j.learninstruc.2006.03.001

Ainsworth, S. E., & Scheiter, K. (2021). Learning by drawing visual representations: Potential, purposes, and practical implications. *Current Directions in Psychological Science*, *30*(1), 61-67. https://doi.org/10.1177/0963721420979582

Bezemer, J., & Kress, G. (2008). Writing in multimodal texts: A social semiotic account of designs for learning. *Written Communication, 25*(2), 166-195. https://doi.org/10.1177/0741088307313177

Chesterman, A., & Williams, J. (2002). *The map: A beginner's guide to doing research in translation studies*. St. Jerome Publishing.

Dechamps, T., & Skulmowski, A. (2025). The effective design of tasks involving learning by drawing: Current trends and methodological progress in research on drawing to learn. *Educational Psychology Review, 37*, 50. https://doi.org/10.1007/s10648-025-10026-2

Dreyfus, H. L., & Dreyfus, S. E. (2005). Peripheral vision: Expertise in real world contexts. *Organization Studies, 26*(5), 779-792. https://doi.org/10.1177/0170840605053102

Hansen, G. (2003). Controlling the process: Theoretical and methodological reflections on research into translation processes. En F. Alves (Ed.), *Triangulating translation: Perspectives in process oriented research* (pp. 25-42). John Benjamins.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.

Hattie, J. (2023). *Visible learning: The sequel: A synthesis of over 2,100 meta-analyses relating to achievement*. Routledge. https://doi.org/10.4324/9781003380542

Jahr, S. (2020). *Wirtschaftsdeutsch UNI? SICHER! Band 4 (B2-C1-C2): Grundlagen zur Volkswirtschaft, Finanzpolitik und zum Unternehmertum*. Booksbaum Verlag.

Jewitt, C. (2008). Multimodality and literacy in school classrooms. *Review of Research in Education*, *32*, 241-267. https://doi.org/10.3102/0091732X07310586

Kiraly, D. (2013). Towards a view of translator competence as an emergent phenomenon: Thinking outside the box(es) in translator education. En D. Kiraly, S. Hansen-Schirra & K. Maksymski (Eds.), *New prospects and perspectives for educating language mediators* (pp. 197-219). Narr.

Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design* (2nd ed.). Routledge.

PACTE Group. (2003). Building a translation competence model. En F. Alves (Ed.), *Triangulating translation: Perspectives in process oriented research* (pp. 43-66). John Benjamins.

PACTE Group, Hurtado Albir, A., Galán-Mañas, A., Kuznik, A., Olalla-Soler, C., Rodríguez-Inés, P., & Romero, L. (2020). Translation competence acquisition. Design and results of the PACTE group's experimental research. *The Interpreter and Translator Trainer, 14*(2), 95-233. https://doi.org/10.1080/1750399X.2020.1732601

Palumbo, G. (2009). *Key terms in translation studies*. Continuum.

Rodríguez Diéguez, J. L. (1988). *El cómic y su utilización didáctica: Los tebeos en la enseñanza*. Gustavo Gili.

Saldanha, G., & O'Brien, S. (2013). *Research methodologies in translation studies*. Routledge.

Shreve, G. M., Lacruz, I., & Angelone, E. (2011). Cognitive effort, syntactic disruption, and visual interference in a sight translation task. En G. Shreve & E. Angelone (Eds.), *Translation and Cognition* (pp. 63-84). John Benjamins.

Siegner, I. (2002). *Der kleine Drache Kokosnuss. Seine ersten Abenteuer*. cbj.
Stieff, M., Scheiter, K., & Ainsworth, S. (2026). Optimising learning from interactive visualisations with drawing prompts. *International Journal of Science Education, 48*(6), 786-815. https://doi.org/10.1080/09500693.2024.2430803
Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

12. La inteligencia artificial generativa en la enseñanza del Derecho: una estrategia para desarrollar las habilidades de pensamiento crítico que sustentan el razonamiento jurídico[1]

Lucía Molina Martínez[1], María Cristina Berenguer Albaladejo[1], Andrea García Jurado[1], Amanda Judith Gil Marín[1], María Remedios Guilabert Vidal[2], Lucía López Gutiérrez[1], Isabela Erika Ungureanu Orosz[1]

[1]*Universidad de Alicante*
[2]*Universidad Miguel Hernández de Elche*

RESUMEN

La incorporación de la inteligencia artificial generativa en la educación superior plantea el desafío de aprovechar su potencial educativo sin sustituir los procesos cognitivos implicados en el pensamiento crítico. En este contexto, el presente estudio analiza una estrategia de integración crítica de sistemas de inteligencia artificial generativa en la enseñanza del Derecho basada en el análisis, el contraste y la fundamentación jurídica de las respuestas generadas por estas herramientas. La experiencia se desarrolló durante los cursos académicos 2024/2025 y 2025/2026 en seis grupos docentes de la Universidad de Alicante y de la Universidad Miguel Hernández de Elche, con una muestra global de 316 estudiantes. La investigación siguió un diseño metodológico mixto, que combinó grupos de discusión, una rúbrica de evaluación de evidencias de aprendizaje y un cuestionario dirigido al alumnado. Los resultados muestran una evolución favorable en habilidades relacionadas con la evaluación crítica de la información, el contraste de fuentes normativas, doctrinales y jurisprudenciales, la fundamentación jurídica de las respuestas y la coherencia argumentativa. Sin embargo, estas mejoras no se traducen en una percepción equivalente del desarrollo del pensamiento crítico considerado como competencia global. Los hallazgos apuntan a que el valor educativo de la inteligencia artificial generativa depende de las condiciones pedagógicas de su utilización y, en particular, de metodologías que sitúen sus respuestas como objeto de análisis, verificación y fundamentación, manteniendo al estudiante como sujeto activo del razonamiento jurídico y limitando el riesgo de delegación intelectual.

PALABRAS CLAVE: inteligencia artificial generativa, razonamiento jurídico, pensamiento crítico, enseñanza del Derecho, innovación docente.

1. INTRODUCCIÓN

La irrupción de la inteligencia artificial generativa constituye uno de los fenómenos más transformadores que ha experimentado la educación superior en los últimos años. Su utilización por parte del alumnado se ha extendido de manera notable, convirtiéndose en un recurso habitual para la búsqueda de información, la elaboración de textos y la resolución de actividades académicas. Más que una innovación tecnológica puntual, constituye una transformación estructural que está modificando las formas de acceso al conocimiento y la relación del alumnado con la información. En este

1 El presente trabajo ha contado con una ayuda del Programa de Redes de investigación en docencia universitaria del Instituto de Ciencias de la Educación de la Universidad de Alicante (convocatoria 2024). Ref.: 6268.

contexto, las instituciones universitarias deben adaptar sus metodologías docentes a un entorno en el que los sistemas de inteligencia artificial forman ya parte del ecosistema habitual de aprendizaje (UNESCO, 2023).

Esta transformación obliga a reconsiderar algunas de las premisas sobre las que tradicionalmente se ha sustentado la enseñanza universitaria. Si durante décadas una parte importante de la actividad académica estuvo vinculada a procesos de búsqueda, selección y elaboración de información, la capacidad de los sistemas de inteligencia artificial para generar respuestas complejas de forma inmediata modifica la naturaleza de estas tareas y desplaza parte de la actividad del estudiante hacia la valoración de la información obtenida. En este nuevo escenario, la utilidad de una respuesta generada por inteligencia artificial no depende únicamente de la información que proporciona, sino de la capacidad del estudiante para analizarla, interpretarla, contrastarla y fundamentarla antes de incorporarla a su proceso de aprendizaje. Estas operaciones forman parte de las habilidades cognitivas asociadas al pensamiento crítico, entendido como un constructo multidimensional que implica procesos de análisis, interpretación, evaluación y reflexión orientados a la formulación de juicios fundamentados (Facione, 2023).

Precisamente porque el pensamiento crítico ocupa una posición central en los procesos de aprendizaje universitario, la relación entre inteligencia artificial generativa y el desarrollo de las habilidades que lo integran se ha convertido en uno de los principales debates en el ámbito de la educación superior. La literatura científica ha señalado que la inteligencia artificial generativa puede favorecer procesos de reflexión, aprendizaje activo y construcción del conocimiento cuando se integra en metodologías que exigen la participación activa del alumnado, el contraste de la información y la fundamentación de las respuestas (Atencio-González et al., 2023; Melisa et al., 2025; Mollick y Mollick, 2023). Este potencial educativo no deriva, sin embargo, de la tecnología por sí misma, sino de las condiciones didácticas que estructuran su utilización (Torrico Cuevas, 2026). En la enseñanza del Derecho, ello exige concebir la inteligencia artificial como un recurso de apoyo al razonamiento del estudiante, no como un sustituto de este, favoreciendo el desarrollo de una mentalidad jurídica crítica basada en la interpretación, el análisis y la argumentación (Alemán Madrigal, 2024; Carretero Sánchez, 2025).

Estas potencialidades conviven, sin embargo, con importantes riesgos cuando el estudiante delega en la inteligencia artificial procesos de razonamiento que debería desarrollar por sí mismo. Una utilización acrítica o excesivamente dependiente puede favorecer la aceptación automática de la información, reducir la implicación cognitiva del alumnado y limitar el desarrollo de capacidades analíticas y reflexivas (AlAfnan et al., 2023; Melisa et al., 2025). En esta línea, un reciente estudio experimental desarrollado por Kosmyna et al. (2025) ha puesto de manifiesto que el uso de modelos de lenguaje puede asociarse a una menor implicación cognitiva cuando sustituyen procesos de razonamiento que deberían ser desarrollados por el propio estudiante. A estos riesgos se añaden las posibles deficiencias de fiabilidad de las respuestas, la reproducción de sesgos, la utilización de referencias inexactas y las implicaciones de su uso para la integridad académica, circunstancias que evidencian la necesidad de establecer orientaciones pedagógicas claras y promover un uso crítico y ético de estas herramientas (Berzunza Criollo, 2025; Galli y Kanobel, 2023).

Esta necesidad de orientar pedagógicamente el uso de la inteligencia artificial plantea un reto específico en la enseñanza del Derecho, donde las habilidades de análisis, interpretación, evaluación y fundamentación propias del pensamiento crítico se proyectan sobre el razonamiento jurídico y permiten al estudiante interpretar las normas, contrastar las distintas fuentes jurídicas y fundamentar

las soluciones adoptadas (Alemán Madrigal, 2024). La cuestión pasa, así, por determinar bajo qué planteamientos didácticos la incorporación de estas herramientas puede integrarse en la formación jurídica de manera compatible con las exigencias propias de esta disciplina.

Sin embargo, pese al creciente interés que esta cuestión ha suscitado, continúan siendo limitadas las evidencias empíricas sobre las estrategias pedagógicas que permiten integrar la inteligencia artificial sin desplazar la actividad cognitiva del estudiante y, en particular, sobre su aplicación a la enseñanza del Derecho. La mayor parte de las investigaciones se ha centrado en analizar las percepciones del profesorado y del alumnado o en describir las ventajas y riesgos potenciales de estas herramientas, mientras que siguen siendo escasos los estudios que evalúan, a partir de experiencias desarrolladas en el aula, qué metodologías favorecen una integración crítica y cuáles son sus efectos sobre competencias esenciales para la formación jurídica. Persiste, por tanto, un vacío de conocimiento acerca de las condiciones pedagógicas que permiten incorporar estas tecnologías sin comprometer las competencias que constituyen la base del razonamiento jurídico.

Con el propósito de contribuir a este ámbito de investigación, el presente estudio examina una estrategia de integración crítica de sistemas de inteligencia artificial generativa en la enseñanza del Derecho. En concreto, analiza si una metodología basada en el contraste sistemático de las respuestas generadas por la inteligencia artificial con fuentes normativas, doctrinales y jurisprudenciales contribuye al desarrollo de habilidades del pensamiento crítico implicadas en el razonamiento jurídico, al tiempo que permite abordar el riesgo de delegación intelectual asociado a una utilización acrítica de estas herramientas.

A partir de estos planteamientos, la investigación persigue tres objetivos: (1) analizar la evolución del desempeño del alumnado, antes y después de la intervención, en determinadas habilidades vinculadas al pensamiento crítico y aplicadas al razonamiento jurídico; (2) analizar la percepción del alumnado sobre la utilidad de la metodología y sobre su contribución al desarrollo de tales habilidades; y (3) identificar, a partir de la experiencia desarrollada, sus principales fortalezas, limitaciones y posibilidades de mejora para la integración crítica de la inteligencia artificial generativa en la enseñanza del Derecho.

2. MÉTODO

2.1. Descripción del contexto y de los participantes

La investigación se ha desarrollado en asignaturas adscritas al área de Derecho civil impartidas en la Universidad de Alicante y en la Universidad Miguel Hernández de Elche, pertenecientes a los grados en Derecho, Derecho y Relaciones Internacionales (DERRI), Derecho y Administración y Dirección de Empresas (DADE) y Turismo. La intervención se ha implementado en seis grupos docentes correspondientes a asignaturas de carácter básico, obligatorio y optativo, impartidas en distintos cursos académicos y dirigidas a estudiantes con diferentes niveles de formación jurídica.

La muestra ha estado integrada por 316 estudiantes, con una distribución relativamente equilibrada por sexo (61 % mujeres y 39 % hombres). Asimismo, el 6 % del alumnado participante presentaba necesidades específicas de apoyo educativo (NEAE), circunstancia que introduce diversidad en el contexto de aplicación de la experiencia.

La heterogeneidad de la muestra, tanto en términos de titulación como de nivel formativo y perfil del estudiantado, resulta especialmente relevante para los objetivos de la investigación, al permitir

examinar la aplicación de la experiencia en contextos académicos caracterizados por diferentes grados de complejidad en las tareas de interpretación, argumentación y razonamiento jurídico.

2.2. Instrumentos

La investigación se articuló a partir de un diseño metodológico mixto de carácter secuencial, combinando técnicas cualitativas y cuantitativas de recogida y análisis de datos. Esta estrategia permitió abordar el fenómeno objeto de estudio desde perspectivas complementarias y reforzar la consistencia de los resultados mediante la triangulación de técnicas e instrumentos.

Como instrumento de recogida de información cualitativa en la fase *pretest*, se empleó el *focus group* para explorar las percepciones del alumnado acerca del uso de sistemas de inteligencia artificial generativa en el contexto universitario. Se realizaron seis grupos de discusión, uno por cada grupo docente participante, integrados por cinco estudiantes seleccionados mediante participación voluntaria y una persona moderadora. La discusión en cada grupo fue conducida mediante un guion semiestructurado compuesto por preguntas abiertas relativas al conocimiento y uso de estas herramientas, su utilidad académica, su influencia en los procesos de aprendizaje, la fiabilidad de las respuestas generadas, las implicaciones éticas asociadas a su utilización y las expectativas sobre su integración en la educación superior.

Para la evaluación de las evidencias de aprendizaje generadas antes y después de la implementación de la metodología (*pretest-postest*), se diseñó una rúbrica *ad hoc* integrada por indicadores representativos de las habilidades del pensamiento crítico que sustentan el razonamiento jurídico. Las dimensiones evaluadas comprendieron la identificación de errores, limitaciones y sesgos en la información generada por la inteligencia artificial, el contraste con fuentes normativas, doctrinales y jurisprudenciales, la fundamentación jurídica de las respuestas, la coherencia argumentativa y la adecuación de las soluciones propuestas a los supuestos planteados. La rúbrica permitió comparar el desempeño observado antes y después de la intervención.

Como instrumento de recogida de información cuantitativa en la fase *postest,* se administró un cuestionario *ad hoc* dirigido al alumnado para conocer su valoración sobre la metodología de integración crítica de la inteligencia artificial generativa implementada. El cuestionario fue cumplimentado por 105 estudiantes de los 316 que integraban la muestra participante (n = 316) y permitió evaluar la percepción del alumnado acerca de la contribución de esta metodología al aprendizaje, al desarrollo de habilidades vinculadas al pensamiento crítico y al razonamiento jurídico, así como el grado de satisfacción con la experiencia desarrollada. El cuestionario estuvo compuesto por doce ítems medidos mediante una escala Likert de cinco puntos (1 = totalmente en desacuerdo; 5 = totalmente de acuerdo) e incorporó un apartado final con preguntas abiertas destinadas a recoger observaciones, fortalezas, limitaciones y propuestas de mejora de la experiencia desarrollada.

La participación del alumnado en las distintas técnicas de recogida de información fue voluntaria y anónima, garantizándose en todo momento la confidencialidad de la información y el tratamiento de los datos exclusivamente con fines de investigación y mejora docente. Los datos cualitativos fueron analizados con el software QDA Miner, mientras que los datos de naturaleza cuantitativa se procesaron mediante el programa estadístico SPSS (versión 22).

2.3. Procedimiento

La investigación se desarrolló durante los cursos académicos 2024/2025 y 2025/2026 mediante la implementación de una estrategia de integración crítica de sistemas de inteligencia artificial generativa en la enseñanza del Derecho. La intervención se diseñó para examinar la evolución de determinadas habilidades de pensamiento crítico implicadas en el razonamiento jurídico a partir del contraste de las respuestas generadas por estas herramientas.

La experiencia se estructuró en tres fases sucesivas. En una primera fase, previa a la intervención, se exploraron las percepciones iniciales del alumnado acerca del uso de la inteligencia artificial generativa en el ámbito universitario, con el fin de conocer los patrones de utilización de estas herramientas y las expectativas desde las que los estudiantes afrontaban la experiencia.

La segunda fase correspondió al desarrollo de la intervención docente. Para ello, el profesorado diseñó actividades prácticas integradas en las distintas unidades didácticas de las asignaturas participantes. Los supuestos planteados se caracterizaban por la existencia de controversias doctrinales, divergencias jurisprudenciales o modificaciones normativas recientes, con el propósito de situar al alumnado ante problemas que exigieran interpretar normas, contrastar fuentes y fundamentar jurídicamente las soluciones adoptadas.

A lo largo del semestre, el alumnado utilizó sistemas de inteligencia artificial generativa para obtener una primera propuesta de resolución de cada caso práctico. Dicha respuesta actuó como punto de partida del proceso de aprendizaje. Sobre ella, los estudiantes desarrollaron un proceso de revisión y validación jurídica orientado a detectar errores, limitaciones o sesgos, verificar la corrección de la información proporcionada, contrastarla con fuentes normativas, doctrinales y jurisprudenciales y elaborar una respuesta propia suficientemente fundamentada. De este modo, la inteligencia artificial pasó de ser un mecanismo de obtención de respuestas a convertirse en un recurso sometido al análisis crítico del alumnado.

Las actividades se realizaron de forma individual o colaborativa, en función de las características de cada asignatura, y culminaron en sesiones de discusión dirigidas por el profesorado, orientadas a confrontar distintas interpretaciones jurídicas y a reforzar los procesos de argumentación, deliberación y fundamentación desarrollados durante la resolución de los casos.

Finalmente, una vez concluida la intervención, se procedió al análisis comparativo de las evidencias de aprendizaje generadas antes y después de la aplicación de la metodología y a la recogida de la valoración del alumnado sobre la experiencia desarrollada. La aplicación continuada de esta estrategia durante dos cursos académicos consecutivos permitió examinar su funcionamiento en distintos contextos de formación jurídica.

3. RESULTADOS

Con el fin de dar respuesta a los objetivos de investigación planteados, los resultados se estructuran en tres apartados. En primer lugar, se examinan los cambios observados en el desempeño del alumnado tras la intervención; en segundo término, se aborda su valoración de la experiencia y de la utilidad de la metodología; finalmente, se recogen las principales fortalezas, limitaciones y posibilidades de mejora identificadas a partir de su aplicación.

3.1. Evolución de las habilidades del pensamiento crítico implicadas en el razonamiento jurídico

El primer objetivo de la investigación consistió en analizar la incidencia de la metodología implementada sobre las habilidades del pensamiento crítico que, en el ámbito jurídico, sustentan el razonamiento jurídico. Para ello, se compararon las evidencias de aprendizaje generadas por el alumnado antes y después de la intervención mediante la rúbrica elaborada *ad hoc*. Los resultados muestran una evolución favorable en todas las dimensiones evaluadas (Tabla 1)

Tabla 1. Comparación *pretest-postest* de las dimensiones evaluadas mediante la rúbrica (%)

Dimensión evaluada	Pretest (%)	Postest (%)
Contraste con fuentes normativas, doctrinales y jurisprudenciales	37	82
Fundamentación jurídica de las respuestas	46	78
Coherencia argumentativa	43	74
Adecuación de las soluciones propuestas	39	71
Identificación de errores, limitaciones y sesgos en respuestas generadas por IA	28	68

Nota: Los porcentajes reflejan la proporción de actividades que alcanzaron niveles satisfactorios de desempeño en cada dimensión evaluada.

En conjunto, los resultados ponen de manifiesto un incremento generalizado en todas las dimensiones evaluadas, cuya frecuencia de aparición prácticamente se duplica tras la implementación de la metodología. Los incrementos más acusados se observan en el contraste con fuentes normativas, doctrinales y jurisprudenciales, que pasa del 37 % al 82 %, y en la fundamentación jurídica de las respuestas, que aumenta del 46 % al 78 %. Ambas dimensiones guardan una relación directa con la evaluación crítica de la información y con la construcción de argumentos jurídicamente fundamentados. La coherencia argumentativa y la adecuación de las soluciones propuestas experimentan, por su parte, incrementos superiores a treinta puntos porcentuales.

La evolución observada en la identificación de errores, limitaciones y sesgos en las respuestas generadas por la inteligencia artificial merece una consideración específica. Aunque esta dimensión registra el mayor incremento relativo, al pasar del 28 % al 68 %, continúa siendo la que presenta el menor nivel de desempeño al finalizar la intervención. Este resultado sugiere que la supervisión crítica de las respuestas constituye una de las habilidades más complejas de desarrollar, incluso en el marco de una metodología diseñada específicamente para favorecerla.

Los resultados obtenidos mediante la rúbrica son parcialmente congruentes con las percepciones sobre las competencias adquiridas expresadas por el alumnado en el cuestionario final (Tabla 2). Las puntuaciones más elevadas corresponden a la mejora de la fundamentación y justificación de las respuestas (M = 4.00) y a la capacidad para evaluar críticamente la información generada por la inteligencia artificial (M = 3.93), precisamente las habilidades que registran una evolución más favorable en las evidencias de aprendizaje. Asimismo, el alumnado percibe una mejora en la construcción de un conocimiento más completo y fundamentado (M = 3.54) y en la comprensión de las posibilidades y limitaciones de estas tecnologías en el aprendizaje jurídico (M = 3.54).

Tabla 2. Percepción del alumnado sobre competencias adquiridas tras la intervención (%)

Ítems	Totalmente en desacuerdo (%)	En desacuerdo (%)	Ni de acuerdo ni en desacuerdo (%)	De acuerdo (%)	Totalmente de acuerdo (%)	Media	Desviación típica
1. Mejora del pensamiento crítico	6.70	20.00	26.70	46.70	0.00	3.14	0.96
2. Evaluación crítica de la información generada por la IA	0.00	6.70	13.30	60.00	20.00	3.93	0.78
3. Conocimiento más completo y fundamentado	6.70	6.70	26.70	46.70	13.30	3.54	1.03
4. Comprensión de los límites y oportunidades de la IA	6.70	6.70	26.70	46.70	13.30	3.54	1.03
5. Fundamentación y justificación de las respuestas	0.00	0.00	20.00	60.00	20.00	4.00	0.63
6. Identificación de errores y sesgos en las respuestas de la IA	0.00	13.30	26.70	33.30	26.70	3.73	0.95

En contraste con la elevada valoración obtenida por habilidades concretas como la evaluación crítica de la información o la fundamentación jurídica, la percepción de mejora del pensamiento crítico considerado como competencia global resulta sensiblemente más moderada (M = 3.14). Este resultado sugiere que el alumnado distingue entre el desarrollo de habilidades específicas vinculadas al análisis, la evaluación, el contraste y la argumentación jurídica y el desarrollo del pensamiento crítico entendido como una competencia transversal de mayor complejidad.

En conjunto, la rúbrica y el cuestionario convergen en señalar avances percibidos y observados en habilidades relacionadas con la evaluación crítica de la información, el contraste de fuentes jurídicas y la fundamentación argumentativa. La diferencia entre las puntuaciones obtenidas en habilidades específicas y la percepción global del pensamiento crítico aconseja mantener ambos planos analíticos diferenciados.

3.2. Percepción del alumnado sobre la inteligencia artificial generativa y la utilidad de la metodología

Los grupos de discusión realizados antes de la intervención mostraron un uso habitual de herramientas de inteligencia artificial generativa con fines académicos, principalmente para la búsqueda de información, la resolución de dudas y la elaboración de tareas. Al mismo tiempo, los participantes manifestaron reservas sobre la fiabilidad de las respuestas y reconocieron la necesidad de contrastar críticamente la información obtenida. Estas aportaciones tuvieron carácter exploratorio y sirvieron para contextualizar la intervención.

Tras la implementación de la metodología, los resultados del cuestionario reflejan una valoración globalmente positiva de la experiencia desarrollada (Tabla 3). Las puntuaciones medias más elevadas corresponden a la utilidad para el aprendizaje (M = 3.80) y a la comprensión de los contenidos de la asignatura (M = 3.80). También se valoran positivamente la posibilidad de extender esta metodología a otras asignaturas del grado (M = 3.73), su contribución a un uso responsable y ético de la inteligencia artificial (M = 3.73) y el fomento del trabajo activo y colaborativo (M = 3.60).

Tabla 3. Percepción del alumnado sobre la utilidad de la metodología (%)

Ítems	Totalmente en desacuerdo (%)	En desacuerdo (%)	Ni de acuerdo ni en desacuerdo (%)	De acuerdo (%)	Totalmente de acuerdo (%)	Media	Desviación típica
1. Utilidad para el aprendizaje	0.00	13.30	26.70	26.70	33.30	3.80	0.99
2. Comprensión de los contenidos de la asignatura	6.70	6.70	20.00	33.30	33.30	3.80	1.05
3. Fomento del trabajo activo y colaborativo	0.00	33.30	13.30	13.30	40.00	3.60	1.18
4. Uso responsable y ético de la IA	0.00	20.00	20.00	26.70	33.30	3.73	1.07
5. Accesibilidad y adaptación a NEAE	0.00	6.70	40.00	53.30	0.00	3.47	0.68
6. Utilidad para otras asignaturas	13.30	0.00	20.00	33.30	33.30	3.73	1.19

La accesibilidad de la metodología y su adaptación a las necesidades específicas del alumnado obtiene una valoración más moderada (M = 3.47), lo que pone de manifiesto la conveniencia de continuar avanzando en el diseño de estrategias inclusivas que permitan responder a la diversidad presente en las aulas universitarias.

3.3. Fortalezas, limitaciones y posibilidades de mejora de la metodología

La experiencia permitió identificar diversas fortalezas, limitaciones y oportunidades de mejora de la metodología implementada. Entre las primeras destaca la toma de conciencia del alumnado sobre la necesidad de contrastar las respuestas generadas por la inteligencia artificial antes de incorporarlas a la resolución de problemas jurídicos, junto con una mayor atención a la fundamentación del razonamiento jurídico y al uso responsable de estas herramientas. Entre las limitaciones destacan la participación desigual del alumnado, la posibilidad de un sesgo de autoselección en los grupos de discusión, la necesidad de un acompañamiento docente más intenso durante las primeras fases de implementación y la mayor dificultad que plantea la identificación de errores, limitaciones y sesgos

cuando las respuestas generadas por la inteligencia artificial presentan una apariencia de corrección jurídica. Estas limitaciones se ven, además, acentuadas por la utilización, por parte del alumnado, de distintos sistemas y versiones de inteligencia artificial generativa, cuyas diferencias en capacidad de razonamiento, precisión y calidad de las respuestas hacen que los estudiantes no se enfrenten a un mismo nivel de complejidad en las tareas de evaluación crítica.

A partir de estas observaciones, se identifican como principales líneas de mejora la incorporación de una formación inicial para el alumnado sobre el uso crítico de la inteligencia artificial, así como el diseño de una progresión metodológica basada en la utilización planificada de sistemas y versiones con niveles crecientes de complejidad, comenzando por aquellos cuyas respuestas presenten errores y limitaciones fácilmente identificables y avanzando progresivamente hacia sistemas que generen respuestas cada vez más difíciles de someter a una evaluación crítica.

4. DISCUSIÓN Y CONCLUSIONES

La principal aportación del presente estudio no reside en afirmar que la inteligencia artificial generativa mejore el pensamiento crítico, sino en mostrar una forma de integrarla en la enseñanza del Derecho que preserve la participación cognitiva del estudiante en la elaboración de la respuesta jurídica. Los resultados obtenidos muestran una evolución favorable en determinadas habilidades vinculadas al pensamiento crítico y aplicadas al razonamiento jurídico, en particular la evaluación crítica de la información, el contraste con fuentes jurídicas y la fundamentación de las soluciones adoptadas. Esta evolución no se corresponde, sin embargo, con una percepción equivalente de mejora del pensamiento crítico entendido como competencia global, lo que refuerza la necesidad de distinguir entre el desarrollo de habilidades específicas, observables en tareas concretas de razonamiento jurídico, y la adquisición de un constructo competencial más amplio. Esta interpretación resulta compatible con la concepción de Facione (2023), que entiende el pensamiento crítico como un constructo multidimensional integrado por diversas habilidades cognitivas cuya aplicación permite valorar información, formular juicios fundamentados y abordar problemas complejos.

Los resultados permiten, además, precisar algunas de las aportaciones de la literatura previa. En línea con Mollick y Mollick (2023) y Torrico Cuevas (2026), los hallazgos sugieren que el potencial educativo de la inteligencia artificial depende en buena medida de las condiciones en las que se integra en la actividad docente y del grado de implicación que se exige al alumnado. En la formación jurídica, dicha implicación debe traducirse en procesos de verificación del marco jurídico aplicable y construcción argumentativa de las respuestas. Esta interpretación resulta igualmente coherente con Alemán Madrigal (2024), para quien la incorporación de la inteligencia artificial a la enseñanza del Derecho adquiere valor formativo cuando complementa, y no sustituye, el razonamiento jurídico del estudiante. Del mismo modo, los resultados pueden interpretarse a la luz de las conclusiones de Kosmyna et al. (2025), que advierten de una posible reducción de la implicación cognitiva cuando los modelos de lenguaje sustituyen parte del razonamiento que corresponde desarrollar al estudiante. Frente a este riesgo, la metodología analizada sitúa las respuestas generadas por la inteligencia artificial como punto de partida de una actividad que exige su análisis, verificación y fundamentación jurídica.

Ahora bien, los resultados también ponen de manifiesto que las distintas habilidades evaluadas no evolucionan de manera homogénea. La identificación de errores, limitaciones y sesgos en las respuestas generadas por la inteligencia artificial continúa siendo la dimensión en la que el alumnado alcanza

un menor nivel de desempeño al finalizar la intervención, lo que refuerza las preocupaciones planteadas por Berzunza Criollo (2025) y Galli y Kanobel (2023) acerca de los riesgos de una aceptación acrítica de estas respuestas. Esta dificultad puede adquirir mayor relevancia a medida que los sistemas de inteligencia artificial generan contenidos más elaborados, coherentes y con una mayor apariencia de corrección jurídica, pues aumenta también la exigencia cognitiva necesaria para someterlos a un análisis crítico. Desde una perspectiva metodológica, ello sugiere la conveniencia de introducir progresivamente herramientas con distintos niveles de sofisticación, de modo que el alumnado pueda consolidar inicialmente estrategias de identificación y contraste ante errores y limitaciones más manifiestos antes de enfrentarse a modelos más avanzados.

La interpretación de los resultados debe hacerse, no obstante, dentro de los límites de la experiencia. Su aplicación en contextos docentes concretos y con una muestra heterogénea aconseja prudencia al trasladar los hallazgos a otros escenarios. La participación voluntaria en los grupos de discusión puede haber introducido un sesgo de autoselección, especialmente al no haberse establecido criterios específicos destinados a garantizar una determinada diversidad de perfiles. También debe tenerse en cuenta que la rúbrica y el cuestionario fueron elaborados específicamente para esta experiencia y que, aunque la primera permite apreciar la evolución del desempeño en las habilidades trabajadas, ninguno de los instrumentos constituye por sí solo una medida del pensamiento crítico entendido globalmente. Por otra parte, el diseño de la investigación no permite determinar si las mejoras observadas se mantienen a largo plazo ni si son plenamente transferibles a otros contextos de resolución de problemas jurídicos. Futuras investigaciones deberían incorporar diseños longitudinales o cuasiexperimentales, criterios de selección más estructurados e instrumentos sometidos a procedimientos de validación que permitan examinar con mayor precisión tanto la evolución de habilidades específicas como su posible relación con el desarrollo más amplio del pensamiento crítico.

En conclusión, la incorporación de la inteligencia artificial generativa a la enseñanza del Derecho no debe valorarse únicamente por las posibilidades que abre, sino también por las capacidades intelectuales cuya conservación puede exigir su propia presencia. Los resultados obtenidos apuntan a que una integración pedagógicamente orientada, en línea con las recomendaciones de la UNESCO (2023), puede hacer compatible el uso de estas herramientas con el mantenimiento de la exigencia intelectual propia de la formación jurídica. Desde esta perspectiva, preservar no significa mantener al estudiante al margen de la innovación, sino impedir que la automatización de la respuesta termine haciendo prescindible el juicio que debe preceder a su aceptación. Es precisamente en ese espacio, entre la respuesta que la tecnología ofrece y la razón por la que el estudiante decide asumirla como válida, donde se sitúa el valor formativo de su incorporación a la enseñanza del Derecho.

5. REFERENCIAS

AlAfnan, M. A., Dishari, S., Jovic, M., & Lomidze, K. (2023). ChatGPT as an educational tool: Opportunities, challenges, and recommendations for communication, business writing, and composition courses. *Journal of Artificial Intelligence and Technology*, *3*(2), 60–68. https://doi.org/10.37965/jait.2023.0184

Alemán Madrigal, L. E. (2024). Innovación educativa: ChatGPT y el Derecho de la Seguridad Social en Nicaragua. *Revista Científica de Estudios Sociales, 4*(4), 68–85. https://doi.org/10.62407/rces.v4i4.116

Atencio-González, R. E., Bonilla-Ron, D. E., Miles-Flores, M. V., & López-Zavala, S. A. (2023). Chat GPT como recurso para el aprendizaje del pensamiento crítico en estudiantes universitarios. *Cienciamatria. Revista Interdisciplinaria de Humanidades, Educación, Ciencia y Tecnología, 9*(17), 36-44. https://doi.org/10.35381/cm.v9i17.1121

Berzunza Criollo, M. C. (2025). Percepciones y usos de ChatGPT en la construcción del pensamiento crítico de estudiantes universitarios. *Ciencia y Reflexión, 4*(3), 201–235. https://doi.org/10.70747/cr.v4i3.413

Carretero Sánchez, S. (2025). Las prácticas en teoría del derecho desde la inteligencia artificial ChatGPT: Una propuesta didáctica razonable. *Derecho Digital e Innovación*, (23), 1–18.

Facione, P. A. (2023). *Critical thinking: What it is and why it counts* (2023 update). Insight Assessment.

Galli, M. G., & Kanobel, M. C. (2023). ChatGPT en Educación Superior: Explorando sus potencialidades y sus limitaciones. *Revista de Educación Superior y Sociedad (ESS), 35*(2), 174–195. https://doi.org/10.54674/ess.v35i2.815

Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.-H., Beresnitzky, A. V., Braunstein, I., & Maes, P. (2025). Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing tasks. *arXiv.* https://doi.org/10.48550/arXiv.2506.08872

Melisa, R., Ashadi, A., Triastuti, A., Hidayati, S., Salido, A., Ero, P. E. L., Marlini, C., Zefrin, Z., & Fuad, Z. A. (2025). Critical thinking in the age of AI: A systematic review of AI's effects on higher education. *Educational Process: International Journal,* 14, e2025031. https://doi.org/10.22521/edupij.2025.14.31

Mollick, E., & Mollick, L. (2023). Assigning AI: Seven approaches for students, with prompts. *arXiv.* https://doi.org/10.48550/arXiv.2306.10052

Torrico Cuevas, D. (2026). Pensamiento crítico en la aplicación de la inteligencia artificial en estudiantes de la carrera de lingüística e idiomas sede Uyuni Universidad Autónoma Tomás Frías. *Revista Ciencia & Sociedad, 6*(S1), 5–16.

UNESCO. (2023). *ChatGPT e inteligencia artificial en la educación superior*: *Guía de inicio rápido.* UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000385146_spa

13. Aplicación de metodologías activas en la universidad y su relevancia en el contexto educativo superior

Ignasi Navarro Soria, Inés Rubio Sánchez Molina, Alba Gutiérrez Martínez, Manuel Torrecillas Martínez

Universidad de Alicante, Departamento de Psicología Evolutiva y Didáctica

RESUMEN

La presente investigación se justifica ante la creciente necesidad de analizar las diferencias en las percepciones del alumnado en función de la metodología empleada en la Enseñanza Superior, especialmente en la formación inicial del profesorado, contrastándolas con enfoques tradicionales. El estudio compara ambos modelos atendiendo a variables como la motivación académica, las estrategias de aprendizaje, los objetivos extrínsecos, la ansiedad, la percepción de la carga de trabajo y la idoneidad tanto de la evaluación como de la docencia. La muestra se compone de 88 estudiantes de tercer y cuarto curso de los grados de Maestro en Educación Infantil y Primaria de la Universidad de Alicante, distribuidos en un grupo experimental con los que se aplicaron metodologías activas, y un grupo control en el que se trabajó de manera expositiva. Para la recogida de datos se emplearon los cuestionarios *Motivated Strategies for Learning Questionnaire (MSLQ)* y el *Course Experience Questionnaire (CEQ)*, complementados con ítems de elaboración propia. La administración de los instrumentos se realizó de forma colectiva y anónima al finalizar la intervención docente. La hipótesis planteaba que el alumnado expuesto a metodologías activas lograría mejores resultados. Los resultados obtenidos confirmaron en gran medida estas expectativas, aunque se recomienda incorporar un pre y post test para reducir posibles sesgos. En conjunto, los hallazgos evidencian diferencias en las percepciones del alumnado que apuntan a mayores beneficios asociados al uso de metodologías activas en contextos universitarios.

PALABRAS CLAVE: Escuela Viva y Activa, Metodologías, Universidad, Enseñanza, Formación.

1. INTRODUCCIÓN

A lo largo de los últimos años, el continuo desarrollo social ha incrementado las exigencias para la inserción y desarrollo laboral de los profesionales, demandando no sólo una sólida formación en conocimientos teóricos y comunicativos, sino también habilidades para la resolución de problemas y el aprendizaje autónomo (Grosemans et al., 2017). Este contexto ha intensificado la competitividad en el acceso al mercado laboral (Rocha Gomes Varela, 2022), impulsando una mayor demanda de educación superior orientada a la formación de una ciudadanía preparada para una sociedad basada en el conocimiento (Revesado Caballares, 2018). Estas transformaciones se reflejan en las directrices del Espacio Europeo de Educación Superior, que promueve el protagonismo del estudiantado mediante el aprendizaje activo y la incorporación de estrategias formativas que favorezcan el desarrollo de competencias profesionales (Guerra Santana et al., 2019).

En el contexto educativo de educación superior, se ha detectado que las universidades españolas no potencian habilidades necesarias para esta INTRODUCCIÓN. A pesar de la abundante presencia de conocimientos teóricos y académicos impartidos, los estudiantes y titulados universitarios precisan de unas competencias más polivalentes, interpersonales y transversales, que faciliten la inserción en el mercado profesional actual (Marzo Navarro et al., 2008). En contraste con las metodologías di-

dácticas tradicionalmente empleadas en el sistema universitario, los enfoques activos satisfacen estas necesidades, obteniendo resultados más favorables en cuanto al desarrollo de competencias y los beneficios que aportan en los ámbitos académico, social, personal y profesional (Calvo Bernardino y Mingorance Arnáiz, 2009; González López et al., 2025; Hartikainen et al., 2019; Jiménez Hernández et al., 2020; Muntaner-Guasp et al., 2022; Palomares Ruiz, 2011; Orozco y Moriña, 2020).

Las metodologías activas promueven una conexión teórico-práctica de los contenidos en contextos significativos, materializándose en una mejor transferencia de los aprendizajes (Cabanillas-García, 2025). Además de ello, la literatura científica respalda que este enfoque favorece la autonomía (Luelmo del Castillo, 2018), el pensamiento creativo, actividades prácticas profesionales (López-Alegría y Fraile, 2023; O'Leary et al., 2023; Pinheiro y Valente, 2024), la participación (Hartikainen et al., 2019), la capacidad crítica (Zuñiga, 2025), el trabajo en equipo (Delgado Aza et al., 2024) y la motivación del alumnado (Tigse Parreño, 2019), mediante la adopción de un rol activo (Jiménez-Hernández et al., 2020). Siendo evidentes los beneficios que comporta la aplicación de metodologías activas en la enseñanza, tanto en la implicación educativa (Theobald et al., 2020), como en la adquisición de competencias del estudiantado (Mystakidis et al., 2019), resulta necesario plantear un enfoque metodológico centrado en el estudiante. De este modo, será imprescindible la adopción de una perspectiva acorde con las demandas actuales que permita ofrecer una educación de calidad en el ámbito de la Educación Superior (Jiménez Hernández et al., 2020).

Así pues, con el propósito de aportar evidencias para la mejora de la calidad educativa, resulta necesario analizar la percepción del estudiantado universitario sobre su desempeño en contextos con distintos enfoques metodológicos. La valoración de la aplicabilidad y utilidad de los contenidos permitirá establecer conclusiones acerca de las diferencias en las percepciones del alumnado y las potencialidades que presentan las metodologías activas frente a las tradicionales en la enseñanza universitaria.

2. MÉTODO

2.1. Descripción del contexto y de los participantes

La presente investigación se ha llevado a cabo en la Universidad de Alicante, concretamente en la Facultad de Educación. La muestra se compuso originalmente de un total de 88 estudiantes de tercer y cuarto curso de los grados de Maestro en Educación Infantil y en Educación Primaria, de todas las especialidades. Estos participantes fueron divididos en dos grupos según la asignatura cursada con un enfoque metodológico y profesorado específico. El grupo control, identificado como el Grupo 1, está constituido por alumnado que ha cursado asignaturas del primer y segundo año del grado, las cuales han empleado enfoques didácticos tradicionales basados en la clase magistral o exposición de la información. Originalmente este grupo contó con 50 participantes, pero se eliminó un sujeto debido a sus valores atípicos en el análisis estadístico, quedando un grupo de 49 participantes (nótese que el número de participantes final del Grupo 1 coincide numéricamente con el identificador del sujeto excluido; ambos son 49, pero hacen referencia a conceptos distintos).

Respecto a la comparabilidad entre grupos, conviene precisar que ambos pertenecen a los Grados de Maestro en Educación Infantil y en Educación Primaria de la Universidad de Alicante, aunque difieren en la asignatura cursada, el curso académico y el/la docente responsable. Dado que el equipo investigador impartió docencia en el grupo experimental, la investigación debe interpretarse como

un pilotaje exploratorio que prepara un diseño con medidas pretest y asignación controlada. Los resultados se leen, en consecuencia, como diferencias en las percepciones del alumnado según el enfoque metodológico recibido, sin que sea posible establecer relaciones causales a partir del diseño empleado.

En cuanto al grupo experimental, nombrado Grupo 2, está conformado por 38 estudiantes que cursaron la asignatura Problemas y Trastornos de Conducta a través del empleo de metodologías activas, conformando una muestra total de 87 sujetos.

2.2. Instrumentos

Para la recopilación de los datos necesarios para la valoración prevista, se ha diseñado un cuestionario a partir de la plataforma *Google Forms*, incluyendo ítems relevantes para la investigación presentes en los cuestionarios *Motivated Strategies for Learning Questionnaire (*MSLQ*,* Cardeñoso Ramírez et al., 2022), y *Course Experience Questionnaire (*CEQ*,* González et al., 2012), incluyendo preguntas de elaboración propia. El primero de los cuestionarios, *MSLQ*, se ha utilizado para medir la motivación (M), estrategias de aprendizaje (EA), objetivos extrínsecos (OE) y ansiedad (A). En cuanto al segundo, este se ha empleado para valorar la carga de trabajo (CT), la idoneidad de la evaluación (IE) y la adecuación de la docencia (AD). Algunos ítems de los cuestionarios se mostraron como invertidos, por lo que se tuvieron en cuenta para el análisis estadístico. Finalmente, las preguntas de elaboración propia (EP) se han diseñado para completar el estudio con otras percepciones relevantes. Cabe destacar que todos los ítems se responden en una escala *Likert* del 1 al 7, donde 1 es *Totalmente en desacuerdo* y 7 es *Totalmente de acuerdo*. Aunque la fiabilidad de estos instrumentos ha sido documentada en validaciones previas, se reportan a continuación los coeficientes de consistencia interna obtenidos en la presente muestra: los índices de consistencia interna calculados sobre la muestra del estudio (N = 88) fueron los siguientes: motivación $\alpha = .963$, estrategias de aprendizaje $\alpha = .921$, objetivos extrínsecos $\alpha = .708$, ansiedad $\alpha = .856$, carga de trabajo $\alpha = .741$, idoneidad de la evaluación $\alpha = .921$, adecuación de la docencia $\alpha = .965$. Todos los valores superan el umbral convencional de .70 (Nunnally, 1978), confirmando una consistencia interna adecuada de los instrumentos en esta muestra.

2.3. Procedimiento

Para recoger la máxima muestra posible, el equipo de investigación solicitó a través de la plataforma UACloud y aplicaciones de mensajería la participación del estudiantado, además de acudir a sesiones de clases. Tras recopilar los datos necesarios a partir del cuestionario, estos se analizaron según el grupo de investigación para su posterior comparación y valoración.

En lo relativo al análisis estadístico, previamente al mismo se examinó normalidad, homogeneidad de varianzas y la presencia de valores atípicos mediante diagramas de caja (boxplots). Se evaluó la normalidad de las distribuciones mediante la prueba de Shapiro–Wilk, así como a través del análisis de curtosis, asimetría e inspección visual de los gráficos Q-Q. Aunque algunas variables mostraron desviaciones respecto a la normalidad, ninguna alcanzó niveles suficientemente severos como para desaconsejar el uso de pruebas paramétricas teniendo en cuenta el tamaño de la muestra. Para controlar la homogeneidad de la varianza se empleó la *t* de Welch, recomendada por su robustez frente

a la heterogeneidad de varianzas y desviaciones moderadas de la normalidad (Delacre et al., 2017). Respecto a la identificación de los outliers, se identificaron pocos sujetos atípicos. Únicamente un participante presentó una puntuación estandarizada superior a |3| en una de las variables (sujeto 49), por lo que fue excluido de la muestra original para el análisis. Así pues, la muestra final quedó compuesta por 49 participantes en el Grupo 1 y 38 participantes en el Grupo 2. Para cada dimensión (M, EA, OE, AT, CT, IE y AD), se calcularon puntuaciones totales mediante la suma de los ítems correspondientes a cada constructo. En el caso de los ítems de elaboración propia (EP), estos se analizaron de forma independiente. Adicionalmente se realizó la prueba no paramétrica *U* de Mann–Whitney como análisis de sensibilidad, aunque fue la *t* de Welch la prueba principal que se utilizó.

3. RESULTADOS

Para una mejor visualización de los resultados obtenidos, se ha elaborado el siguiente gráfico que compara el grupo 1 control con el grupo experimental, mostrando una mejor valoración del alumnado del grupo experimental en la mayoría de los ítems, destacando una menor ansiedad, así como una puntuación más alta relacionada con los objetivos extrínsecos.

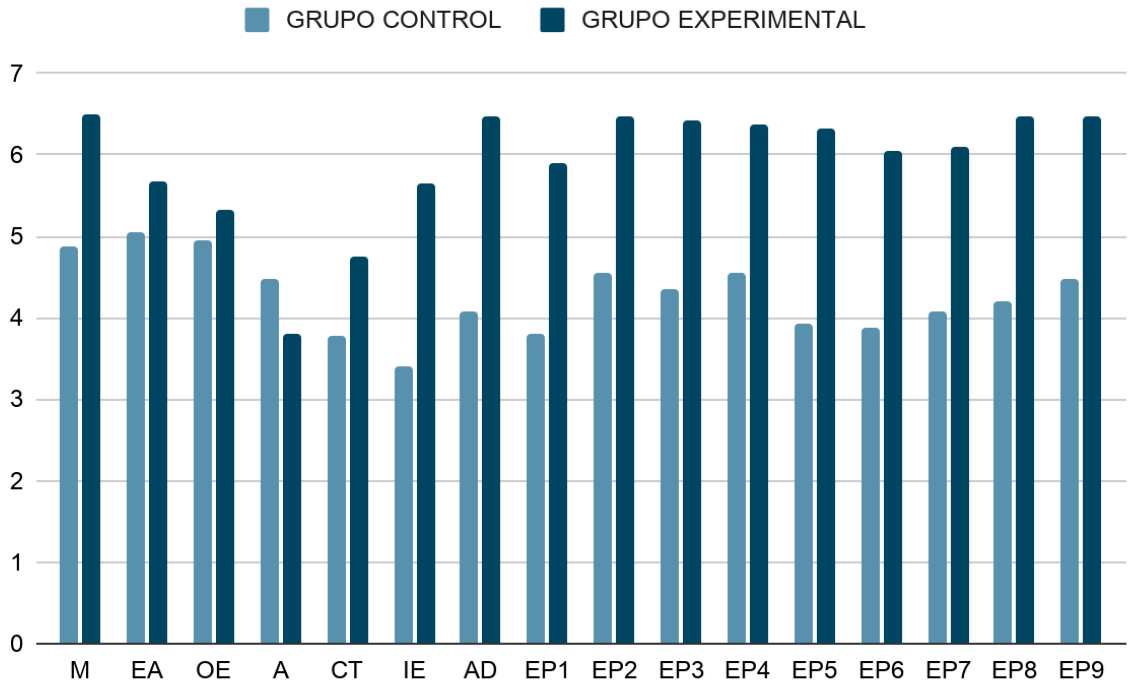


Gráfico 1. Comparación de los grupos a partir de las medias obtenidas en las dimensiones evaluadas.

A partir de la siguiente tabla se puede observar que todas las dimensiones mostraron diferencias estadísticamente significativas entre grupos, excepto la dedicada a los objetivos extrínsecos (OE) y la ansiedad (A). El tamaño de efecto obtenido a partir de las diferencias es considerablemente grande.

Tabla 1. Resultados obtenidos en los ítems de MSLQ y CEQ.

	Grupo 1		Grupo 2		t	gl	p	g de Hedges
	Media	**DT**	**Media**	**DT**				
M	98,57	23,34	130,26	7,40	-8,941	59,945	<,001	-1,725
EA	181,69	28,47	202,16	23,67	-3,659	84,552	<,001	-,766
OE	20,14	4,95	21,32	4,12	-1,206	84,528	,231	
A	17,88	6,01	15,21	7,92	1,727	67,128	,089	
CT	14,94	5,39	18,47	4,08	-3,478	84,964	<,001	-,720
IE	10,37	5,12	16,95	4,62	-6,284	82,993	<,001	-1,329
AD	20,57	9,01	32,42	5,54	-7,990	74,201	<,001	-1,586

M = motivación; EA = estrategias de aprendizaje; OE = objetivos extrínsecos; A = ansiedad; CT = carga de trabajo; IE = idoneidad de la evaluación; AD = adecuación de la docencia; DT = desviación típica. Nota. La reducida DT de motivación en el Grupo 2 refleja un efecto techo: la media (130,26) se aproxima al máximo teórico (140), concentrando las puntuaciones en la parte alta de la escala.

Respecto a los ítems de elaboración propia (EP), todos mostraron diferencias significativas entre grupos tal como se puede observar en la Tabla 2; destacando un tamaño del efecto muy grande.

Tabla 2. Resultados obtenidos en los ítems de elaboración propia.

	Grupo 1		Grupo 2		t	gl	p	g de Hedges
	Media	**DT**	**Media**	**DT**				
EP1	3,82	1,822	5,89	1,134	-6,521	81,536	<,001	-1,320
EP2	4,59	1,682	6,47	,687	-7,104	66,864	<,001	-1,389
EP3	4,41	1,892	6,42	,758	-6,778	66,246	<,001	-1,324
EP4	4,59	1,719	6,37	,883	-6,249	74,958	<,001	-1,243
EP5	3,98	1,909	6,32	1,233	-6,908	82,531	<,001	-1,404
EP6	3,92	1,801	6,05	1,293	-6,429	84,558	<,001	-1,322
EP7	4,12	1,844	6,11	1,085	-6,257	79,796	<,001	-1,260
EP8	4,24	1,786	6,47	,687	-8,006	65,008	<,001	-1,560
EP9	4,51	1,647	6,47	,687	-7,540	67,549	<,001	-1,476

EP = elaboración propia; DT = desviación típica.

Tanto el análisis con la prueba paramétrica (*t* de Welch) como la prueba no paramétrica (*U* de Mann-Whitney) coincidieron en la identificación de diferencias significativas entre grupos en las mismas variables.

4. DISCUSIÓN Y CONCLUSIONES

Tras la obtención de los resultados, se pueden observar diferencias individuales que pueden ser justificadas con las evidencias obtenidas en otras investigaciones a favor del uso de metodologías activas.

En primer lugar, la motivación medida muestra una diferencia significativa entre los dos grupos, considerando el grupo experimental aquel que presenta una mayor motivación hacia la asignatura y la enseñanza. Así pues, se observa una concordancia con afirmaciones que señalan una mayor motivación en el estudiantado que recibe una enseñanza activa (González López at al., 2025). Por otra parte, aunque las diferencias observadas sean mínimas entre los grupos control y experimental, siguen siendo superiores las percepciones del alumnado en referencia a las estrategias de aprendizaje, evidenciando una mayor toma de conciencia de su propio proceso de aprendizaje o metacognición, defendiendo el uso de las metodologías activas (Hartikainen et al., 2019). De igual manera, este último grupo muestra una adquisición mayor, aunque poco significativa, de objetivos extrínsecos, asociando su rendimiento hacia la obtención de refuerzos externos.

Por otra parte, es necesario destacar que el alumnado del grupo control presenta valores de ansiedad superiores en lo que respecta a la realización de exámenes, vinculándose con la realización de pruebas escritas como instrumento de evaluación con mayor valor en la asignatura, midiendo conocimientos principalmente teóricos, asociados a modelos de enseñanza tradicionales (Holmos-Flores et al., 2023).

La percepción de la carga de trabajo recogida en el segundo cuestionario muestra diferencias claras entre los grupos. Considerando que cuanto más elevada sea la puntuación total obtenida mejor es la percepción del alumnado, el grupo experimental posee una mejor percepción de la carga de trabajo a diferencia del grupo control, enfatizando en la consideración de una mayor carga de trabajo en este último. Además, este resultado contrasta con la literatura reciente, puesto que no existe consenso en lo que se refiere a la carga de trabajo que supone la evaluación formativa, promovida con el uso de las metodologías activas. Aunque los resultados obtenidos defienden que la evaluación formativa no incrementa de manera significativa el esfuerzo requerido por parte del alumnado, otros autores afirman que este enfoque implica una mayor dedicación de trabajo en contraste con los MÉTODOs tradicionales, los cuales suelen basarse en exámenes finales, y conllevar una carga mayor en el último período cuatrimestral (Romero-Martín et al., 2024). Sin embargo, para la obtención de información relevante, sería conveniente estudiar este aspecto teniendo en cuenta el sistema de evaluación por asignatura y el grado universitario cursado.

Continuando con la idoneidad de la evaluación (IE), los resultados obtenidos muestran diferencias notables. Siguiendo un modelo de evaluación formativa que caracteriza al uso de metodologías activas, se observa que el alumnado del grupo experimental presenta una actitud más positiva respecto al mismo, a diferencia de la de carácter sumativo empleada habitualmente en la metodología tradicional. Por ello, es necesario destacar que esta evaluación formativa comporta beneficios considerables que posiblemente han favorecido esta mejor percepción. Además de permitir la obtención de una visión global del progreso de los aprendices en relación a la comprensión de relaciones, procesos de razonamiento y destrezas complejas, contrastando con el mero análisis superficial del recuerdo de fragmentos de conocimientos (Manassero-Mas y Vázquez-Alonso, 2020), esta continuidad de la formación favorece el aprendizaje permanente a lo largo de la vida (OECD, 2008).

Por último, en la evaluación de la percepción de la adecuación de la docencia (AD), se observa una diferencia significativa a favor de las metodologías activas. Así como el tipo de docencia impartida varía de forma considerable entre las metodologías activas y las tradicionales, cabe destacar que, en

este último enfoque, el docente emplea una estrategia discursiva expositiva y de forma ininterrumpida, así como el alumnado asume un papel pasivo de memorización de información (Martínez Cervantes, 2024), justificando con estas afirmaciones los resultados obtenidos en el grupo control.

Los ítems de elaboración propia (EP1-EP9) apuntan en la misma dirección que las escalas estandarizadas. El grupo experimental percibe mayor autonomía, libertad de elección y pensamiento crítico (EP1, EP3), en consonancia con el rol activo que promueven estas metodologías (Delgado Aza et al., 2024). La percepción de trabajo en equipo y habilidades interpersonales también resulta más positiva (EP2, EP4, EP8, EP9), lo que refuerza la idea de que la interacción grupal favorece la cooperación, las habilidades comunicativas y la retención de conocimientos (Jiménez-Hernández et al., 2020; Guerra Santana et al., 2019). Finalmente, los contenidos se perciben como más útiles (EP5), aplicables a situaciones reales (EP6) y mejor conectados con los conocimientos previos (EP7), apoyando la transferencia de aprendizajes característica de estos enfoques (González López et al., 2025; Delgado Aza et al., 2024).

A partir del análisis exhaustivo de los resultados, no es posible establecer conclusiones completamente sólidas acerca de las percepciones del alumnado respecto al aprendizaje en relación con la metodología didáctica empleada, dado el reducido tamaño muestral. De manera general, el grupo experimental obtuvo puntuaciones superiores en todas las dimensiones, excepto en la relativa a la ansiedad, interpretándose como un resultado positivo para este grupo. Sin embargo, para poder extrapolar los hallazgos a otros contextos, resultaría necesario considerar un aumento del tamaño muestral para aumentar la significatividad de los resultados. Además, a pesar de las limitaciones presentes en el pilotaje de la prueba, los resultados obtenidos contribuyen a esclarecer la posible relación entre la aplicación de metodologías activas, el desarrollo competencial y los logros de aprendizaje (Hartikainen et al., 2019), asentando las bases para futuras investigaciones. Así pues, esta investigación se concentra en el pilotaje que permitirá impulsar la medición, de manera más precisa, de las percepciones del alumnado con respecto al aprendizaje. De esta manera, las siguientes investigaciones se orientan a la realización de un pretest al inicio de la asignatura, de manera adicional, comparando también los resultados al finalizarla, reduciendo las posibilidades de sesgo, así como facilitando una comparación más precisa de los resultados y perfilando el instrumento que permita continuar el estudio con eficacia.

Desde un punto de vista de las implicaciones educativas, los hallazgos orientan el diseño de la docencia universitaria en los grados de formación del profesorado. La mayor motivación y las estrategias de aprendizaje más elaboradas percibidas por el grupo experimental sugieren que las metodologías activas pueden contribuir a desarrollar un perfil docente más reflexivo y autónomo (Jiménez-Hernández et al., 2020). La reducción de la ansiedad ante la evaluación en ese mismo grupo apunta a que los sistemas de evaluación formativa generan entornos percibidos como más seguros y justos (Holmos-Flores et al., 2023; Romero-Martín et al., 2024). Por último, las diferencias en la percepción de adecuación docente indican que el alumnado valora positivamente el rol facilitador del profesorado, lo que refuerza la necesidad de programas de formación continuada centrados en la transferencia metodológica activa al aula universitaria.

En definitiva, esta investigación ofrece un marco sólido para comprender cómo las metodologías didácticas empleadas inciden en el aprendizaje del estudiantado universitario a través de una comparativa de sus percepciones. Sus aportaciones pueden orientar el diseño de propuestas pedagógicas innovadoras y estimular nuevas líneas de investigación centradas en la formación y actualización docente. Así pues, el estudio contribuye a impulsar una transformación educativa más reflexiva y sostenible en el ámbito de la educación superior, incidiendo en el mayor protagonismo del estudiante en el proceso de aprendizaje.

5. REFERENCIAS.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., y Rothstein, H. R. (2009). *Introduction to meta-analysis*. Wiley. https://doi.org/10.1002/9780470743386

Cabanillas-García, J. L. (2025). The Application of Active Methodologies in Spain: An Investigation of Teachers' Use, Perceived Student Acceptance, Attitude, and Training Needs Across Various Educational Levels. *Education Sciences*, *15*(2), 210. https://doi.org/10.3390/educsci15020210

Calvo Bernardino A. y Mingorance Arnáiz A. C. (2009). La estrategia de las universidades frente al Espacio Europeo de Educación Superior. *Revista Complutense de Educación*, *20*(2), 319-342. https://revistas.ucm.es/index.php/RCED/article/view/RCED0909220319A

Cardeños Ramírez, O., Larruzea-Urkixo, N., y Bully Garay, P. B. (2022). Adaptation to the Spanish university context and psychometric properties of the MSLQ: Contributions to the measurement and analysis of gender differences of self-regulated learning. *Anales de Psicología, 38*(2), 295-306. https://doi.org/10.6018/analesps.444851

Delacre, M., Lakens, D., y Leys, C. (2017). Why Psychologists Should by Default Use Welch's t-test Instead of Student's t-test. *International Review of Social Psychology, 30*(1), 92-101. https://doi.org/10.5334/irsp.82

Delgado Aza, T. S., Rosero Bustos, A. M., Rocha Tayupanta, G. E., Suasnavas Reina, S. A., Maldonado Cruz, M. de J., y Santillán Vaca, H. S. (2024). Implementación de Metodologías Activas en el Proceso de Enseñanza-Aprendizaje en el Aula. *Revista Científica Multidisciplinaria G-nerando, 5*(1), 1027-1046. https://doi.org/10.60100/rcmg.v5i1.238

González, C., López, L., y Montenegro, H. (2012) Análisis de confiabilidad y de validez del instrumento Course Experience Questionnaire (CEQ). *Educación Y Educadores*, *15*(1), 63-78. https://doi.org/10.5294/edu.2012.15.1.4

González López, A. M., Pascual Sevillano, M. Á., y Sorzio, P. (2025). Use of Active Methodologies in Basic Education: An Umbrella Review. *Education Sciences*, *15*(11), 1536. https://doi.org/10.3390/educsci15111536

Grosemans, I., Coertjens, L., y Kyndt, E. (2017). Exploring learning and fit in the transition from higher education to the labour market: A systematic review. *Educational Research Review*, *21*, 67-84.https://doi.org/10.1016/j.edurev.2017.03.001

Guerra Santana, M., Rodríguez Pulido, J., y Artiles Rodríguez, J. (2019). Aprendizaje colaborativo: experiencia innovadora en el alumnado universitario. *Revista de Estudios y Experiencias En Educación*, *18*(36), 269-281. https://doi.org/10.21703/rexe.20191836guerra5

Hartikainen, S., Rintala, H., Pylväs, L., y Nokelainen, P. (2019). The Concept of Active Learning and the Measurement of Learning Outcomes: A Review of Research in Engineering Higher Education. *Education Sciences*, *9*(4), 276. https://doi.org/10.3390/educsci9040276

Holmos-Flores, E., Atencio-Gonzáles, R. E., Espinoza-Moreno, T. M., y Abarca-Arias, Y. M. (2023). Evaluación alternativa y evaluación tradicional en el contexto de la educación universitaria. *Revista Arbitrada Interdisciplinaria KOINONIA, 8*(16), 220-237. https://doi.org/10.35381/r.k.v8i16.2546

Jiménez-Hernández, D., González Ortiz, J. J., y Tornel Abellán, M. (2020). Metodologías activas en la universidad y su relación con los enfoques de enseñanza. *Profesorado. Revista de currículum y formación del profesorado*, *24*(1), 76-94. https://doi.org/10.30827/profesorado.v24i1.8173

López-Alegría, F., y Fraile, C. (2023). Metodologías didácticas activas frente a paradigma tradicional. Una revisión sistemática. *Educación Médica*, *26*(1), 5. https://doi.org/10.33588/fem.261.1255

Luelmo del Castillo, M. J. (2018). Origen y desarrollo de las metodologías activas dentro del sistema educativo español. *Encuentro*, *27*, 4-21. https://doi.org/10.37536/ej.2018.27.1890

Manassero-Mas, M. A., y Vázquez-Alonso, Á. (2020). Desarrollo curricular de las competencias clave: su evaluación para el aprendizaje desde la normativa educativa. *Enseñanza & Teaching Revista Interuniversitaria de Didáctica*, *38*(1), 29-48. https://doi.org/10.14201/et20203812948

Martínez Cervantes, A. (2024). Metodologías activas: estrategias didácticas utilizadas por docentes en educación básica en Oaxaca. *Revista Reflexión e Investigación Educacional, 6*(2), 1-14. https://doi.org/10.22320/reined.v6i1.6491

Marzo Navarro, M., Pedraja Iglesias, M., y Rivera Torres, P. (2008). Habilidades y competencias demandadas por el mercado laboral. *ICADE. Revista Cuatrimestral de las Facultades de Derecho y Ciencias Económicas y Empresariales,* (73), 347-363. https://revistas.comillas.edu/revistaicade/article/view/381

Muntaner-Guasp, J. J., Mut-Amengual, B., y Pinya-Medina, C. (2022). Las metodologías activas para la implementación de la educación inclusiva. *Revista electrónica Educare*, *26*(2), 1–21. https://doi.org/10.15359/ree.26-2.5

Mystakidis, S., Berki, E., y Valtanen, J. (2019). The Patras Blended Strategy Model for Deep and Meaningful Learning in Quality Life-Long Distance Education. *The Electronic Journal Of e-Learning*, *17*(2). https://doi.org/10.34190/jel.17.2.01

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). McGraw-Hill.

OECD. (2008). *Assessment for Learning Formative Assessment. OECD/CERI International Conference Learning in the 21st Century: Research, Innovation and Policy.* [Archivo PDF]. https://one.oecd.org/document/EDU/CERI/CD(2008)18/en/pdf

O'Leary, M., Cui, V., Kiem, M. T., Dang, D. T., Nguyen, G. T. H., y Thi Hoang, K. H. (2023). The role of classroom observation in the development and assessment of schoolteachers in Vietnam: a review of national policy and research. *Asia-Pacific Journal of Teacher Education*, *51*(4), 395–407. https://doi.org/10.1080/1359866X.2023.2191307

Orozco, I., y Moriña, A. (2020). Estrategias metodológicas que promueven la inclusión en educación infantil, primaria y secundaria. *Revista Internacional de Educación para la Justicia Social, 9*(1), 81–98.https://doi.org/10.15366/riejs2020.9.1.004

Palomares Ruiz, A. (2011). El modelo docente universitario y el uso de nuevas metodologías en la enseñanza, aprendizaje y evaluación. *Revista de Educación*, (355), https://doi.org/10.4438/1988-592X-RE-2011-355-038

Pinheiro, W. S., y Valente, E. A. T. (2024). Metodologias ativas no âmbito da Educação Básica: uma revisão sistemática de literatura. *Revista Caribeña de Ciencias Sociales*, *13*(12), 4385. https://doi.org/10.55905/rcssv13n12-010

Revesado Carballares, D. (2018). El proceso de Bolonia y la admisión a la universidad. La realidad del sistema universitario español en perspectiva comparada. *Revista Española de Educación Comparada* (32), 169-180. https://doi.org/10.5944/reec.32.2018.20948

Rocha Gomes Varela, T. M. (2022). Creative projects in training in working context. *Conference Proceedings. CIVINEDU 2022*, 136-141. https://doi.org/10.58909/adc22935203

Romero-Martín, R., Fraile-Aranda, A., López-Pastor, V-M., y Castejón-Oliva, F-J. (2024). The relationship between formative assessment systems, academic performance and teacher and student workloads in higher education. *Infancia y Aprendizaje, 37*(2), 310-341. https://doi.org/10.1080/02103702.2014.918818

Theobald, E. J., Hill, M. J., Tran, E., Agrawal, S., Arroyo, E. N., Behling, S., Chambwe, N., Cintrón, D. L., Cooper, J. D., Dunster, G., Grummer, J. A., Hennessey, K., Hsiao, J., Iranon, N., Jones II, L., Jordt, H., Keller, M., Lacey, M. E., Littlefield, C. E., … Freeman, S. (2020). Active learning narrows achievement gaps for underrepresented students in undergraduate science, technology, engineering, and math. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(12), 6476-6483. https://doi.org/10.1073/pnas.1916903117

Tigse Parreño, C. M. (2019). El constructivismo, según bases teóricas de César Coll. *Revista Andina de Educación*, *2*(1), 25-28. https://doi.org/10.32719/26312816.2019.2.1.4

Zuñiga, J. (2025). Metodologías activas para el desarrollo de pensamiento crítico en la universidad. Una revisión de literatura. *Espacios*, *46*(4), 129-140. https://doi.org/10.48082/espacios-a25v46n04p13

14. Programación colaborativa en ingeniería apoyada por inteligencia artificial: innovación educativa para un aprendizaje significativo[1]

J. Pérez Beltrán[1], R.P. Ñeco García[2], N.M. García Aracil[2], J.M. Catalán Orts[2], A. Muñoz Lozano[1]

[1]*Universidad de Alicante*
[2]*Universidad Miguel Hernández de Elche*

RESUMEN

En las titulaciones de ingeniería en las que la programación tiene un carácter instrumental, el alumnado suele encontrar dificultades para traducir problemas técnicos descritos en lenguaje natural a modelos lógicos, algoritmos y programas funcionales. Este trabajo presenta una metodología de programación colaborativa apoyada por herramientas digitales y simulaciones conversacionales con inteligencia artificial, orientada a mejorar la formalización de requisitos y el diseño de soluciones de automatización. La experiencia se desarrolla principalmente en una asignatura universitaria de Automatización Industrial de un grado de ingeniería eléctrica y se complementa con adaptaciones a educación secundaria, con el fin de valorar la transferencia del enfoque entre niveles educativos. El alumnado trabaja en grupos reducidos para analizar procesos industriales que deben ser automatizados y, para ello, deben programarse autómatas programables. Como apoyo, se incorpora un agente conversacional que simula a un cliente no técnico y favorece la formulación de preguntas, la negociación de interpretaciones y el consenso sobre el comportamiento del sistema. Los resultados muestran una mejora de las calificaciones, una disminución del número de estudiantes con resultados insuficientes y una mayor implicación del alumnado en las actividades colaborativas. Se concluye que la programación colaborativa, reforzada mediante el uso complementario de inteligencia artificial conversacional, contribuye a mejorar la formalización de problemas y acerca el aprendizaje a dinámicas propias de la práctica profesional en ingeniería.

PALABRAS CLAVE: aprendizaje colaborativo, inteligencia artificial, automatización industrial, enseñanza de programación.

1. INTRODUCCIÓN

La enseñanza de la programación en titulaciones de ingeniería en las que esta materia no constituye el eje central del plan de estudios plantea dificultades que van más allá del aprendizaje de la sintaxis de un lenguaje o del manejo de un entorno de desarrollo. En asignaturas aplicadas, como las relacionadas con la automatización industrial, el alumnado debe analizar procesos físicos, identificar entradas y salidas, establecer secuencias de funcionamiento, prever situaciones alternativas y traducir todo ello a modelos formales y programas ejecutables. Esta transición desde una descripción inicial, generalmente expresada en lenguaje natural, hasta una solución técnica estructurada constituye uno de los aspectos más complejos del aprendizaje, especialmente cuando el estudiante todavía no distingue con seguridad entre la información esencial, la información ambigua y los

1 El presente trabajo ha contado con una ayuda del Programa de Redes de investigación en docencia universitaria del Instituto de Ciencias de la Educación de la Universidad de Alicante (convocatoria 2024). Ref.: 6193

aspectos que deben ser aclarados antes de programar.

Esta dificultad resulta relevante porque el trabajo profesional en ingeniería rara vez parte de enunciados perfectamente cerrados. En situaciones reales, los requisitos de un sistema suelen aparecer de forma incompleta, informal o incluso contradictoria, y deben clarificarse mediante diálogo con clientes, responsables de planta, operadores u otros agentes implicados. Sin embargo, en el contexto académico, los ejercicios suelen presentarse como enunciados ya depurados, lo que facilita la corrección, pero limita el desarrollo de competencias vinculadas al análisis de requisitos y la toma de decisiones compartida. Por ello, este trabajo plantea una metodología basada en programación colaborativa, apoyada por simulaciones conversacionales con inteligencia artificial, para que el alumnado afronte de forma guiada la formalización y programación de problemas de automatización.

Diversos estudios sobre enseñanza de la programación coinciden en que buena parte de las dificultades del alumnado no se limita al aprendizaje de la sintaxis, sino que aparece al analizar el problema, formalizarlo y trasladarlo a una solución estructurada (Chuang & Chang, 2024; Gao et al., 2025; Huang et al., 2024; Qian & Choi, 2023). Estas dificultades pueden acentuarse cuando la programación se imparte como herramienta instrumental dentro de titulaciones técnicas no informáticas, ya que el alumnado debe relacionar simultáneamente conceptos de programación, funcionamiento físico de sistemas, representación formal y validación práctica. En este contexto, los enfoques exclusivamente individuales pueden reducir las oportunidades de contrastar interpretaciones y corregir ideas parciales mediante la interacción con otros compañeros.

Frente a este enfoque de trabajo individual, el aprendizaje colaborativo se ha consolidado como una estrategia eficaz para favorecer la comprensión de conceptos básicos, la motivación y el desarrollo de competencias transversales. Diversos trabajos han mostrado que la colaboración estructurada favorece la reflexión compartida, la revisión de las soluciones y un aprendizaje más activo y significativo (Alacapınar & Uysal, 2020; Lai & Wong, 2022; Mercier et al., 2023; Ralston et al., 2017). En el ámbito concreto de la programación, la programación colaborativa y la resolución conjunta de problemas han demostrado tener efectos positivos sobre el rendimiento académico (Krismadinata et al., 2023; Silva et al., 2020; Sun & Xu, 2025; Xu & Correia, 2024). Estas aportaciones resultan especialmente coherentes con la automatización industrial, donde el diseño de un programa no puede separarse del análisis previo del proceso y la validación del comportamiento esperado.

Las herramientas digitales refuerzan este enfoque al facilitar la documentación compartida y el seguimiento del trabajo de los grupos. A ello se suma la inteligencia artificial conversacional, que puede utilizarse para simular interlocutores no técnicos que describen un problema de manera informal. De este modo, la IA no sustituye el razonamiento del estudiante, sino que introduce incertidumbre controlada y estimula el diálogo técnico dentro del equipo.

El presente trabajo se sitúa en esta línea y propone una experiencia docente centrada en la programación colaborativa en ingeniería, ampliada mediante un agente conversacional que simula el papel de un cliente o responsable de planta no técnico. El agente proporciona información parcial o ambigua sobre el funcionamiento deseado de un sistema industrial y los estudiantes, organizados en grupos reducidos, deben entrevistarlo, identificar lagunas en la especificación, contrastar interpretaciones y acordar una descripción funcional suficientemente precisa. Posteriormente, esta descripción se traduce en herramientas propias de la automatización, como listas de entradas y salidas, diagramas GRAFCET (*Graphe Fonctionnel de Commande Étape/Transition*, usados para representar de forma estructurada el funcionamiento secuencial del sistema). La experiencia se

desarrolla principalmente en una asignatura universitaria de Automatización Industrial y se considera también su posible transferencia a contextos de educación secundaria relacionados con la enseñanza de la programación.

El objetivo general de este trabajo es analizar en qué medida una metodología de programación colaborativa, apoyada por simulaciones conversacionales con inteligencia artificial, puede favorecer un aprendizaje significativo en la programación aplicada a sistemas de automatización industrial. De forma específica, se pretende: identificar las dificultades del alumnado al pasar de descripciones informales a modelos formales de funcionamiento; diseñar una secuencia de trabajo colaborativo en la que el agente conversacional actúe como desencadenante de preguntas y discusión técnica; valorar el impacto de la metodología sobre la calidad de los diseños, la implicación del alumnado y el rendimiento académico; y estudiar la posible transferencia del enfoque a otros niveles educativos. La hipótesis de partida es que la programación colaborativa, reforzada con inteligencia artificial conversacional, puede ayudar al alumnado a comprender mejor los problemas y trabajar de una forma más próxima a la práctica profesional de la ingeniería. Esta hipótesis se fundamenta en las evidencias sobre los beneficios de la resolución colaborativa de problemas de programación (Lai & Wong, 2022; Sun & Xu, 2025). A partir de estos antecedentes, se propone utilizar el agente conversacional como interlocutor para promover la discusión y la especificación de requisitos dentro del grupo.

2. MÉTODO

La metodología empleada se basa en una experiencia docente real, orientada a analizar el efecto de la programación colaborativa en el aprendizaje de la programación aplicada a la automatización industrial. El estudio adopta un enfoque mixto, en el que los resultados académicos se complementan con evidencias cualitativas sobre el trabajo desarrollado por los grupos y la valoración del alumnado. La inteligencia artificial conversacional se utiliza para simular una situación profesional realista: un cliente o responsable de planta explica lo que necesita desde un punto de vista no técnico, y el grupo debe hacer preguntas para concretar cómo debe funcionar el sistema antes de diseñarlo y programarlo.

La actividad se organiza como una secuencia en la que los estudiantes parten de una descripción inicial de un sistema industrial y avanzan hasta obtener una solución formal, programada y verificable. Para ello, los estudiantes trabajan en grupos reducidos, utilizan herramientas propias de la automatización industrial e interactúan con el agente conversacional para aclarar requisitos. La metodología permite valorar tanto la solución final como el proceso seguido por el grupo para transformar el problema inicial en un programa funcional.

2.1. Descripción del contexto y de los participantes

La experiencia principal se ha desarrollado en una asignatura de Automatización Industrial de un grado de ingeniería eléctrica impartido en una universidad pública española. Se trata de una asignatura aplicada, situada en un curso intermedio del plan de estudios (tercer curso de un total de cuatro cursos), en la que uno de los objetivos es el aprendizaje del diseño de automatismos industriales mediante diagramas GRAFCET, así como su programación en autómatas programables y su validación práctica sobre sistemas físicos o simulados. El estudiantado posee conocimientos

previos básicos de programación y de sistemas (adquiridos en los cursos primero y segundo, fundamentalmente los lenguajes Python y Matlab), pero presenta una experiencia limitada en la traducción de descripciones funcionales abiertas a soluciones de automatización o programación estructuradas.

El análisis del rendimiento académico comprende los cursos 2022-23, 2023-24, 2024-25 y 2025-26, con 19, 12, 18 y 17 estudiantes con calificaciones disponibles en cada curso, respectivamente. Los datos corresponden a las actividades y pruebas realizadas en las convocatorias ordinaria y extraordinaria del correspondiente curso académico. Para el cálculo de las medias y los porcentajes se incluyeron los estudiantes presentados, considerándose como estudiante presentado el que asistió al menos al 70 % de las actividades y realizó el examen final. La descripción de los participantes se basa en sus características académicas, dado que no se dispone de datos sociodemográficos sistematizados sobre su edad o sexo.

La metodología se ha aplicado de forma progresiva en sesiones prácticas de laboratorio, mediante grupos de tres o cuatro estudiantes. Esta organización favorece que el alumnado participe activamente y construya las soluciones mediante discusión y trabajo compartido. La experiencia descrita parte de dinámicas colaborativas aplicadas en ediciones anteriores de la asignatura y las adapta a nuevas actividades de análisis y diseño, incorporando como ampliación metodológica el agente conversacional de inteligencia artificial. La metodología colaborativa se introdujo progresivamente desde el curso académico 2022-23, mientras que las actividades con el agente conversacional se incorporaron en el curso 2025-26 y contaron con la participación de 17 estudiantes.

De forma complementaria, se consideran experiencias relacionadas en educación secundaria, dentro de asignaturas de programación de cursos previos a la universidad. En este contexto, la metodología se adapta al nivel del alumnado y a las herramientas disponibles, manteniendo sus principios básicos: trabajo colaborativo, análisis compartido del problema y formalización progresiva de soluciones. Esta extensión permite valorar la posible transferencia del enfoque a etapas educativas en las que la programación también se aprende como herramienta para resolver problemas, aunque con menor complejidad técnica que en el ámbito universitario.

2.2. Instrumentos

En el contexto universitario se han utilizado los recursos habituales para el diseño y validación de automatismos industriales, combinando herramientas de representación y programación con maquetas físicas o sistemas simulados. Entre ellos se incluyen los diagramas GRAFCET, el entorno TIA Portal, los autómatas Siemens S7-1200 y los sistemas de laboratorio empleados para comprobar el funcionamiento de los programas. Estos recursos permiten vincular la programación colaborativa con situaciones próximas a la práctica profesional, en las que el diseño debe validarse sobre un sistema físico o simulado.

A lo largo de la experiencia se han planteado distintas prácticas de automatización, como el control del tráfico en una carretera en obras, el lavado industrial de piezas metálicas o la dosificación y transporte mediante tolvas, depósito y cintas transportadoras. Estos casos permiten trabajar problemas secuenciales con distintos niveles de complejidad. Como ejemplo representativo, la figura 1 muestra el esquema funcional y el equipo de laboratorio correspondiente a la práctica de dosificación y transporte.

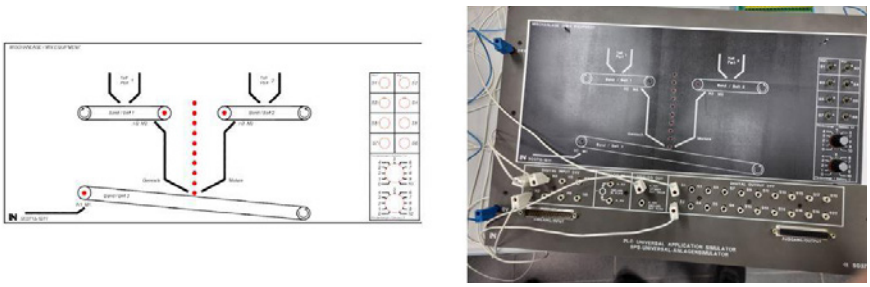


Figura 1. Sistema de dosificación y transporte empleado en las actividades prácticas: esquema funcional (izquierda) y equipo de laboratorio (derecha). Fuente: esquema adaptado del manual del simulador universal de sistemas de Lucas-Nülle (Fleumer, 2006); fotografía propia.

La figura 2 muestra una representación más próxima al proceso industrial que sirve de referencia para la misma actividad. Esta imagen permite contextualizar el problema y facilita la INTRODUCCIÓN de descripciones iniciales incompletas o ambiguas, similares a las que pueden aparecer en una situación profesional. A partir de estos sistemas, los grupos deben formular preguntas y aclarar requisitos antes de transformar la descripción inicial en un diseño GRAFCET y en un programa de control.



Figura 2. Representación del proceso industrial de dosificación y transporte utilizado para el análisis de requisitos. Fuente: imagen generada con ChatGPT y modificada y revisada por los autores.

Junto a los instrumentos técnicos se ha utilizado *Moodle* como plataforma de apoyo para distribuir materiales, recoger entregas y realizar cuestionarios. Además, las *wikis* y documentos compartidos han permitido documentar el proceso seguido por cada grupo.

El agente conversacional se utiliza como apoyo en la fase de análisis de requisitos. Como ya se ha comentado previamente, sus respuestas reproducen el tipo de explicación que podría ofrecer un cliente o responsable de planta desde una perspectiva no técnica, de modo que los grupos deben formular preguntas y concretar el funcionamiento esperado antes de diseñar el automatismo. Su papel no es resolver la actividad, sino hacer más realista el proceso de clarificación previo a la programación. Este agente conversacional se implementó mediante una versión personalizada de ChatGPT (aunque podrían emplearse otras plataformas de inteligencia artificial conversacional que permitan configurar el papel del interlocutor).

La información analizada en este trabajo comprende las calificaciones de las actividades prácticas y las pruebas de evaluación, complementadas por la observación en el laboratorio y la revisión de los diseños y la documentación elaborada por los grupos.

2.3. Procedimiento

El procedimiento seguido en la experiencia universitaria se organiza en varias fases. En la fase inicial, el alumnado trabaja con problemas de automatización planteados mediante enunciados cerrados, lo que le permite familiarizarse con el entorno de programación, el uso de los autómatas y la representación mediante GRAFCET. Esta primera etapa también sirve para detectar dificultades habituales, como comenzar a programar sin haber definido previamente el funcionamiento del sistema o confundir entradas (sensores) y salidas (actuadores) del sistema que se va a controlar.

En una segunda fase se introducen actividades de análisis colaborativo a partir de descripciones en lenguaje natural de procesos industriales que deben automatizarse. Antes de elaborar el GRAFCET o definir las entradas y salidas, cada grupo debe discutir qué información resulta suficiente, qué aspectos necesitan aclaración y qué preguntas conviene formular. De este modo, la actividad se centra primero en comprender el problema y construir una interpretación compartida, antes de pasar al diseño y la programación.

Posteriormente, en la tercera fase, se incorpora el agente conversacional de inteligencia artificial. Los grupos interactúan con él como si se tratara de un cliente o responsable de planta que conoce el funcionamiento deseado del sistema, pero no siempre lo expresa con precisión técnica. La interacción se realiza a nivel de grupo, no como consulta individual, para preservar el carácter colaborativo de la actividad. El alumnado debe formular preguntas, analizar las respuestas y decidir si la información disponible es suficiente para avanzar, evitando que la inteligencia artificial actúe como sustituto del razonamiento propio.

Una vez aclarados los requisitos, cada grupo elabora la documentación técnica del automatismo, que incluye la descripción funcional acordada, la lista de entradas y salidas, el diagrama GRAFCET y la justificación de las principales decisiones de diseño. Durante esta fase, el profesorado acompaña el trabajo de los grupos, revisa su avance y plantea preguntas que ayudan a detectar inconsistencias, sin proporcionar directamente la solución.

A continuación, los estudiantes transforman el diseño realizado en un programa para el autómata. La programación se desarrolla de forma colaborativa, con reparto de tareas y revisión conjunta del resultado. Después, el programa se prueba sobre la maqueta física o el sistema simulado, comprobando si el comportamiento obtenido coincide con los requisitos acordados. Cuando aparecen errores, los grupos deben analizar si el problema está en la programación, en el diseño GRAFCET o en una interpretación insuficiente de los requisitos. Durante esta fase, el agente conversacional sigue disponible para que el alumnado pueda plantear nuevas preguntas si necesita aclarar algún aspecto del funcionamiento del sistema.

Finalmente, se evalúa tanto el resultado técnico de la práctica como el proceso seguido por cada grupo. Una de las principales novedades de la metodología es que no solo se valora la solución final, sino también la interacción con el agente conversacional. En este sentido, se tiene en cuenta si el grupo ha sabido formular preguntas útiles, aclarar el funcionamiento del sistema e incorporar esa información al diseño del automatismo. La evaluación se completa con la revisión del GRAFCET, del programa desarrollado y de la documentación colaborativa elaborada durante la actividad. En la adaptación a secundaria se mantiene la misma lógica general, aunque con una menor complejidad técnica y con ejercicios de programación ajustados al nivel del alumnado.

2.4 Análisis de datos

El análisis de datos realizado tiene un alcance descriptivo y exploratorio. Para examinar la evolución del rendimiento académico entre los cursos considerados, se comparan la calificación media global de la asignatura y las distribuciones porcentuales del alumnado por intervalos de calificación, tanto en la asignatura como en el módulo de diseño. Las observaciones realizadas en el laboratorio y la revisión de los trabajos se utilizan como información complementaria para contextualizar las tendencias observadas. La comparación entre cursos permite describir la evolución del rendimiento, pero no aislar el efecto de cada componente de la metodología ni atribuir causalmente las diferencias a la incorporación del agente conversacional.

3. RESULTADOS

3.1 Evolución del rendimiento académico

Los resultados permiten analizar la evolución del rendimiento académico tras la incorporación progresiva de la metodología colaborativa. Además de la nota media global, se considera la distribución de calificaciones y, de forma específica, el módulo de diseño, por ser la parte de la asignatura más vinculada al análisis del problema, la definición de requisitos y la elaboración del GRAFCET.

La figura 3 muestra una evolución positiva de las calificaciones medias durante los cursos analizados. La nota media pasa de 6,08 en 2022-23 a 7,1 en 2023-24, 7,7 en 2024-25 y 7,9 en 2025-26. En conjunto, se observa una evolución progresivamente favorable del rendimiento académico durante el periodo analizado.

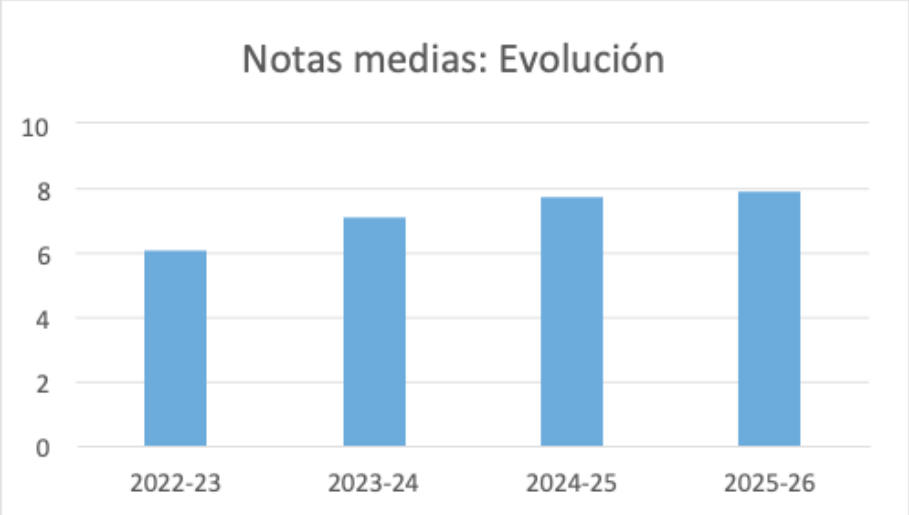


Figura 3. Evolución de la nota media en la asignatura durante los cursos analizados.
Fuente: elaboración propia a partir de las calificaciones de la asignatura.

La figura 4 completa este análisis mediante la distribución de calificaciones por rangos. En comparación con el curso inicial, se aprecia una reducción del porcentaje de estudiantes con calificaciones inferiores a 5, que pasa del 26,3 % del alumnado en 2022-23 al 11,8 % en 2025-26. Al mismo tiempo, aumenta la proporción de estudiantes situados en los rangos superiores, especialmente en las calificaciones iguales o superiores a 9. Por tanto, la mejora no se limita a un incremento de la media, sino que refleja un desplazamiento general del grupo hacia niveles de rendimiento más altos.

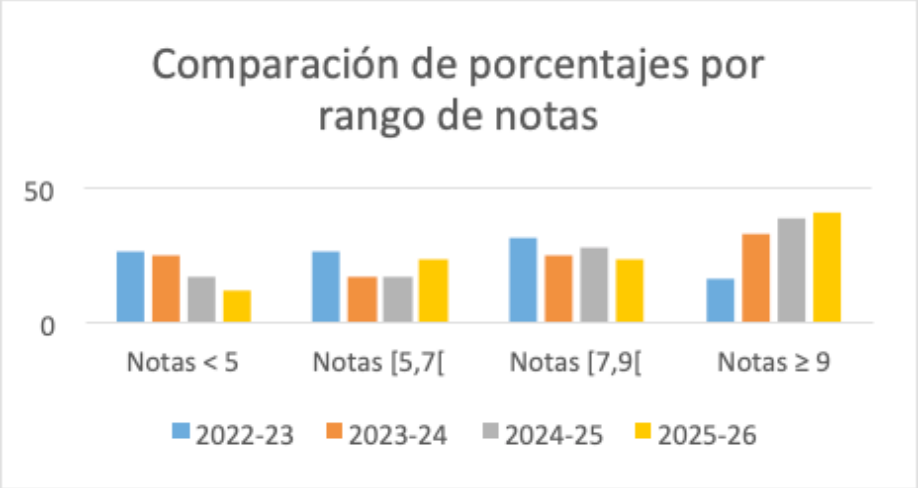


Figura 4. Distribución porcentual del alumnado por rangos de calificación en la asignatura.
Fuente: elaboración propia a partir de las calificaciones de la asignatura.

La figura 5 recoge los resultados del módulo de diseño, que es la parte de la asignatura más directamente relacionada con la fase previa a la programación. En este módulo, el alumnado debe comprender el problema, concretar los requisitos y representar el funcionamiento del sistema mediante GRAFCET antes de desarrollar el programa. Los datos muestran una mejora más clara que en el rendimiento global: disminuye de forma notable el porcentaje de estudiantes con calificaciones inferiores a 5 y aumenta el alumnado situado en los rangos altos. Este resultado sugiere que la metodología colaborativa se asocia con una evolución especialmente favorable en la fase de análisis y diseño, donde suelen aparecer muchas de las dificultades de formalización.

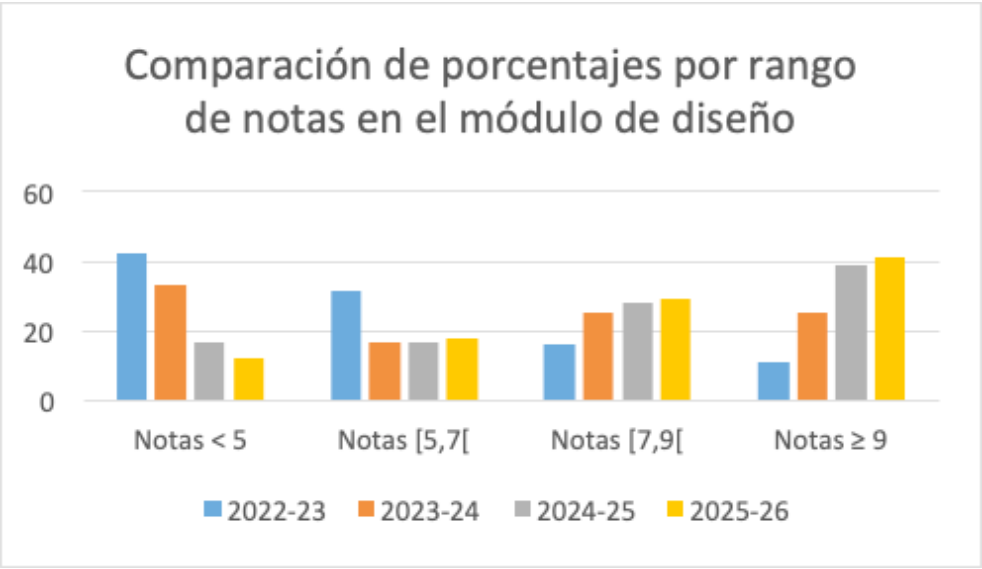


Figura 5. Distribución porcentual del alumnado por rangos de calificación en el módulo de diseño.
Fuente: elaboración propia a partir de las calificaciones de la asignatura.

3.2 Análisis de resultados y observaciones cualitativas

Consideradas conjuntamente, las figuras 3, 4 y 5 sugieren una evolución favorable del rendimiento académico respecto al curso inicial. La distribución por intervalos complementa la información proporcionada por la media al mostrar cómo cambia la proporción de estudiantes con calificaciones bajas y altas. Por su parte, los resultados del módulo de diseño permiten examinar específicamente la fase de comprensión y formalización del problema, que constituye uno de los objetivos de la experiencia.

La observación en el laboratorio y la revisión de los trabajos entregados coinciden con los resultados obtenidos en las calificaciones. En los grupos con mejores resultados se observa que el GRAFCET y el programa final reflejan mejor el funcionamiento que el grupo había definido al analizar el problema. Estas observaciones sugieren que el trabajo colaborativo ayuda a que el alumnado no vea la práctica como una simple tarea de programación, sino como un proceso completo en el que primero se entiende el problema, después se diseña la solución y finalmente se comprueba su funcionamiento.

Las experiencias en educación secundaria muestran una tendencia parecida. Aunque las herramientas y la dificultad técnica son diferentes, el alumnado que trabaja de forma colaborativa obtiene mejores resultados que el que trabaja individualmente, especialmente cuando la tarea exige pensar el problema y construir la solución paso a paso. En conjunto, los resultados respaldan la utilidad de la programación colaborativa para favorecer un aprendizaje significativo. La inteligencia artificial conversacional aporta un apoyo adicional en la aclaración de requisitos, pero el eje de la propuesta sigue siendo el trabajo en grupo aplicado a la resolución de problemas de programación y automatización.

4. DISCUSIÓN Y CONCLUSIONES

Los resultados obtenidos permiten valorar positivamente la metodología propuesta en relación con los objetivos planteados. La mejora observada en el rendimiento académico, especialmente en las tareas de diseño, sugiere que la programación colaborativa ayuda al alumnado a entender que programar no consiste solo en escribir código. Antes de programar es necesario comprender el problema, aclarar los requisitos, representar el funcionamiento del sistema y comprobar después que la solución responde a lo previsto. Esta forma de trabajar favorece una comprensión más profunda de la relación entre requisitos, diseño e implementación. Este resultado es coherente con estudios que señalan que muchas dificultades en el aprendizaje de la programación no se deben solo al dominio de la sintaxis, sino también a la capacidad de comprender el problema, organizarlo de forma adecuada y trasladarlo a una solución programada (Chuang & Chang, 2024; Gao et al., 2025; Huang et al., 2024; Qian & Choi, 2023).

La metodología también refuerza el trabajo colaborativo y la discusión técnica. Las calificaciones muestran una evolución favorable y la observación en el aula indica una mayor implicación del alumnado, especialmente cuando debe analizar el problema antes de programar. En lugar de resolver los ejercicios por ensayo y error, los estudiantes tienen que explicar sus ideas al grupo, contrastarlas con las de sus compañeros y revisar la solución a medida que avanzan. Este proceso les ayuda a identificar mejor sus propias dificultades y a mejorar progresivamente la calidad de sus respuestas, algo especialmente relevante en el aprendizaje de conceptos de programación (Florou et al., 2025). Estos resultados se alinean con investigaciones que destacan el valor del aprendizaje colaborativo en educación técnica y superior, así como los efectos positivos de la programación colaborativa sobre el rendimiento, la motivación y la construcción compartida de soluciones (Alacapınar & Uysal, 2020; Krismadinata et al., 2023; Lai & Wong, 2022; Mercier et al., 2023; Silva et al., 2020; Sun & Xu, 2025; Xu & Correia, 2024).

La incorporación del agente conversacional añade un apoyo específico a la metodología, pero el eje principal sigue siendo el trabajo colaborativo. Su función es recrear una situación parecida a la que puede darse en la práctica profesional, cuando un cliente o responsable de planta explica lo que necesita con un lenguaje no técnico. A partir de esa interacción, el alumnado debe aclarar el funcionamiento del sistema y tomar decisiones de diseño de forma consensuada. Así, la inteligencia artificial no se utiliza para obtener una solución automática, sino para generar una situación de aprendizaje más realista antes de programar.

Los resultados también sugieren que este enfoque puede adaptarse a otros niveles educativos. Las experiencias complementarias en educación secundaria muestran que la colaboración resulta útil cuando la programación se plantea como una herramienta para resolver problemas, no solo como aprendizaje de un lenguaje. Para ello, no es necesario reproducir exactamente la experiencia universitaria, sino ajustar la dificultad técnica y mantener la misma idea de fondo: que el alumnado comprenda primero el problema, lo discuta con sus compañeros y construya progresivamente una solución.

No obstante, el estudio presenta algunas limitaciones. La experiencia se ha desarrollado en grupos reducidos y en un contexto docente determinado, por lo que los resultados deben interpretarse con prudencia. Como futuras líneas de investigación, se propone comparar actividades colaborativas equivalentes con y sin apoyo de IA, analizar la relación entre las preguntas formuladas al agente y la calidad de los diseños, y evaluar la transferencia del aprendizaje a tareas individuales sin asistencia. Asimismo, sería conveniente ampliar la experiencia a otras asignaturas y estudiar sistemáticamente su adaptación a educación secundaria.

En conclusión, la metodología colaborativa propuesta favorece un aprendizaje más significativo en la enseñanza de la programación aplicada en grados de ingeniería no especializados en informática, y en enseñanza secundaria. Los resultados muestran que el alumnado comprende mejor los problemas cuando debe analizarlos con sus compañeros, discutir posibles soluciones y comprobar después si el programa funciona como estaba previsto. La incorporación de inteligencia artificial conversacional refuerza este proceso al situar al alumnado ante una tarea más próxima a la práctica profesional, en la que programar exige comprender el problema antes de construir la solución.

5. REFERENCIAS

Alacapinar, F. G., & Uysal, H. (2020). The effect of cooperative learning in education: A meta-analysis study. *Research on Education and Psychology (REP), 4*(1), 54-72. https://izlik.org/JA66ED39BE

Chuang, Y. T., & Chang, H. Y. (2024). Analyzing novice and competent programmers' problem-solving behaviors using an automated evaluation system. *Science of Computer Programming*, *237*, Artículo 103138. https://doi.org/10.1016/j.scico.2024.103138

Fleumer, T. (2006). *PLC universal system simulator: Supplementary set of templates CLC 34* (2.ª ed., SH5009-6C). Lucas-Nülle.

Florou, C., Stamoulis, G., Xenakis, A., & Plageras, A. (2025). The role of educators in facilitating students' self-assessment in learning computer programming concepts: Addressing students' challenges and enhancing learning. *Education and Information Technologies, 30*(7), 8567–8590. https://doi.org/10.1007/s10639-024-13172-2

Gao, Z., Yan, H., Liu, J., Zhang, X., Lin, Y., Zhang, Y., Sun, X. & Feng, J. (2025). Tracing distinct learning trajectories in introductory programming course: a sequence analysis of score, engage-

ment, and code metrics for novice computer science vs. math cohorts. *International Journal of STEM Education*, *12*(1), Artículo 27. https://doi.org/10.1186/s40594-025-00546-2

Huang, Y., Schunn, C.D.,Guerra, J. & Brusilovsky, P. (2024). Why students cannot easily integrate component skills: An investigation of the composition effect in programming. *ACM Transactions on Computing Education*, *24*(3), Artículo 38. https://doi.org/10.1145/3673239

Krismadinata, E., Chérifa Boudia, C., J. J., & Saputra, A. Y. (2023). Effect of collaborative programming on students achievement learning object-oriented programming course. *International Journal of Information and Education Technology, 13*(5), 792-800. 10.18178/ijiet.2023.13.5.1869

Lai, X. & Wong, G. K. (2022). Collaborative versus individual problem solving in computational thinking through programming: A meta-analysis. *British Journal of Educational Technology*, *53*(1), 150–170. https://doi.org/10.1111/bjet.13157

Mercier, E., Goldstein, M.H., Baligar, P., & Rajarathinam, R. J. (2023). Collaborative learning in engineering education. In A. Johri (Ed.), *International Handbook of Engineering Education Research* (pp. 402-432). Routledge. https://doi.org/10.4324/9781003287483-23

Qian, Y. & Choi, I. (2023). Tracing the essence: Ways to develop abstraction in computational thinking. *Educational Technology Research and Development*, *71*(3), 1055–1078. https://doi.org/10.1007/s11423-022-10182-0

Ralston, P. A. S., Tretter, T. R., & Kendall-Brown, M. (2017). Implementing collaborative learning across the engineering curriculum. *Journal of the Scholarship of Teaching and Learning, 17*(3), 89-108. https://doi.org/10.14434/v17i3.21323

Silva, L., Mendes, A. J., & Gomes, A. (2020). Computer-supported collaborative learning in programming education: A systematic literature review. In *2020 IEEE Global Engineering Education Conference (EDUCON)* (pp. 1086-1095). IEEE. https://doi.org/10.1109/EDUCON45650.2020.9125237

Sun, D., & Xu, F. (2025). Real-time collaborative programming in undergraduate education: A comprehensive empirical analysis of its impact on knowledge, behaviors, and attitudes. *Journal of Educational Computing Research, 63*(1), 33–63. https://doi.org/10.1177/07356331241295739

Xu, F., & Correia, A. P. (2024). Adopting distributed pair programming as an effective team learning activity: A systematic review. *Journal of Computing in Higher Education, 36*(2)*,* 320–349. https://doi.org/10.1007/s12528-023-09356-3

15. AI-Assisted Active Learning and Scientific Competence in Sport Science Education

Basilio Pueo, Alfonso Penichet-Tomas, Lamberto Villalon-Gasch,
Jose Manuel Jimenez-Olmedo.

Departamento Didáctica General y Didácticas Específicas, Facultad de Educación,
Universidad de Alicante

ABSTRACT

The integration of artificial intelligence (AI) in higher education raises critical questions about its impact on deep learning and autonomous competence. While AI can accelerate short-term performance, its effect on long-term knowledge retention remains poorly understood. This concern is particularly relevant in fields like Sport Sciences, where students must develop transferable skills for unaided professional practice. This study examined the effect of AI-assisted versus non-AI-assisted active learning on the scientific competence of undergraduate students (n = 50). In a six-week quasi-experimental design, an AI group (n = 26) used AI tools during practice sessions, while a non-AI group (n = 24) used conventional resources. Both groups followed the same curriculum, and all assessments (pre-test, post-test, retest) were completed without AI access. The groups were statistically equivalent at baseline (p = .243). A 2 × 3 mixed ANOVA revealed a significant Group × Time interaction (F(2, 96) = 4.63, p = .012, η^2_p = .088). Post-hoc analyses showed that at post-test, the AI group significantly outperformed the non-AI group (p < .001, Cohen's d = 1.26). At the delayed retest, however, this advantage disappeared, with the difference no longer being significant (p = .325). These findings suggest that while AI assistance accelerates initial performance, it may not confer a lasting advantage for unaided, long-term competence.

KEYWORDS: artificial intelligence; scientific competence; active methodologies; desirable difficulties; sport sciences.

1. INTRODUCTION

Generative artificial intelligence (AI) has sparked a critical debate in higher education about the nature of student learning (Johnson et al., 2024a). While discussions about new technologies are not new, AI presents a unique challenge because it does not just retrieve information but generates novel, reasoned content (Capinding & Dumayas, 2024). This capability forces the academic community to confront a complex landscape of opportunities, ethical concerns, and student anxieties (Aljuaid, 2024; Iranzo et al., 2024; Saúde et al., 2024).

This issue is especially relevant in fields like Physical Activity and Sport Sciences (PASS), where scientific competence is a core professional skill (Potop et al., 2025). Students in this area have long struggled with research methods and statistics, disciplines often marked by low self-efficacy and high anxiety (Lane et al., 2004; Von Roten & De Roten, 2023). Mastering these skills is essential for evidence-based practice, yet the gap between classroom knowledge and its real-world application remains a persistent problem that can limit students' critical thinking abilities (Turkay, 2021).

In response, modern health and sport science curricula have widely adopted active learning methodologies (Herrera et al., 2020; Mitre et al., 2008). Approaches like problem-based and case-based learning are proven to build complex skills by immersing students in authentic tasks (Gontijo & Finkler, 2025; Simon et al., 2014; Veloso et al., 2019). It is into this active learning environment that AI is now being introduced.

Proponents view AI as a powerful scaffolding tool. It can function as an on-demand virtual tutor, offering immediate feedback and automating routine procedures (Hemachandran et al., 2022; Lin et al., 2023). This process can reduce extraneous cognitive load, freeing up students' mental bandwidth for higher-order thinking (Sevilla et al., 2025). This potential for cognitive augmentation is the strongest case for AI's pedagogical use.

The risk, however, is that students may use AI not to augment their thinking but to replace it entirely, a phenomenon known as cognitive offloading (Risko & Gilbert, 2016). By outsourcing mental effort, students may engage in shallower processing, leading to a form of technological dependency that hinders the development of autonomous skills (García, 2025; Jose et al., 2025). This can create a "metacognitive laziness" where the student loses a sense of ownership over their work (Batista-Toledo & Gavilan, 2026; Estrada-Araoz, 2026), a dynamic previously observed with the "Google effect" on memory (Sparrow et al., 2011).

This tension between short-term efficiency and long-term learning is explained well by the "desirable difficulties" framework (Bjork, 1994). The theory predicts that conditions requiring greater cognitive effort during practice, while often hindering immediate performance, lead to more robust, long-term knowledge retention. Conversely, conditions that make performance easy may result in more fragile learning (Soderstrom & Bjork, 2015).

This study, therefore, examines the effects of AI-assisted versus non-assisted active learning on the development of scientific competence. Drawing from the desirable difficulties framework, we hypothesized that: (1) students using AI would perform significantly better on an immediate post-test, reflecting an accelerated but more superficial learning process; and (2) this advantage would disappear or even reverse on a delayed retest, where the non-AI group, having engaged in more effortful practice, would show superior long-term retention.

2. METHOD

2.1. Participants and Context

We used a quasi-experimental, parallel-group design conducted over six weeks in a required research methods course for sport science undergraduates. The design was quasi-experimental as it involved pre-existing, intact classroom groups. The sample consisted of 50 students (29 males, 21 females) who were assigned to either an AI-assisted group ($n = 26$) or a non-AI-assisted control group ($n = 24$). The study received ethical approval from the institutional review board, and all participants provided informed consent.

2.2. Instruments

The primary dependent variable was scientific competence, which we measured through student performance on three parallel assessments scored on a 0–10 scale. These tests required students to apply statistical concepts to solve problems relevant to sport sciences. The pre-test was administered after

an introductory module on descriptive statistics. The post-test was given at the end of the six-week intervention, covering inferential statistics. The retest coincided with the official end-of-course examination, administered three weeks after the post-test, and covered the full syllabus from descriptive statistics to multiple regression.

2.3. Procedure

Both groups attended the same theoretical lectures and worked on identical active learning tasks (Ballen et al., 2017). The intervention was confined to the practical, problem-solving sessions. The AI group was trained and encouraged to use generative AI tools (e.g., ChatGPT) for assistance. Students in the AI group attended a brief orientation session before the intervention began. During this session, they were introduced to basic strategies for formulating effective prompts, critically evaluating AI-generated responses, and incorporating those responses into their own reasoning rather than accepting them uncritically. After the initial training, no restrictions were placed on either the type or the frequency of AI queries, allowing participants to use the system freely and reflecting a more natural pattern of unmonitored AI use. The non-AI group performed the same tasks using standard digital tools (e.g., Jamovi software, search engines) without AI assistance. Crucially, all assessments were conducted in a controlled environment without AI access for either group, ensuring that the tests measured unaided, transferable competence.

2.4. Statistical Analysis

A 2 (Group: AI, non-AI) × 3 (Time: Pre-test, Post-test, Retest) mixed-design Analysis of Variance (ANOVA) was conducted to examine the effects of the intervention on scientific competence scores. The within-subjects factor was Time, and the between-subjects factor was Group. Assumptions of sphericity and homogeneity of variances were checked using Mauchly's test and Levene's test, respectively. Post-hoc pairwise comparisons with Bonferroni correction were used to decompose any significant interaction effects. An independent samples t-test was used to calculate the effect size (Cohen's *d*) for the key comparison at post-test. All statistical analyses were performed using R (Version 4.3.1). The alpha level for all statistical tests was set at .05.

3. RESULTS

Descriptive statistics for the scientific competence scores are presented in Table 1. The data met all necessary assumptions for the planned analysis. Mauchly's test indicated that the assumption of sphericity was not violated, W = 0.929, $p = .177$. Levene's test confirmed homogeneity of variances at all three time points (all p values > .25).

Table 1. Descriptive Statistics of Scientific Competence Scores by Group and Time.

Group	n	Pre-test M (SD)	Post-test M (SD)	Retest M (SD)
AI Group	26	5.12 (1.36)	7.52 (0.88)	7.00 (1.15)
Non-AI Group	24	4.66 (1.40)	6.41 (0.89)	7.29 (0.83)

The mixed ANOVA revealed a large and significant main effect of Time, $F(2, 96) = 59.94$, $p < .001$, $\eta^2_p = 0.555$, showing that scores changed significantly across the measurements. A significant main effect of Group was also found, $F(1, 48) = 6.41$, $p = .015$, $\eta^2_p = 0.118$. Most importantly, the analysis yielded the predicted significant Time × Group interaction, $F(2, 96) = 4.63$, $p = .012$, $\eta^2_p = 0.088$. This interaction demonstrates that the learning trajectories of the two groups differed significantly over time, a pattern illustrated in Figure 1.

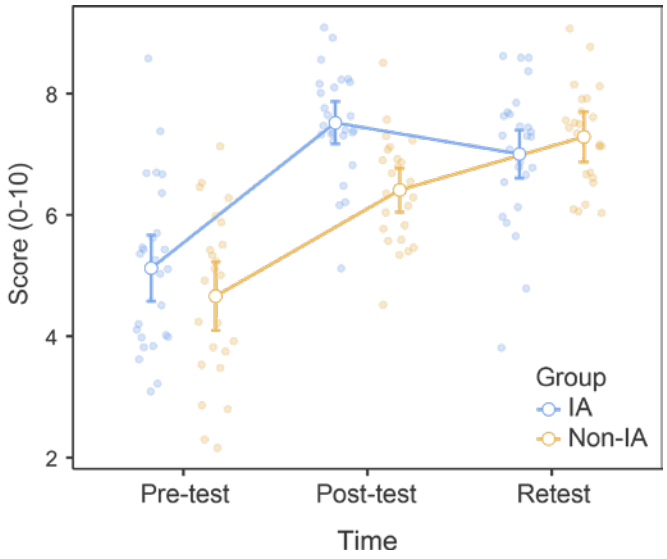


Figure 1. Estimated marginal means of scientific competence scores across the three time points (Pre, Post, Retest) for the AI-assisted and non-AI-assisted groups. Error bars represent 95% confidence intervals.

Post-hoc pairwise comparisons were conducted to interpret the interaction. At Pre-test, the AI and non-AI groups were equivalent ($p = .243$), establishing a common baseline. At Post-test, the AI group significantly outperformed the non-AI group ($p < .001$), with the difference representing a large effect size (Cohen's $d = 1.26$), which supports our first hypothesis. At the delayed Retest, however, this advantage disappeared, and no significant difference was found between the groups ($p = .325$), supporting our second hypothesis.

4. DISCUSSION AND CONCLUSIONS

The results of this study confirm our primary hypotheses and align with the theoretical frameworks of cognitive offloading (Risko & Gilbert, 2016) and desirable difficulties (Bjork, 1994). The significant Group × Time interaction ($F(2, 96) = 4.63$, $p = .012$) reveals a critical pattern: AI assistance accelerates short-term performance but fails to confer a lasting advantage in long-term, unaided competence. This finding resonates with recent studies describing the effect of generative AI as providing "short-term gains" that result in "long-term gaps" (Akgun & Toker, 2025) and acting as a "cognitive crutch" that, while immediately helpful, may not lead to durable knowledge retention (Barcaui, 2025; Kalam et al., 2025).

The AI group's superior performance at post-test is consistent with the scaffolding function that AI tools can provide: immediate feedback, procedural support, and reduced extraneous load during complex tasks (Kasneci et al., 2023). However, this raises a distinction that is easy to overlook in

practice, the difference between assisted performance and autonomous competence. Students may have learned to operate the tool effectively without fully internalizing the underlying reasoning. This concern is not merely theoretical; Bastani et al. (2025) demonstrated experimentally that unrestricted AI use in educational settings can reduce student persistence and impair independent performance, suggesting that ease of access to answers may come at a cognitive cost.

The erosion of the AI group's advantage at retest is, in our view, the most theoretically significant finding of this study. The non-AI group, denied the shortcut, had to wrestle with problems, tolerate uncertainty, and construct solutions independently. This is precisely the kind of effortful engagement that the desirable difficulties framework predicts will produce more durable learning (Soderstrom & Bjork, 2015). The implication is uncomfortable but important: the cognitive effort that AI saves during practice may be the very effort that consolidates knowledge. As De Bruin (2023) has argued in medical education, helping students accept rather than avoid difficulty is one of the most undervalued pedagogical challenges of our time.

Our findings contribute to a growing body of literature examining the impact of generative AI on knowledge worker productivity and academic performance. While studies often highlight significant productivity gains in professional tasks when using AI (Dell'Acqua et al., 2026; Noy & Zhang, 2023), our results add a crucial nuance: performance gains measured during assisted practice may not translate to improved unaided performance in subsequent assessments. This is consistent with findings by Wecks et al. (2024), who also found that generative AI usage did not necessarily lead to better exam performance, suggesting a disconnect between AI-assisted task completion and genuine knowledge consolidation.

The quantitative results gain texture when set alongside what was observed directly in the classroom. Students in the AI group repeatedly turned to the tool for queries that were already addressed in the practical guide, questions about notation, about which test to run, about how to interpret an output they had not yet tried to read themselves. The AI answered before the student had fully formed the question, and in doing so, may have preempted the productive struggle that learning requires. This pattern aligns with what Galindo-Domínguez et al. (2026) describe as a paradox of overreliance: the more capable the tool, the less students feel compelled to regulate their own learning process. The AI, in effect, was solving problems the students needed to solve themselves.

These findings do not advocate for the exclusion of AI from education but rather for its thoughtful and strategic integration. The results suggest that the optimal approach may be a scaffolded model where AI is used to reduce barriers in the initial phases of learning but is progressively withdrawn to foster autonomous competence (Kasneci et al., 2023; Sun & Wu, 2026; Wang et al., 2025).

Finally, several limitations of this study must be acknowledged. First, the quasi-experimental design, while ecologically valid, does not permit the same degree of causal inference as a randomized controlled trial. Second, the sample size is modest and drawn from a single academic discipline at one university, which may limit the generalizability of the findings. Third, the assessments, while designed to be parallel, inherently increased in cumulative content knowledge, which could be a confounding factor in the retest scores. Fourth, baseline differences in statistical knowledge and digital competence were not assessed directly. Although the two groups performed similarly on the pre-test, unmeasured variation in familiarity with statistical software or AI tools may still have influenced how effectively participants engaged with the learning activities. As a result, the possibility of residual confounding cannot be completely excluded. Future research should include baseline assessments of these variables to better account for their potential influence. Future research should replicate

these findings with larger, more diverse samples and employ randomized designs to strengthen causal claims.

This study offers empirical evidence for a pattern that many instructors may already sense intuitively: AI assistance can accelerate performance without deepening learning. The Group × Time interaction observed here is not simply a statistical result, it reflects a genuine tension at the heart of AI integration in education. When students use AI to bypass the difficult parts of learning, they may be trading long-term competence for short-term fluency.

The pedagogical implication is not to prohibit AI but to be deliberate about when and how it is available. A scaffolded model, where AI support is present in early phases and progressively withdrawn as competence develops, may allow educators to capture the motivational and efficiency benefits of these tools without undermining the effortful practice that durable learning requires. In our specific context, a sport science research methods course, this might mean permitting AI assistance in early descriptive tasks while requiring fully independent work as students move into inferential reasoning.

What remains open is how students themselves experience this transition. Whether they perceive the withdrawal of AI as a productive challenge or as an arbitrary restriction is a question that future research should take seriously. The answer may matter as much as the pedagogical design itself.

5. REFERENCES

Akgun, M., & Toker, S. (2025). Short-Term Gains, Long-Term Gaps: The Impact of GenAI and Search Technologies on Retention. *ArXiv.org*. https://doi.org/10.48550/arxiv.2507.07357

Aljuaid, H. (2024). The Impact of Artificial Intelligence Tools on Academic Writing Instruction in Higher Education: A Systematic Review. *Arab World English Journal, 1*(1), 26-55. https://doi.org/10.24093/awej/chatgpt.2

Ballen, C. J., Wieman, C., Salehi, S., Searle, J. B., & Zamudio, K. R. (2017). Enhancing diversity in undergraduate science: Self-efficacy drives performance gains with active learning. *CBE—Life Sciences Education, 16*(4), ar56. https://doi.org/10.1187/cbe.17-02-0031

Barcaui, A. (2025). ChatGPT as a cognitive crutch: Evidence from a randomized controlled trial on knowledge retention. *Social Sciences & Humanities Open, 12*, 102287. https://doi.org/10.1016/j.ssaho.2025.102287

Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., & Mariman, R. (2025). Generative AI without guardrails can harm learning: Evidence from high school mathematics. *Proceedings of the National Academy of Sciences, 122*(26), e2422633122. https://doi.org/10.1073/pnas.2422633122

Batista-Toledo, S., & Gavilan, D. (2026). ¿Lo hice yo o lo hizo la inteligencia artificial?: impacto de la inteligencia artificial en la propiedad psicológica y el síndrome «impostor» de los estudiantes en la educación superior. *Tecnología, Ciencia y Educación, 33*, 7-33. https://doi.org/10.51302/tce.2026.24625

Bjork, R. A. (1994). Memory and Metamemory Considerations in the Training of Human Beings. In J. Metcalfe & A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185–206). The MIT Press.

Capinding, A. T., & Dumayas, F. T. (2024). Transformative pedagogy in the digital age: Unraveling the impact of artificial intelligence on higher education students. *Problems Of Education In The 21st Century, 82*(5), 630-657. https://doi.org/10.33225/pec/24.82.630

De Bruin, A. B. H. (2023). Dealing with Desirable Difficulties: Supporting Students to Accept, Reduce, or Silence Effort. *Medical Science Educator, 33*(Suppl 1), 5-9. https://doi.org/10.1007/s40670-023-01911-y

Dell'Acqua, F., McFowland, E., Mollick, E., Lifshitz, H., Kellogg, K. C., Rajendran, S., Krayer, L., Candelon, F., & Lakhani, K. R. (2026). Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of Artificial Intelligence on Knowledge Worker Productivity and Quality. *Organization Science, 37*(2), 403-423. https://doi.org/10.1287/orsc.2025.21838

Estrada-Araoz, E. G. (2026). Pereza metacognitiva y control cognitivo en la era de la inteligencia artificial. *Revista de Neuro-Psiquiatría, 89*(1), 102-104. https://doi.org/10.20453/rnp.v89i1.7535

Galindo-Domínguez, H., Delgado, N., Sainz-De-La-Maza, M., & Etxabe, J. (2026). Self-regulation and overreliance on artificial intelligence: Unpacking a paradox through a mixed-methods study in higher education. *Computers In Human Behavior, 181*, 108985. https://doi.org/10.1016/j.chb.2026.108985

García, A. E. (2025). Generative artificial intelligence: between enhancement and cognitive offloading. *Cultura Ciencia y Deporte, 20*(66). https://doi.org/10.12800/ccd.v20i66.2698

Gontijo, L. P. T., & Finkler, M. (2025). Preparing health educators: the role of active teaching-learning methodologies and interprofessional education. *Revista Ibero-Americana de Estudos Em Educação, 20*, e20284. https://doi.org/10.21723/riaee.v20i00.2028402

Hemachandran, K., Verma, P., Pareek, P., Arora, N., Kumar, K. V. R., Ahanger, T. A., Pise, A. A., & Ratna, R. (2022). Artificial Intelligence: A universal virtual tool to augment tutoring in higher education. *Computational Intelligence and Neuroscience, 2022*, 1-8. https://doi.org/10.1155/2022/1410448

Herrera, Y. R., Sánchez, M. L. Z., Valdivia-Moral, P., Marín-Marín, J., & García, S. A. (2020). Active Methodologies in the Training of Future Health Professionals: Academic Goals and Autonomous Learning Strategies. *Sustainability, 12*(4), 1485. https://doi.org/10.3390/su12041485

Iranzo, R. G., Gutiérrez-Ujaque, D., & Cairol, M. T. (2024). De la ansiedad al empoderamiento: Impacto del uso de la inteligencia artificial en la percepción de los estudiantes en educación superior. *REDU Revista de Docencia Universitaria, 22*(2), 85-104. https://doi.org/10.4995/redu.2024.22009

Johnson, N., Seaman, J., & Seaman, J. (2024a). The anticipated impact of artificial intelligence on higher education. *Online Learning, 28*(3). https://doi.org/10.24059/olj.v28i3.4646

Jose, B., Cherian, J., Verghis, A. M., Varghise, S. M., S, M., & Joseph, S. (2025). The cognitive paradox of AI in education: Between enhancement and erosion. *Frontiers In Psychology, 16*, 1550621. https://doi.org/10.3389/fpsyg.2025.1550621

Kalam, K. A., Masoud, F. D., Muntaser, A., Ranga, R., Geng, X., & Goyal, M. (2025). ChatGPT as a learning tool for medical students: Results from a randomized controlled trial. *Cureus, 17*(6), e85767. https://doi.org/10.7759/cureus.85767

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., & Kasneci, 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

Lane, A. M., Hall, R., & Lane, J. (2004). Self-efficacy and statistics performance among Sport Studies students. *Teaching In Higher Education, 9*(4), 435-448. https://doi.org/10.1080/1356251042000252372

Lin, C., Huang, A. y. Q., & Lu, O. H. T. (2023). Artificial intelligence in intelligent tutoring systems toward sustainable education: A systematic review. *Smart Learning Environments, 10*(1). https://doi.org/10.1186/s40561-023-00260-y

Mitre, S. M., Siqueira-Batista, R., Girardi-De-Mendonça, J. M., De Morais-Pinto, N. M., De Almeida Brandão Meirelles, C., Pinto-Porto, C., Moreira, T., & Hoffmann, L. M. A. (2008). Metodologias ativas de ensino-aprendizagem na formação profissional em saúde: debates atuais. *Ciência & Saúde Coletiva, 13*(suppl 2), 2133-2144. https://doi.org/10.1590/s1413-81232008000900018

Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science, 381*(6654), 187-192. https://doi.org/10.1126/science.adh2586

Potop, V., Mihailescu, L. E., Mihai, I., Ciorbă, C., Chernozub, A., Cieslicka, M., Jagiello, W., Matac, L. M., & Pricop, A. D. (2025). Validation of research activity in the field of sports science and physical education for students in higher education. *Physical Education Of Students, 29*(1), 85-94. https://doi.org/10.15561/20755279.2025.0109

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences, 20*(9), 676–688. https://doi.org/10.1016/j.tics.2016.07.002

Saúde, S., Barros, J. P., & Almeida, I. (2024). Impacts of Generative Artificial Intelligence in Higher Education: Research Trends and Students' Perceptions. *Social Sciences, 13*(8), 410. https://doi.org/10.3390/socsci13080410

Sevilla, A., Cuevas-Ruiz, P., Rello, L., & Sanz, I. (2025). Artificial Intelligence in Education: Computer-Assisted Learning and AI-guided Tutors. *Italian Economic Journal, 12*(1), 1-38. https://doi.org/10.1007/s40797-025-00354-1

Simon, E., Jezine, E., Vasconcelos, E. M., & Ribeiro, K. S. Q. S. (2014). Metodologias ativas de ensino-aprendizagem e educação popular: encontros e desencontros no contexto da formação dos profissionais de saúde. *Interface - Comunicação Saúde Educação, 18*(suppl 2), 1355-1364. https://doi.org/10.1590/1807-57622013.0477

Soderstrom, N. C., & Bjork, R. A. (2015). Learning versus performance: An integrative review. *Perspectives on Psychological Science, 10*(2), 176–199. https://doi.org/10.1177/1745691615569000

Sparrow, B., Liu, J., & Wegner, D. M. (2011). Google Effects on Memory: Cognitive Consequences of Having Information at Our Fingertips. *Science, 333*(6043), 776-778. https://doi.org/10.1126/science.1207745

Sun, Y., & Wu, Y. (2026). The efficacy of artificial intelligence - powered scaffolding in individual acquisition efficiency of EFL in tertiary educational context. *Frontiers In Psychology, 16*, 1613285. https://doi.org/10.3389/fpsyg.2025.1613285

Turkay, H. (2021). The impact of the sports science students' attitudes towards the scientific research methods course on their critical thinking skills. *Pakistan Journal of Medical & Health Sciences, 15*(10), 3004-3008. https://doi.org/10.53350/pjmhs2115103004

Veloso, M. D., Pequeno, A. M. C., & Da Silva Negreiros, F. D. (2019). Metodologias ativas de aprendizagem no ensino superior de saúde: o fazer pedagógico. *Revista Ibero-Americana de Estudos Em Educação, 14*(2), 354-370. https://doi.org/10.21723/riaee.v14i2.11860

Von Roten, F. C., & De Roten, Y. (2023). Undergraduate sport sciences students' attitudes towards statistics: A gender perspective. *Journal Of Hospitality Leisure Sport & Tourism Education, 33*, 100452. https://doi.org/10.1016/j.jhlste.2023.100452

Wang, F., Zhou, X., Li, K., Cheung, A. C. K., & Tian, M. (2025). The effects of artificial intelligence-based interactive scaffolding on secondary students' speaking performance, goal setting, self-evaluation, and motivation in informal digital learning of English. *Interactive Learning Environments, 33*(7), 4633-4652. https://doi.org/10.1080/10494820.2025.2470319

Wecks, J. O., Voshaar, J., Plate, B. J., & Zimmermann, J. (2024). Generative AI Usage and Exam Performance. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2404.19699

16. Evaluación continua con validación formativa en Python: sostenibilidad, escalabilidad e integridad académica en la enseñanza del cálculo numérico en el Grado en Matemáticas [1]

J. A. Vargas Alemañy, L. Moreno Martínez, J.M. Sayol España, I. Vigo Aguiar, D. García García

Departamento de Matemática Aplicada e Ingeniería Aeroespacial, Escuela Politécnica Superior, Universidad de Alicante.

RESUMEN

El presente trabajo analiza la viabilidad, escalabilidad y capacidad discriminatoria de un modelo de evaluación continua basado en la validación formativa mediante Python en la asignatura Cálculo Numérico II del Grado en Matemáticas de la Universidad de Alicante. A lo largo de tres cursos académicos (2023-24 a 2025-26), se estudió la transición desde una corrección tradicional manual hacia un entorno de andamiaje cognitivo basado en Jupyter Notebooks que provee al alumnado de ejemplos numéricos intermedios para el autodiagnóstico de sus scripts. Los resultados apoyan la viabilidad y sostenibilidad del modelo al duplicar el volumen de prácticas evaluadas en la fase de consolidación, logrando mantener la excelencia académica general con una mediana en las calificaciones de 9,45. Asimismo, el incremento de entregas se asoció con una reducción del efecto techo observado en la fase piloto y aumentó la dispersión estadística de las notas en prácticas complejas (alcanzando una desviación típica de 1,85), favoreciendo una mayor diferenciación entre niveles de desempeño del alumnado. Finalmente, la optimización del tiempo de corrección permitió dedicar una mayor atención a la revisión cualitativa del código, lo que favoreció la detección de posibles conductas contrarias a la integridad académica como el plagio. En conjunto, los resultados apoyan que el modelo constituye una alternativa viable, escalable y potencialmente eficaz para la enseñanza de MÉTODOs numéricos.

PALABRAS CLAVE: Evaluación Continua, Validación Formativa, Integridad Académica, Python, Cálculo Numérico.

1. INTRODUCCIÓN

En el ámbito de la educación superior en STEM, la enseñanza del cálculo numérico se enfrenta al desafío de articular una sólida base teórica con la resolución práctica de problemas complejos a través de la programación. En asignaturas masificadas con una alta demanda técnica, la implementación de un modelo de evaluación continua que ofrezca una retroalimentación ágil y constructiva resulta a menudo inviable desde el punto de vista logístico. Los docentes se ven frecuentemente obligados a adoptar una evaluación punitiva y desfasada en el tiempo, o bien a reducir el número de entregas prácticas para mitigar una carga de corrección inasumible. Esta problemática resta eficiencia al proceso evaluativo, ya que el profesorado invierte la mayor parte de su tiempo en detectar y penalizar

1 El presente trabajo ha contado con una ayuda del Programa de Redes de investigación en docencia universitaria del Instituto de Ciencias de la Educación de la Universidad de Alicante (convocatoria 2024). Ref.: 6235.

fallos triviales o errores de cálculo numérico básicos, en lugar de evaluar la comprensión algorítmica profunda y la calidad estructural del código. Ante esta situación, la estrategia de investigación de este trabajo se plantea como un estudio cuasiexperimental basado en la comparación de tres cohortes independientes de estudiantes a lo largo de tres cursos académicos consecutivos (2023-24 a 2025-26). En él se analiza el impacto pedagógico y operativo de transicionar desde un modelo de evaluación tradicional hacia una metodología apoyada por herramientas de validación y autodiagnóstico formativo en Python, evaluando su sostenibilidad al escalar el nivel de exigencia y duplicar el volumen de prácticas evaluadas.

La literatura académica reciente corrobora que la retroalimentación inmediata y el diseño de entornos de autodiagnóstico son factores críticos para el éxito del aprendizaje autónomo en disciplinas de programación y MÉTODOs numéricos (Morris et al., 2021). En el contexto de las asignaturas STEM, los sistemas de validación que ofrecen pistas y ejemplos de los resultados numéricos esperados actúan como un andamiaje cognitivo indispensable que promueve la autorregulación y la corrección interactiva por parte del alumnado (Becker et al., 2023). No obstante, la implantación de estas metodologías en la actualidad se enfrenta a un cambio de paradigma debido a la irrupción de las herramientas de Inteligencia Artificial Generativa, como los grandes modelos de lenguaje aplicados a la generación de código, lo que está obligando a redefinir el papel del profesorado y el diseño de las actividades evaluativas en las aulas de programación (Prather et al., 2023).

Estudios recientes advierten que el libre acceso a estas tecnologías puede inducir a un efecto techo —fenómeno estadístico donde las calificaciones se acumulan artificialmente en el valor máximo de la escala, perdiendo capacidad de discriminación (Staus et al., 2021)— y al camuflaje del aprendizaje real. El acceso a herramientas de IA generativa puede facilitar la obtención de resultados numéricos correctos sin que ello implique necesariamente una comprensión equivalente del razonamiento algorítmico subyacente (Finnie-Ansley et al., 2023; Prather et al., 2024). Asimismo, la literatura enfocada en la integridad académica en la era de la IA (Denny et al., 2024) señala que la masificación en las aulas incrementa exponencialmente el riesgo de plagio y de automatización acrítica de las entregas si el docente carece de la disponibilidad temporal para auditar las fuentes de código fuente. Limitarse a contrastar exclusivamente la validez del output numérico entregado resulta insuficiente, haciendo imperativo un modelo evaluativo continuo con un mayor número de hitos de control que libere tiempo docente para la fiscalización cualitativa del desarrollo lógico.

Para dar respuesta a estos retos, este trabajo se articula en torno a un objetivo principal: evaluar la viabilidad, escalabilidad y capacidad para diferenciar niveles de desempeño académico de un modelo de evaluación continua basado en la validación formativa en Python tras duplicar la carga docente evaluada, pasando de un enfoque piloto de 2 prácticas (curso 2024-25) a un diseño consolidado de 4 prácticas (curso 2025-26). Como objetivos específicos, se propone verificar si el rendimiento académico se sostiene al elevar las exigencias prácticas, analizar si el incremento de la continuidad en las entregas aumenta la dispersión de las calificaciones mejorando la detección de estudiantes rezagados, y evaluar si la reducción del tiempo invertido en corregir errores básicos permite al docente realizar una fiscalización cualitativa más rigurosa para preservar la integridad académica. Al respecto, se formulan dos hipótesis de investigación: de un lado, se espera que la provisión de ejemplos de validación intermedia en Python mitigue la carga de corrección básica del docente, haciendo sostenible la duplicación del volumen de prácticas sin mermar la excelencia académica general (mediana); y, de otro lado, que este incremento de la continuidad pedagógica

dota al sistema de una mayor diferenciación estadística entre los resultados académicos del alumnado (desviación típica) y, de forma complementaria, libera tiempo docente que permite una fiscalización cualitativa más detenida susceptible de identificar conductas contrarias a la honestidad académica, como el plagio de código.

2. MÉTODO

2.1. Descripción del contexto y de los participantes

La presente innovación docente se ha llevado a cabo en el contexto del Grado en Matemáticas de la Universidad de Alicante, concretamente en la asignatura Cálculo Numérico II (código 25040), impartida en el primer cuatrimestre del cuarto curso. Esta asignatura forma parte del módulo de Asignaturas Obligatorias y tiene una carga de 6 créditos ECTS. Su objetivo principal es que el alumnado adquiera competencias en la resolución numérica de ecuaciones diferenciales ordinarias, utilizando para ello técnicas de programación orientadas a la precisión, eficiencia y estabilidad de los algoritmos implementados. El estudio adopta un diseño cuasiexperimental con comparación de cohortes independientes de estudiantes a lo largo de tres cursos académicos consecutivos (2023-24, 2024-25 y 2025-26), con el fin de evaluar la transición y escalabilidad de un modelo de evaluación apoyado en herramientas de validación intermedia en Python. La muestra total está constituida por los estudiantes matriculados en la asignatura que completaron el itinerario de evaluación continua obligatoria en sus respectivos años, distribuidos de la siguiente manera: 33 estudiantes en el curso de partida 2023-24 (modelo tradicional sin herramientas de validación); 62 estudiantes en el curso 2024-25 (fase piloto con INTRODUCCIÓN de herramientas básicas); y 45 estudiantes en el curso 2025-26 (fase de consolidación con duplicación de la carga práctica). El notable incremento de matriculados en el curso 2024-25 frente al curso de partida resulta llamativo y excede el objeto de este estudio; de hecho, esa misma masificación obligó a habilitar un grupo adicional de prácticas ese curso. Una posible explicación, que no ha sido contrastada empíricamente en este trabajo, es el efecto de arrastre de las cohortes de la pandemia de COVID-19, cuyas modalidades de evaluación no presencial en cursos previos podrían haber facilitado una mayor progresión académica y, con ello, un mayor número de estudiantes alcanzando cuarto curso en ese año académico. Esta circunstancia debe tenerse en cuenta al interpretar las comparaciones entre cohortes. Dado que el alumnado pertenece a la misma titulación y nivel académico, se asume un perfil formativo de partida semejante en cuanto a sus prerrequisitos matemáticos y de programación en Python, si bien debe tenerse en cuenta la cautela propia de la comparación entre cohortes independientes.

Por último, cabe señalar las garantías éticas y de diseño metodológico vinculadas al tratamiento de los datos y al rol del profesorado. Dado que la investigación adopta un diseño transversal de cohortes independientes orientadas a un análisis global, las calificaciones y registros académicos extraídos fueron tratados de forma agregada y anónima, garantizando el pleno cumplimiento de la normativa sobre protección de datos personales (RGPD, UE 2016/679). Asimismo, para atenuar el posible sesgo de expectativa derivado del doble rol del equipo docente como diseñador de la innovación y evaluador de los resultados, la calificación de las entregas se basó de manera estricta en las rúbricas numéricas predefinidas y en los criterios de validación objetiva, realizándose la explotación de los datos con fines de investigación una vez cerradas las actas oficiales de cada curso

2.2. Instrumentos

El instrumento pedagógico principal consiste en un entorno de prácticas basadas en Jupyter Notebooks y scripts de Python, diseñado para la resolución computacional de problemas complejos (como interpolación avanzada, diferencias finitas y ecuaciones diferenciales), siguiendo el enfoque de aprendizaje por indagación basado en notebooks interactivos propuesto para la enseñanza de MÉTODOs numéricos (Ketcheson, 2014). La herramienta clave introducida en la fase piloto (24-25) y consolidada en el último curso (25-26) consiste en un set de datos de prueba, ejemplos numéricos e hitos de validación intermedia proporcionados a los estudiantes junto con el enunciado de cada práctica. Específicamente, estas validaciones consisten en tablas de ejecución de referencia que detallan de forma explícita el comportamiento y el *output* esperado del algoritmo —como los valores exactos de los pasos de integración (), estimaciones de error local o parámetros de control (véase la Figura 1)— ante diferentes condiciones de entrada de los programas numéricos. Este instrumento de autodiagnóstico permite que los alumnos contrasten si los resultados de sus programas coinciden con las soluciones físicas y analíticas esperadas antes de efectuar la entrega definitiva.

Para el análisis de los resultados académicos y de la integridad del proceso, se recopilaron las calificaciones numéricas individualizadas de cada hito de evaluación continua (escala de 0 a 10) obtenidas a partir de los registros de entregas oficiales de los tres cursos.

Ejemplo de comprobación dado a los alumnos.

Comprobaciones

Funciones RKF, RKV

In [146]:
```
def f(x,y):
    return -2*x*y**2+y
rng = [-1,1]
pvi=[f,(0,2)]
S=DESOLVE_RK(pvi,rng,info=3,metodo='RKF45',hini=0.5,
             hmin=10**(-5),hmax=0.5, atol=10**(-5))
```

executed in 42ms, finished 08:37:15 2025-11-27

```
|-----------------------|
|Información de los pasos|
|-----------------------------------------------------------------------
Paso          x                y          h             error        q
---      -------  ---------------   -0.50000  1.1543636713e-02  0.21248
  1     -0.10624  1.7613613254e+00   -0.10624  1.5523498795e-06  1.26356
  2     -0.24048  1.4312480299e+00   -0.13424  4.0050017903e-06  1.04538
  3     -0.38082  1.1147454254e+00   -0.14033  9.4817684821e-07  1.39449
---      -------  ---------------   -0.19570  2.2702928777e-05  0.73889
  4     -0.52542  8.4956238439e-01   -0.14460  5.0094021628e-06  0.99962
  5     -0.66996  6.4704482327e-01   -0.14454  4.9087582326e-06  1.00369
  6     -0.81503  4.9551695364e-01   -0.14508  3.5552233619e-06  1.07058
  7     -0.97035  3.7643162648e-01   -0.15532  3.1165009408e-06  1.09916
  8     -1.00000  3.5768466716e-01   -0.02965  4.8187988647e-10  -------
|-----------------------|
|Información de los pasos|
|-----------------------------------------------------------------------
Paso          x                y          h             error        q
---      -------  ---------------    0.50000  3.9072549901e-03  0.26389
  1      0.13194  2.1984576892e+00    0.13194  3.9979371208e-06  1.04575
  2      0.26992  2.2299356314e+00    0.13798  2.3080631243e-07  1.84988
---      -------  ---------------    0.25525  3.7652551639e-04  0.42134
  3      0.37747  2.1325360311e+00    0.10755  3.0108960328e-06  1.10676
  4      0.49650  1.9431027581e+00    0.11903  6.6308096550e-06  0.94511
  5      0.60899  1.7309148340e+00    0.11249  2.9689281280e-06  1.10987
  6      0.73385  1.4973675890e+00    0.12485  6.6506499526e-07  1.49700
---      -------  ---------------    0.18691  1.7194361476e-05  0.78112
  7      0.87985  1.2550062599e+00    0.14600  4.8701437000e-06  1.00528
  8      1.00000  1.0873099127e+00    0.12015  2.3882050370e-06  -------
Numero total de pasos:  16
```

Figura 1. Ejemplo de interfaz de autodiagnóstico en Jupyter Notebook para la validación intermedia de pasos de integración en MÉTODOs Runge-Kutta.

2.3. Procedimiento

El procedimiento de investigación varía estructuralmente a lo largo de las tres fases del estudio para poner a prueba la sostenibilidad y su capacidad para diferenciar distintos niveles de desempeño académico:

1. Curso 2023-24 (Línea base tradicional): Los estudiantes desarrollaban sus programas de cálculo numérico en Python de forma completamente autónoma y entregaban el software final sin haber recibido guías de validación previas. El docente realizaba una corrección manual integral.
2. Curso 2024-25 (Fase piloto): Se introdujo el instrumento de validación formativa (ejemplos de control intermedio). Con el propósito de testear el sistema de forma controlada, se seleccionaron únicamente dos hitos prácticos críticos para su evaluación cuantitativa exhaustiva (Práctica 3 y Práctica 5). Los estudiantes depuraban sus errores lógicos básicos de forma interactiva antes de la entrega.
3. Curso 2025-26 (Fase de consolidación y escalado): Manteniendo exactamente la misma tecnología de andamiaje y validación formativa en Python del año anterior, se duplicó la exigencia y el volumen del modelo de evaluación continua, pasando de dos a cuatro prácticas evaluadas sistemáticamente con nota numérica (Prácticas 2, 3, 4 y 5).

La estructura del plan de prácticas se mantuvo constante en volumen a lo largo de las fases de estudio, compartiendo el mismo número total de entregas, con la Práctica 1 actuando siempre como una toma de contacto introductoria y no computable en los análisis de rendimiento. La diferencia crítica radicó en el modelo de calificación y la carga de evaluación asociada. En la fase piloto (2024-25), únicamente dos de las prácticas recibieron una calificación numérica exhaustiva, mientras que el resto fue objeto de una evaluación cualitativa simplificada mediante una escala de rangos (A, B, C) con un impacto menor en la nota final. Por el contrario, en la fase de consolidación (2025-26), la optimización del tiempo docente lograda asociada al andamiaje intermedio permitió evaluar el total de las prácticas mediante rúbricas numéricas completas. Este cambio metodológico es de especial relevancia, dado que las revisiones numéricas profundas demandan un tiempo de corrección sustancialmente mayor por estudiante; una exigencia de dedicación que el sistema fue capaz de absorber de forma sostenible.

Para permitir la comparabilidad entre cursos a pesar del diferente número de entregas calificadas cuantitativamente, la nota global del bloque de prácticas se estructuró de forma equivalente en todos los periodos: un 90% de la calificación correspondía a las prácticas evaluadas mediante rúbrica numérica exhaustiva (dividido a partes iguales entre las prácticas calificadas con nota en cada curso: P2, P3, P4 y P5 en 2023-24 y 2025-26; P3 y P5 en 2024-25), mientras que el 10% restante procedía de la evaluación cualitativa en escala categórica (A, B, C) de las entregas restantes. Asimismo, las rúbricas numéricas (escala 0-10), los contenidos temáticos, los algoritmos a implementar y el nivel de dificultad técnico de cada práctica se mantuvieron equivalentes en los tres cursos académicos, evaluando con los mismos criterios de precisión, estabilidad y eficiencia computacional.

A nivel estadístico, el procedimiento de análisis de datos consistió en el cálculo de estadísticos descriptivos robustos (medias, medianas y desviaciones típicas) tanto de las notas globales como de cada práctica de manera pormenorizada. Debido a la naturaleza no paramétrica de las distribuciones, el análisis inferencial se realizó en dos pasos. En primer lugar, se aplicó el test ómnibus de Kruskal-Wallis para evaluar la existencia de diferencias globales entre las tres cohortes. Dado que el resultado fue significativo, se procedió a realizar comparaciones post-hoc por pares mediante el test de suma de rangos de Wilcoxon-Mann-Whitney (ranksum en MATLAB). Al haberse realizado tres compara-

ciones interanuales por pares (23-24 vs. 24-25, 23-24 vs. 25-26 y 24-25 vs. 25-26), los p-valores se ajustaron mediante la corrección de Bonferroni para controlar la tasa de error de tipo I acumulada por comparaciones múltiples.

Cabe señalar que, al tratarse de un diseño cuasiexperimental con cohortes independientes de estudiantes en cada curso académico y no de un seguimiento longitudinal del mismo grupo, las diferencias observadas entre periodos deben interpretarse con cautela. Dichas diferencias no pueden atribuirse de forma inequívoca a la intervención, pudiendo estar influidas por variaciones en la composición o nivel de partida de cada cohorte. Por ello, la significación estadística indica divergencias en las distribuciones, pero no demuestra causalidad directa.

3. RESULTADOS

Los resultados muestran que la implantación del modelo de evaluación continua con validación formativa estuvo asociada a un mantenimiento del rendimiento académico y a cambios en la distribución de las calificaciones. Los hallazgos indican que la provisión de ejemplos numéricos intermedios estuvo acompañada de una reducción de errores triviales observada durante la corrección, lo cual sostuvo la excelencia académica general incluso al duplicar la carga de entregas en el curso 2025-26. Este aumento en la continuidad y exigencia del proceso evaluativo propició un incremento de la dispersión estadística de las notas, sugiriendo una mayor diferenciación entre calificaciones. De forma complementaria, y como hallazgo cualitativo no cuantificado estadísticamente, la reducción del tiempo de corrección de errores básicos permitió al docente dedicar mayor atención a la revisión de la originalidad del código fuente, lo que en el curso 2025-26 condujo a la identificación de tres casos de fraude académico.

Distribución de las Notas de Prácticas por Curso

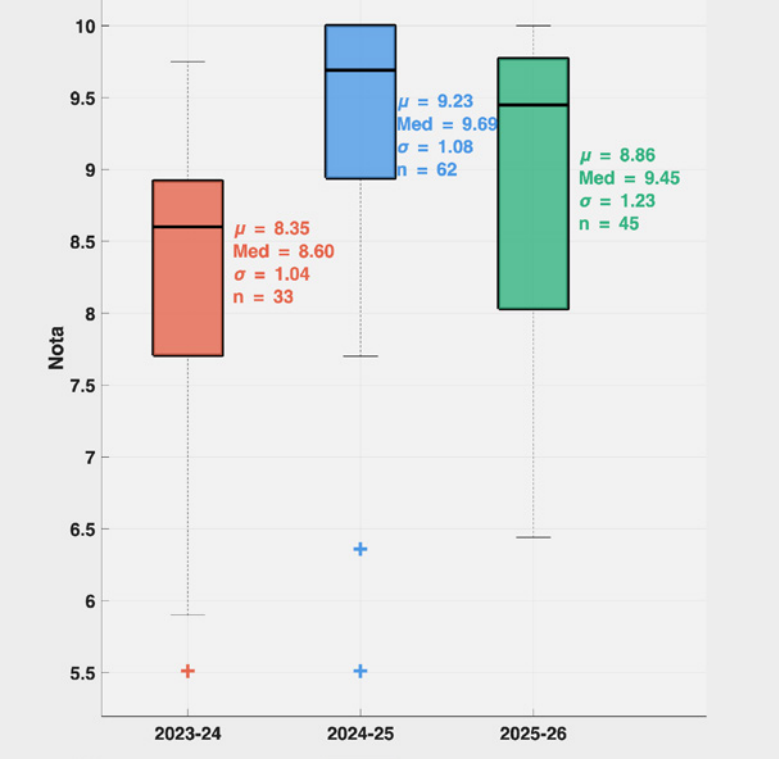


Figura 2. Distribución de las calificaciones obtenidas en las prácticas de Cálculo Numérico II durante los cursos 2023-24, 2024-25 y 2025-26.

El protocolo de actuación ante sospechas de plagio o autoría no legítima —identificadas por la presencia de estructuras lógicas ajenas a los contenidos de la materia, implementaciones de una sofisticación desproporcionada para el nivel del curso o divergencias con el desempeño habitual observado en el aula— consistió en una convocatoria de defensa oral individualizada. Durante esta sesión, los estudiantes debían exponer la arquitectura de su programa y responder a preguntas técnicas planteadas por el profesorado; la falta de dominio y la incapacidad para justificar las decisiones de diseño algorítmico permitieron contrastar y confirmar los casos de copia. Este protocolo es coherente con la evidencia disponible sobre la defensa oral como mecanismo de verificación de autoría e integridad académica en educación superior (Nallaya et al., 2024).

Para justificar estas conclusiones, en primer lugar, se evaluó el rendimiento académico global mediante el análisis de las distribuciones de notas finales de las prácticas. Tal y como se observa en la Figura 2, el curso 2023-24 (línea base tradicional) presentó una mediana de 8,60 y una desviación típica de 1,04 (). La INTRODUCCIÓN de herramientas de validación en el curso 2024-25 generó un evidente efecto techo, elevando la mediana a 9,69 (), en gran parte debido a que la carga evaluativa se concentró únicamente en dos prácticas. Sin embargo, el dato más relevante se encuentra en la fase de consolidación (curso 2025-26): al duplicar el número de prácticas evaluadas, la mediana se estabilizó en un sobresaliente 9,45 (), logrando mantener el alto rendimiento general, pero aumentando la desviación típica global a 1,23. La robustez de estas diferencias se evaluó en dos pasos. En primer lugar, se aplicó el test ómnibus de Kruskal-Wallis sobre la calificación global ponderada del bloque de prácticas de cada estudiante, que confirmó la existencia de diferencias significativas entre los tres grupos ($\chi^2(2)=21{,}38$; $p<0{,}001$). A continuación, se realizaron comparaciones post-hoc por pares mediante el test de suma de rangos de Wilcoxon con corrección de Bonferroni. Los resultados mostraron diferencias significativas entre el modelo tradicional (23-24) y la fase piloto (24-25) ($Z=-4{,}63$; $p<0{,}001$), y entre el modelo tradicional y la fase de consolidación (25-26) ($Z=-2{,}53$; $p=0{,}034$ tras corrección). Por el contrario, la comparación entre las dos fases que emplearon validación formativa (24-25 vs. 25-26) no alcanzó significación estadística tras la corrección ($Z=2{,}06$; p ajustada=0,120), lo que es coherente con la interpretación de que ambas fases comparten un nivel de rendimiento general elevado, siendo la diferencia entre ellas atribuible principalmente al cambio en la exigencia evaluativa y no a una variación sustancial en el rendimiento académico global.

La ausencia de diferencias estadísticamente significativas entre las dos fases con validación formativa (24-25 vs. 25-26) tras la corrección de Bonferroni refuerza la interpretación de sostenibilidad del modelo: pese a duplicar el número de prácticas evaluadas con rúbricas numéricas completas, el rendimiento académico global se mantuvo en niveles de excelencia comparables entre ambas fases. Este resultado es coherente con la hipótesis de que el andamiaje intermedio actuó como factor estabilizador del rendimiento, absorbiendo el incremento en la exigencia evaluativa sin generar un deterioro apreciable en las calificaciones globales del alumnado.

A nivel pormenorizado, la evolución por entregas (ilustrada en la Figura 3 y detallada en la Tabla 1) evidencia la consolidación y el aumento del rigor discriminatorio del modelo continuo. En la Práctica 3, la comparativa entre las cohortes con herramientas refleja que, aunque la mediana se mantuvo en 9,50 tanto en 2024-25 como en 2025-26, la dispersión aumentó (frente a 1,31). Como indicador directo de esta variabilidad, el primer cuartil (o P25) pasó de 9,00 en la fase piloto a 8,68 en la de consolidación, mientras que el rango intercuartílico () se amplió de 1,00 a 1,07 puntos, mostrando una mayor amplitud en el tramo de calificaciones intermedias y bajas. Un comportamiento análogo de atenuación del efecto techo se aprecia en la Práctica 5: mientras que en el curso

2024-25 la distribución se concentró masivamente en el valor máximo (mediana de 10,00, e), en 2025-26 el primer cuartil descendió a y el rango intercuartílico aumentó a puntos, registrando una mayor dispersión en los niveles de desempeño.

Evolución de Notas por Práctica y Curso

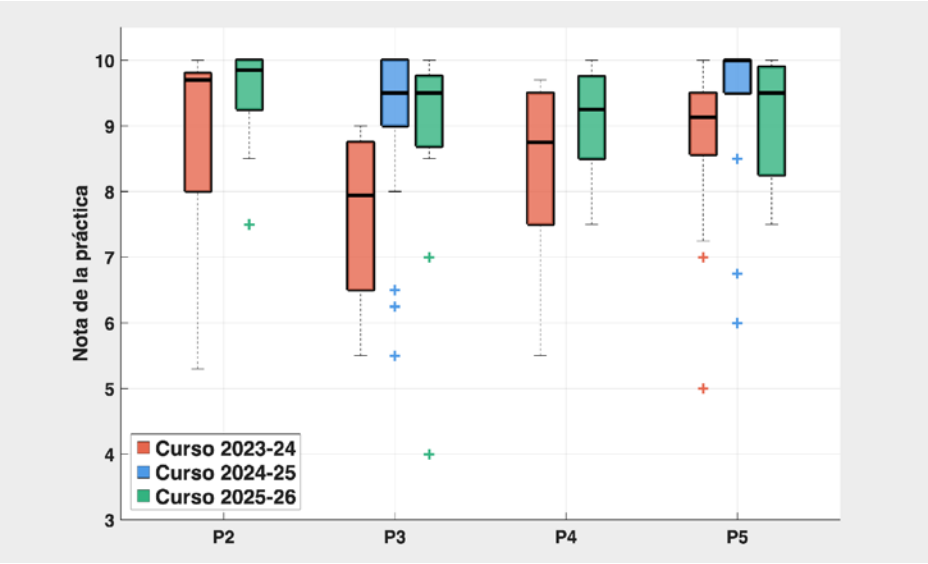


Figura 3. Distribución de las calificaciones para las distintas prácticas en los distintos cursos. En el curso 24-25, dado el gran número de matriculados, solo se evaluaron las prácticas P3 y P5.

De forma complementaria, los datos de las Prácticas 2 y 4 son compatibles con la eficacia del andamiaje algorítmico: en el curso 2025-26, ambas incrementaron sus medianas frente al modelo tradicional (alcanzando 9,85 y 9,25, respectivamente) y experimentaron una fuerte reducción en su desviación típica (0,65 y 0,73 frente al 1,47 y 1,27 del curso 2023-24). Esta reducción de la dispersión en entregas clave indica que una amplia mayoría del alumnado logró asimilar las guías de validación, alcanzando de manera homogénea el estándar de precisión exigido.

Tabla 1. Métricas de las distintas prácticas por curso.

Práctica	Curso	Media	Mediana	Desviación Típica
P2	23-24	8,81	9,70	1,47
	25-26	9,48	9,85	0,65
P3	23-24	7,69	7,94	1,15
	24-25	9,02	9,50	1,31
	25-26	8,64	9,50	1,85
P4	23-24	8,40	8,75	1,27
	25-26	9,16	9,25	0,73
P5	23-24	8,74	9,13	1,26
	24-25	9,58	10	0,89
	25-26	9,09	9,50	0,92

Desde la perspectiva de la práctica docente, se observó cualitativamente una notable optimización del tiempo dedicado a la evaluación de los bloques de código. Si bien no se dispone de un registro cuantitativo formal del tiempo de corrección por expediente, la experiencia del equipo docente indica que la preevaluación algorítmica y el andamiaje previo permitieron focalizar la revisión manual exclusivamente en la estructura del código y los aspectos conceptuales. Al no tener que emplear tiempo en penalizar errores matemáticos o sintácticos básicos en la ejecución de los scripts —ya filtrados por la validación del alumno—, el profesorado pudo aplicar revisiones estructurales profundas en entregas tempranas. Esto permitió detectar patrones anómalos en la lógica de programación durante la evaluación de la Práctica 2 en el curso 2025-26, culminando en la identificación de tres casos de plagio de código, confirmados posteriormente mediante el protocolo de defensa oral descrito anteriormente. Si bien este hallazgo no permite inferir una variación cuantitativa en la tasa global de fraude ni evaluar la presencia de incidencias no detectadas en cohortes previas, sí evidencia que la reorganización del tiempo de corrección facilitó una revisión cualitativa capaz de identificar y verificar activamente casos de plagio en el curso 2025-26.

4. DISCUSIÓN Y CONCLUSIONES

El análisis de la transición e implementación de un modelo de evaluación continua con validación formativa apoyado en herramientas de autodiagnóstico en Python permite dar una respuesta sólida a los objetivos e hipótesis planteados en esta investigación. En primer lugar, los resultados respaldan la hipótesis de sostenibilidad y escalabilidad. Los resultados son compatibles con la interpretación de que los ejemplos de validación actuaron como un andamiaje cognitivo efectivo que permitió a los estudiantes del Grado en Matemáticas depurar de manera autónoma sus errores algorítmicos y de programación básicos. Al delegar la corrección de fallos triviales en los propios alumnos, se optimizó significativamente el tiempo de dedicación docente por entrega. Esta reorganización del proceso de corrección parece haber hecho viable la duplicación del volumen del modelo de evaluación continua en el curso 2025-26 —pasando de 2 a 4 prácticas exhaustivas— sin deteriorar la excelencia académica general, la cual se mantuvo firmemente en una mediana de 9,45.

En segundo lugar, los resultados son compatibles con la idea de que el incremento del número de hitos de evaluación contribuyó a atenuar el efecto techo observado durante la fase piloto, favoreciendo una distribución más amplia de las calificaciones y una mayor diferenciación entre estudiantes con distintos niveles de desempeño. La literatura reciente advierte que los entornos enriquecidos o la facilidad de acceso a herramientas externas de automatización, como la Inteligencia Artificial Generativa, pueden simular una falsa competencia algorítmica si la evaluación es superficial (Prather et al., 2024). Al evaluar únicamente dos hitos en la fase piloto (2024-25), el sistema mostró una concentración artificial de calificaciones máximas (mediana de 10,00 en la Práctica 5). Al aumentar la continuidad pedagógica a cuatro entregas en el curso 2025-26, el modelo de evaluación continua con validación formativa mostró una mayor variabilidad de las calificaciones observándose un incremento del rango intercuartílico (en P5) y un descenso del primer cuartil (en P5 y en P3). Este comportamiento plantea como una hipótesis plausible que una mayor frecuencia de entregas evaluables podría ayudar a que las diferencias en el ritmo de asimilación algorítmica resulten más visibles, en consonancia con lo sugerido por Morris et al. (2021). No obstante, conviene matizar que una mayor dispersión estadística por sí sola no constituye una prueba inequívoca de una mejor capacidad discriminatoria del sistema, pudiendo estar influida también por la heterogeneidad intrínseca entre las distintas cohortes de estu-

diantes o por variaciones en la dificultad técnica de las tareas. Por otra parte, la drástica reducción de la varianza en las Prácticas 2 y 4 (al comparar los cursos 23-24 y 25-26) es coherente con la hipótesis de que, cuando el alumnado supera de forma rigurosa los criterios de estabilidad y precisión exigidos, se logra una convergencia homogénea hacia el estándar de excelencia fijado en la asignatura.

Finalmente, la reestructuración del esfuerzo de corrección del profesorado generó un impacto crítico sobre la integridad académica. En la era actual de la generación automatizada de código, limitar la evaluación a constatar la validez del output numérico entregado resulta insuficiente para verificar la autoría de los trabajos, dado que las herramientas automáticas de detección de plagio de código presentan limitaciones bien documentadas y pueden ser eludidas con relativa facilidad mediante transformaciones superficiales del código (Novak et al., 2019; Pawelczak, 2018). Al quedar liberado de la búsqueda manual de errores de cómputo básicos, la percepción docente indica que fue posible reasignar tiempo a la fiscalización cualitativa estructural y lógica de los programas fuente en las fases iniciales del cuatrimestre, recurriendo a la defensa oral como mecanismo complementario de verificación de autoría (Nallaya et al., 2024). Este análisis cualitativo facilitó la identificación de patrones anómalos y la detección de tres casos de plagio en la Práctica 2 del curso 2025-26, posteriormente confirmados mediante la defensa oral. Conviene precisar que esta reorganización de la carga de trabajo constituye una observación cualitativa no respaldada por métricas cuantitativas directas del tiempo ahorrado, y que la identificación de estos tres casos no demuestra una reducción general del fraude ni permite evaluar la presencia de incidencias no detectadas en cohortes anteriores; no obstante, la experiencia de aula sugiere que esta redistribución del esfuerzo docente representa una vía prometedora para respaldar la honestidad académica en entornos masificados.

Este estudio presenta algunas limitaciones que conviene considerar al interpretar sus resultados. En primer lugar, se desarrolló en una única asignatura de una única titulación, por lo que la generalización de los hallazgos a otros contextos docentes debe realizarse con cautela. En segundo lugar, el diseño cuasiexperimental con cohortes independientes impide atribuir de forma exclusiva los cambios observados a la intervención, ya que no puede descartarse la influencia de diferencias entre promociones o de otros factores no controlados. A estas salvedades debe añadirse la limitación inherente a la doble función del profesorado como diseñador de la propuesta y evaluador de la experiencia, si bien el uso de rúbricas objetivas y el tratamiento anónimo y agregado de los datos contribuyeron a minimizar posibles sesgos en la interpretación de los resultados. Finalmente, algunos de los beneficios descritos, como la optimización del tiempo de corrección o la mejora en la detección de posibles casos de plagio, se sustentan en observaciones cualitativas derivadas de la práctica docente y no fueron cuantificados mediante indicadores específicos, por lo que sería conveniente abordarlos de forma más sistemática en investigaciones futuras.

A modo de conclusión, los resultados obtenidos en este contexto respaldan la viabilidad del modelo de validación formativa mediante Python y su potencial escalabilidad para la enseñanza del cálculo numérico de ecuaciones diferenciales ordinarias. La experiencia muestra que fue posible ampliar el número de prácticas calificadas manteniendo un rendimiento académico elevado y con indicios descriptivos de una mayor capacidad para diferenciar niveles de competencia, al tiempo que sugiere una distribución más fluida del trabajo docente para el control cualitativo de las entregas. No obstante, su eficacia pedagógica global y su impacto preciso sobre la carga docente deberán verificarse en estudios adicionales mediante indicadores cuantitativos específicos. Como líneas de trabajo futuro, se plantea la integración de sistemas automáticos de análisis de control de versiones en repositorios de código, lo que permitirá monitorizar de forma continua e incremental el flujo de desarrollo lógico de los estu-

diantes, mitigando aún más el impacto del fraude académico y consolidando el enfoque del proceso de aprendizaje frente al producto final.

5. REFERENCIAS

Becker, B. A., Denny, P., Finnie-Ansley, J., Luxton-Reilly, A., Prather, J., & Santos, E. A. (2023). Programming is hard - or at least it used to be: Educational opportunities and challenges of AI code generation. En *Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1* (pp. 500-506). Association for Computing Machinery. https://doi.org/10.1145/3545945.3569759

Denny, P., Prather, J., Becker, B. A., Finnie-Ansley, J., Hellas, A., Leinonen, J., Luxton-Reilly, A., Reeves, B. N., Santos, E. A., & Sarsa, S. (2024). Computing education in the era of generative AI. Communications of the ACM, 67(2), 56-67. https://doi.org/10.1145/3624720

Finnie-Ansley, J., Denny, P., Luxton-Reilly, A., Santos, E. A., Prather, J., & Becker, B. A. (2023). My AI wants to know if this will be on the exam: Testing OpenAI's Codex on CS2 programming exercises. En Proceedings of the 25th Australasian Computing Education Conference (ACE'23) (pp. 97-104). Association for Computing Machinery. https://doi.org/10.1145/3576123.3576134

Ketcheson, D. I. (2014). Teaching numerical methods with IPython notebooks and inquiry-based learning. En Proceedings of the 13th Python in Science Conference (SciPy 2014) (pp. 96-99).

Morris, R., Perry, T., & Wardle, L. (2021). Formative assessment and feedback for learning in higher education: A systematic review. Review of Education, 9(3), e3292. https://doi.org/10.1002/rev3.3292

Nallaya, S., Gentili, S., Weeks, S., & Baldock, K. (2024). The validity, reliability, academic integrity and integration of oral assessments in higher education: A systematic review. Issues in Educational Research, 34, 629-646.

Novak, M., Joy, M., & Kermek, D. (2019). Source-code similarity detection and detection tools used in academia: A systematic review. ACM Transactions on Computing Education, 19(3), Article 27.

Pawelczak, D. (2018). Benefits and drawbacks of source code plagiarism detection in engineering education. En Proceedings of the IEEE Global Engineering Education Conference (EDUCON 2018) (pp. 1048-1056). IEEE. https://doi.org/10.1109/EDUCON.2018.8363346

Prather, J., Denny, P., Leinonen, J., Becker, B. A., Albluwi, I., Craig, M., Keuning, H., Kiesler, N., Kohn, T., Luxton-Reilly, A., MacNeil, S., Petersen, A., Pettit, R., Reeves, B. N., & Savelka, J. (2023). The robots are here: Navigating the generative AI revolution in computing education. En Proceedings of the 2023 Working Group Reports on Innovation and Technology in Computer Science Education (ITiCSE-WGR'23) (pp. 108-159). Association for Computing Machinery. https://doi.org/10.1145/3623762.3633499

Prather, J., Reeves, B. N., Leinonen, J., MacNeil, S., Randrianasolo, A. S., Becker, B. A., Kimmel, B., Wright, J., & Briggs, B. (2024). The widening gap: The benefits and harms of generative AI for novice programmers. En Proceedings of the 2024 ACM Conference on International Computing Education Research V. 1 (pp. 469-486). Association for Computing Machinery. https://doi.org/10.1145/3632620.3671116

Staus, N. L., O'Connell, K., & Storksdieck, M. (2021). Addressing the ceiling effect when assessing STEM out-of-school time experiences. Frontiers in Education, 6, 690431. https://doi.org/10.3389/feduc.2021.690431

Docencia universitaria en transformación: escenarios emergentes

Esta obra reúne investigaciones y experiencias de innovación que reflejan los cambios que atraviesa actualmente la educación superior. El volumen explora cómo las universidades responden a nuevos escenarios pedagógicos, tecnológicos y sociales, y cómo estas transformaciones modifican las formas de enseñar, aprender y evaluar.

Uno de los ejes centrales es la incorporación de la inteligencia artificial, especialmente la generativa, a diferentes ámbitos de la enseñanza universitaria. Los capítulos analizan sus posibilidades para favorecer el aprendizaje adaptativo, el autoaprendizaje y el desarrollo de competencias, al tiempo que abordan desafíos relacionados con la evaluación, la integridad académica, la trazabilidad de los procesos de aprendizaje y el pensamiento crítico.

Junto a la dimensión tecnológica, la obra aborda propuestas centradas en el aprendizaje autónomo y colaborativo, las metodologías activas, la evaluación formativa y el desarrollo de competencias específicas. Desde una perspectiva interdisciplinar, los capítulos trasladan estas cuestiones a diferentes ámbitos de la educación superior, mostrando la diversidad de contextos y prácticas en los que se concreta la innovación docente.

Las contribuciones comparten una perspectiva que entiende la innovación docente como un proceso de transformación de las prácticas educativas y no únicamente como la incorporación de nuevas herramientas. En este sentido, la obra ofrece una visión plural y transversal de los retos y oportunidades de la universidad contemporánea, destacando la necesidad de promover aprendizajes significativos, pensamiento crítico, autonomía y una integración reflexiva y responsable de las tecnologías emergentes.



Rosana Satorre Cuerda es Doctora en Ingeniería Informática, Titular de Universidad en Ciencia de la Computación e Inteligencia Artificial de la Universidad de Alicante [UA]. Ha ocupado los cargos de Subdirectora (2000-2004) y Directora en funciones (2004-2005) del Departamento, y Subdirectora de las Titulaciones de Informática (2005-2009) y Secretaria (2009-2013) de la Escuela Politécnica Superior de la UA. Actualmente es Directora del Instituto de Ciencias de la Educación de esta