

La Inteligencia Artificial en la formación inicial docente: Elaboración de un instrumento de evaluación basado en el marco de la UNESCO y el modelo TPACK

Victoria Tornel Bastida – Universidad Católica San Antonio de Murcia

 0009-0005-3611-950X

María del Mar Rodríguez Rosell – Universidad Católica San Antonio de Murcia

 0000-0003-0837-8473

Irene Melgarejo Moreno – Universidad Católica San Antonio de Murcia

 0000-0003-0837-8473

Recepción: 22.06.2026 | Aceptado: 02.07.2026

Correspondencia a través de **ORCID**: Victoria Tornel Bastida

 **0009-0005-3611-950X**

Citar: Tornel Bastida, V, Rodríguez Rosell, MM, & Melgarejo Moreno, I (2026). La Inteligencia Artificial en la formación inicial docente: Elaboración de un instrumento de evaluación basado en el marco de la UNESCO y el modelo TPACK. *REIDOCREA*, 15(44), 595-610.

Estudio de investigación de Tesis Doctoral

Área o categoría del conocimiento: La formación docente en Educación Superior

Resumen: La irrupción de la inteligencia artificial en la universidad plantea retos éticos y pedagógicos que requieren reforzar las competencias digitales del profesorado. Este estudio presenta la elaboración y validación de un cuestionario para medir la formación en IA del profesorado universitario que imparte docencia en los grados de Educación Infantil y Primaria en España. Se siguió un diseño mixto, exploratorio y descriptivo, y se desarrolló un instrumento ad hoc organizado por bloques temáticos basados en el modelo TPACK y en el Marco de Competencias en IA de la UNESCO. La validación de contenido se realizó mediante juicio de expertos y un estudio piloto de 35 docentes. Los resultados mostraron una elevada concordancia entre expertos, con un Índice de Validez de Contenido de 1,00 en el 91,8 % de los ítems, y una buena comprensión del cuestionario en la fase piloto. Además, el instrumento presentó una alta consistencia interna, con un alfa de Cronbach global de 0,873. En conjunto, los hallazgos indican que el cuestionario es válido, fiable y operativo para evaluar la capacitación y alfabetización en IA del profesorado universitario, ofreciendo un diagnóstico integral alineado con los estándares vigentes.

Palabra clave: Alfabetización en IA

Artificial Intelligence in initial teacher training: developing an assessment tool based on the UNESCO framework and the TPACK model

Abstract: Artificial intelligence in higher education presents ethical and pedagogical challenges that require strengthening teachers' digital competencies. This study presents the development and validation of a questionnaire designed to assess the level of AI training among university faculty teaching in Early Childhood and Primary Education degree programs in Spain. A mixed, exploratory, and descriptive design was used, and an ad hoc instrument was developed with thematic blocks based on the TPACK model and UNESCO's AI Competency Framework. Content validation was carried out through expert judgment and a pilot study with 35 university teachers. The results showed a high level of agreement among experts, with a Content Validity Index of 1.00 in 91.8% of the items, and good comprehension of the questionnaire in the pilot phase. In addition, the instrument showed high internal consistency, with an overall Cronbach's alpha of 0.873. Overall, the findings indicate that the questionnaire is a valid, reliable, and practical tool for assessing university teachers' AI training and literacy, providing a comprehensive diagnosis aligned with current international standards

Keyword: AI Literacy

Introducción

La formación del profesorado de Educación Superior es crucial para poder atender a los desafíos que se presentan en la actualidad, y como consecuencia atender a las necesidades que los estudiantes universitarios tendrán en su futuro laboral más próximo, o lo que es lo mismo, a las demandas del mercado laboral.

Uno de los desafíos que en este momento ha irrumpido con fuerza en el sistema educativo y como consecuencia en el ámbito universitario es la Inteligencia Artificial (IA).

Debido al uso continuado, excesivo y sin filtro por parte de los estudiantes en tareas propias de este ámbito académico (Jaramillo et al., 2026; Espinoza et al., 2026; Islas, & Rodríguez-Betanzos, 2026), se han observado consecuencias como la dependencia, pereza cognitiva, uso indebido y poco ético de esta tecnología (Condori et al., 2026; Pereira & Idoiaga, 2026).

Para dar respuesta a esta situación, fomentar y favorecer el uso adecuado de la misma, junto con la utilización pedagógica de la IA en el ámbito de enseñanza universitaria por parte del personal docente, se hace imprescindible, según Almenara (2016), tomar como referencia el modelo TPAK o del Conocimiento Tecnológico Pedagógico y Disciplinar propuesto por Koehler y Mishra (2009), basado en el uso adecuado de las TIC para la adquisición de conocimientos instrumentales, disciplinares y metodológicos que permitan integrarlas de manera adecuada en el ámbito educativo (Cabero et al., 2018; Giler et al., 2026). Según Hernández y Domínguez (2025), el Marco de Competencias para docentes en materia de IA de la diseñado por la UNESCO en 2024, es otra de las bases fundamentales a considerar en este escenario, dado que permite abordar de manera directa conocimientos, habilidades y valores, que en este sentido los profesores de cualquier etapa educativa deben de adquirir como propias, dentro de sus competencias profesionales, atendiendo de manera concreta a cinco dimensiones: la ética de la IA, la IA centrada en el ser humano, fundamentos y aplicaciones de la IA, la pedagogía de la IA y la IA para el aprendizaje profesional (Fengchun & Cukurova, 2024).

Estudios realizados por Mondragón et al., (2025), destacan la importancia de implementar programas de formación para el profesorado universitario que lleven al fortalecimiento de las competencias digitales docentes en materia de IA, con la finalidad de sacar el máximo potencial de esta tecnología para su práctica docente. Otros como Berríos (2026), afirman que estos planes de formación no solo benefician al profesorado en su práctica profesional, sino que, además, permitirá integrar la IA de una manera crítica y segura en el proceso de enseñanza y aprendizaje del estudiante universitario.

Tomando como referencia cuestiones como la alfabetización en IA del profesorado de Educación Superior (Vásquez, 2025), el uso ético de esta tecnología por parte del estudiantado universitario (Alastor et al., 2024), así como la actualización de las competencias digitales docentes en esta materia para el profesorado de esta etapa (Riquelme 2022; Ordoñez, 2026; Alcántara, 2026), se plantea la necesidad de conocer cuál es la formación del profesorado universitario de los grados de Educación Infantil y Primaria en Inteligencia Artificial a nivel estatal aplicada como recurso para el proceso de enseñanza-aprendizaje, tratando de dar respuesta a estudios como el desarrollado por Galindo et al., (2024) ante el uso poco ético de sistemas basados en IA como Chat GPT por parte del alumnado del grado de Educación Primaria, así como otros más concretos como el implementado por García et al., (2025), el cual afirma la necesidad imperante de formar al alumnado universitario de los grados de educación infantil y primaria en el uso pedagógico de la IA que les permitan entre otros aspectos reforzar sus competencias digitales como futuros docentes.

Como conclusión, y con motivo de los diversos estudios anteriormente comentados sobre la actualización de las competencias digitales docentes en IA por parte de los estudiantes del grado de Educación Infantil y Primaria, se parte de la siguiente cuestión: “¿cómo están de preparados los profesores universitarios que imparten formación a los futuros maestros de nuestro país en materia de IA aplicándola en el proceso de enseñanza y aprendizaje como recurso educativo?”. Esto permitirá describir y analizar la situación actual de la capacitación en IA del profesorado universitario en España.

Método

El presente estudio adopta un diseño metodológico mixto, de carácter exploratorio-descriptivo y no experimental de corte transversal, que combina de forma complementaria una fase cuantitativa -basada en un cuestionario- y cualitativa – con la ampliación de respuestas abiertas ante cuestiones que inicialmente son de tipo cerradas y que necesitan ser mayormente argumentadas por los participantes - que permite triangular resultados y enriquecer la comprensión del fenómeno estudiado (Arenas, 2021). Además, siguiendo a Bisquerra (2009), se clasifica como exploratorio por la escasa investigación previa existente sobre la IA en la docencia universitaria sobre el profesorado que imparte formación en los grados de Educación Infantil y Primaria en España, y descriptivo ya que trata de exponer la situación actual de este contexto en concreto, sin manipular variables y limitándose a observar la realidad existente (Ochoa & Yunkor, 2019).

La población objeto de estudio estuvo compuesta por docentes que imparten formación a los futuros maestros y maestras en España, tanto en universidades de carácter público y privado, así como centros asociados que también ofertan dichos estudios superiores.

Para ello, se empleó un muestreo no probabilístico estratificado por tipo de universidad -privada o pública- que garantizó la representatividad institucional, teniendo como objetivo mínimo orientativo una muestra de $n \geq 385$ docentes, y tomando como criterio de exclusión “no aceptar el consentimiento informado” que da acceso a la realización del cuestionario.

Instrumento

Para la recolección de datos se empleó como instrumento principal un cuestionario ad hoc, estructurado en seis bloques temáticos -datos sociodemográficos, conocimiento general y formación en IA, uso de la IA en la docencia universitaria, políticas institucionales e impacto de la IA en la educación, pensamiento crítico y ético del profesorado y alumnado, conocimiento avanzado en IA- que permitiesen medir de manera sistemática y abarcar de forma integral todas las variables y dimensiones relevantes para el estudio.

En cuanto al tipo de ítems, este instrumento estuvo principalmente compuesto por 42 preguntas de tipo cerradas, incluyendo preguntas de respuesta única y escalas ordinales tipo Likert, lo que permitió evaluar de manera estandarizada la frecuencia de uso, el grado de acuerdo, así como diversas percepciones y actitudes (Rosa de Saa, 2012). También, algunas preguntas incorporaban la opción “Otro (especificar)”, facilitando la recogida de información adicional sin afectar la consistencia del análisis cuantitativo.

El tiempo estimado para la cumplimentación del cuestionario, se situó en unos 15 minutos aproximadamente. En cuanto a su realización fue en modalidad virtual y en línea, con la finalidad de garantizar el anonimato y confidencialidad de las respuestas. Además, atendiendo al Reglamento General para la Protección de Datos (RGPD) , se incluyó previo a la realización y participación del cuestionario, un texto informativo dirigido a los participantes sobre el tratamiento de sus datos conforme a esta normativa, el objetivo de la investigación, así como el carácter voluntario de su participación.

Revisión y validación del cuestionario

Para garantizar la calidad y pertinencia del instrumento diseñado para el presente estudio, se sometió al cuestionario a un proceso riguroso de revisión y validación del contenido (Montero et al., 2024). En un primer momento, siguiendo a Urrutia et al., (2014), se trató de evaluar el grado en que los ítems reflejaban adecuadamente las dimensiones teóricas y objetivos del estudio, buscando la validez del contenido. Para ello, se recurrió al juicio de expertos, mediante la revisión del borrador del cuestionario por un panel de expertos en educación, estadística y tecnología, incluyendo a profesores e investigadores familiarizados con la temática de la IA en educación, así como expertos en el diseño de instrumentos. Estos evaluadores analizaron cada ítem atendiendo a los criterios de relevancia, calidad, coherencia y cobertura respecto a los objetivos planteados (Escobar-Pérez & Cuervo-Martínez, 2008). Así mismo se solicitó a los expertos que valoraran las cuatro características propuestas en cada ítem para determinar su adecuación a los aspectos que se pretendían medir.

Para fortalecer la validez del contenido, se tomó como referencia el Marco de Competencias en IA para docentes, así como el modelo TPACK anteriormente comentados. Con ello, se pretendía cubrir las cinco dimensiones claves establecidas por la UNESCO (2024), y competencias propias que el docente del siglo XXI debe de adquirir para una correcta función y capacitación docente, siguiendo el modelo de Koehler y Mishra (2009).

También se consideraron aspectos como la claridad semántica y formulación de los ítems, la secuencia y organización de las secciones o bloques en los que estaba estructurado el cuestionario, de tal forma que el lenguaje utilizado fuese comprensible y no excesivamente técnico y ambiguo; Además, esto también ayudó a que la carga cognitiva no fuese excesiva, teniendo en cuenta que los futuros participantes no tenían que porqué dominar el tema principal del contenido del cuestionario.

Para comprobar que hubiera una coherencia entre los objetivos de la investigación y los ítems del cuestionario, se utilizó una matriz de congruencia (Rendón, 2001) -asegurando que ninguno de los objetivos quedara sin cobertura y que cada ítem cumpliera una función- buscando que tanto los objetivos específicos del estudio como las preguntas del instrumento estuviesen adecuadamente representados. Esta correspondencia, permitió fortalecer la validez del constructo del instrumento y que los resultados obtenidos (Gil et al., 2000), pudiesen interpretarse directamente en relación con las preguntas de investigación planteada.

Una vez finalizado el juicio de expertos, se realizó un estudio piloto en el que participaron 35 docentes universitarios que impartían formación en los grados de Educación Infantil y Primaria tanto de universidad públicas, privadas, así como de centros asociados que ofertan esta formación en nuestro país. En este proceso se trató de evaluar la comprensibilidad y funcionamiento de cada ítem, detectar preguntas confusas o difíciles, y obtener datos preliminares para analizar la confiabilidad y estructura interna del instrumento, así como el tiempo promedio de respuesta del participante entre una pregunta y otra, junto con aquellas preguntas susceptibles de mejoras.

Obtenidos los datos de la prueba piloto se calculó la consistencia interna mediante el coeficiente alfa de Cronbach (Oviedo & Campo-Arias, 2005), tratando de alcanzar un valor $>0,7$ que determinara la coherencia, se evaluó de nuevo los ítems que pudiesen ser discordantes, y se realizó un análisis factorial exploratorio (Pérez & Medrano, 2010) con el fin de verificar que los ítems se agrupaban según las dimensiones prevista. Esta

fase final, y necesaria a realizar antes de la aplicación del cuestionario final, garantizó que el instrumento fuese válido, fiable y adecuado para la recolección de datos.

Resultados

Resultados del juicio de expertos

Con el objetivo de valorar la validez de contenido del cuestionario, se realizó el análisis cuantitativo mediante el cálculo de la media de las respuestas obtenidas y desviación típica (medición de la dispersión de datos) para cada dimensión en los distintos ítems del instrumento, junto con las puntuaciones asignadas por los tres expertos - educación, estadística y tecnología- atendiendo a los siguientes criterios:

- Claridad o grado en el que los ítems estaban redactados para ser entendidos de manera comprensible y precisa por parte de los encuestados.
- Relevancia: medida en la que el ítem presenta importancia para los objetivos del estudio.
- Coherencia o grado de lógica de un ítem con el resto de cuestiones planteadas y bloque al que pertenece, dentro del cuestionario.
- Cobertura: grado en el que el ítem representa de manera adecuada los contenidos o dimensiones que se pretendían evaluar dentro del instrumento.

Los resultados mostraron puntuaciones muy altas, teniendo en cuenta que la escala de valoración de tipo Likert se situaba del 1 (nada claro) a 5 (totalmente claro) en cada uno de estos criterios – claridad (4,72), relevancia (4,88), coherencia (4,90) y cobertura (4,71)- indicando un alto grado de adecuación del cuestionario en todas sus dimensiones evaluadas. A nivel de bloques, los bloques 1 y 3 sobre “Datos sociodemográficos y uso de la IA en la docencia universitaria” alcanzaron puntuaciones perfectas, es decir de 5,00 en todos los criterios, destacando su comprensibilidad, pertinencia y alineación de los ítems frente al objetivo del estudio. En los bloques 5 y 6, sobre “práctica docente y competencias” y “competencias avanzadas y aplicación de la IA” también se observó un excelente rendimiento, concretamente en coherencia (5,00) y con medias superiores a 4,8 en todos los criterios.

En el caso del bloque 7 “Pensamiento crítico del profesorado y alumnado” que posteriormente se reubicó junto al seis formando un solo bloque, se obtuvieron valores ligeramente inferiores en los criterios de claridad (4,50), y cobertura (4,43), aunque estos estaban dentro de una franja alta de valoración. También se identificaron ítems con mayor variabilidad entre expertos, reflejada en desviaciones típicas iguales o superiores a 0,94 en los ítems 16, 19, 30, etc. y de forma más acentuada en el ítem 47 con una puntuación de 1,41 en tres de los cuatro criterios, indicando desacuerdo o diferentes interpretaciones entre los diferentes expertos, planteando la necesidad de una revisión o reformulación de los mismos con la finalidad de aportar claridad y precisión al instrumento.

No se observaron medias por debajo del umbral de 3,5 en ningún ítem, lo que reforzó la solidez general del cuestionario. Para reforzar la validez del contenido, se calculó el Índice de Validez de Contenido (IVC) según Lawshe (1975) y adaptado por Tristán-López (2008) para paneles reducidos. Con tres expertos, un ítem se considera válido si todos coinciden en que es “esencial”, correspondiendo a un IVC de 1,00. Los resultados indicaron que 45 de los 49 ítems (91,8 %) alcanzaron un IVC =1, evidenciando un alto grado de consenso sobre su relevancia. Por el contrario, cuatro ítems -7, 19, 30 y 31- obtuvieron un IVC = 0,33, indicando menor acuerdo y sugiriendo la necesidad de revisar su formulación o pertinencia.

Tabla 1.

Trazabilidad de modificaciones en ítems del cuestionario.

Ítems	Criterios comprometidos	IVC	DT ALTA (<0,94)	Propuesta de acción
7	Relevancia	0,33	0,94 (relevancia y cobertura)	Revisión en la redacción, reformulación o eliminación si es redundante.
16	Claridad	1,00	0,94 (claridad)	Reformulación para mejorar la precisión y evitar la ambigüedad.
19	Relevancia y cobertura	0,33	0,94 (relevancia y cobertura)	Comprobación de la alineación con el objetivo propio del bloque.
30	Claridad y relevancia	0,33	0,94 (claridad y cobertura)	Rediseño con lenguaje más accesible.
31	Claridad y relevancia	1,00	0,94 (claridad, relevancia y cobertura)	Reformular; posible fusión con ítems cercanos temáticamente.
33	Claridad	1,00	0,94 (claridad)	Añadir definición de “API” o Interfaz de Programación de aplicaciones para mejorar la accesibilidad al conocimiento y comprensión.
34	Claridad	1,00	0,94 (claridad y cobertura)	Mantener, pero mejorar la comprensión del ítem.
36	Cobertura	1,00	0,94 (cobertura)	Revisar formulación para asegurar inclusión de todos los elementos relevantes.
38	Cobertura	1,00	0,94 (cobertura)	Ampliar opciones de respuesta y alternativas como “otro” o “todos son relevantes”
40	Claridad	1,00	0,94 (claridad)	Revisión de reescritura y vocabulario técnico.
43	Claridad	1,00	0,94 (claridad)	Ajustar semántica
47	Claridad, coherencia y cobertura	1,00	1,41 (claridad, coherencia y cobertura)	Revisión profunda; posible reformulación o división en dos ítems.
—	Homogeneidad del formato	—	—	Convertir los ítems idóneos de los bloques 2 a 7 a escala tipo Likert para mejorar la discriminación de respuestas y la coherencia estructural del cuestionario.

Tras los resultados obtenidos, y anteriormente detallados (Tabla 3), se procedió a una revisión detallada del instrumento, incorporando los siguientes cambios y/o ajustes en: la reformulación de ítems problemáticos con los ítems 7, 16, 19, 30 y 31; la corrección de errores técnicos y de terminología en el ítem 43 cambiando “RPGD” por “RGDP” y revisando el vocabulario de los ítems 33 y 40 para hacerlo más accesible para aquellos docentes universitarios que no contaran con formación o conocimiento en IA; la revisión del contenido y cobertura de los 36 y 38; la reestructuración de ítems complejos -como el caso del 47- en dos, para mejorar su coherencia interna y la precisión en la recogida de información; y la transformación de varios ítems de los bloques 2 a 6 a un formato de escala tipo Likert, con la finalidad de incrementar la consistencia metodológica del instrumento, facilitar la discriminación de respuestas y mejorar la comparabilidad de los resultados, tratando de dar un aspecto más homogéneo al cuestionario en su formato.

Resultados de la prueba piloto

Tras la revisión completa del cuestionario inicial y juicio de expertos, se elaboró una versión revisada del cuestionario, incluyendo los ajustes y cambios anteriormente mencionados, y siendo la versión proporcionada para la prueba piloto, con el objetivo de mejorar la claridad conceptual, homogeneizar formatos de respuesta, reforzar la coherencia entre bloques, así como comprobar aspectos relacionados con el tiempo estimado en la realización del cuestionario, funcionamiento de la plataforma (*Google Forms*) y adecuación general antes de su aplicación definitiva.

Los resultados obtenidos de la prueba piloto, evidenciaron un adecuado funcionamiento general del instrumento, con un elevado nivel de comprensión en la totalidad de sus ítems. Tanto las preguntas sociodemográficas como los bloques relativos al conocimiento, uso e integración de la IA en la docencia fueron interpretados

correctamente, sin detectarse dificultades en la lectura, ambigüedades relevantes ni problemas de redacción. Las escalas de respuesta fueron comprendidas de forma homogénea, lo que refleja coherencia y variabilidad de las respuestas obtenidas.

En términos operativos, no se registraron incidencias durante la administración del cuestionario, confirmando la idoneidad del instrumento en condiciones reales de aplicación. El tiempo de cumplimentación se situó en 12-18 minutos, con una media aproximada de 15 minutos, en línea con la estimación inicial, lo que indica una adecuada extensión del cuestionario y una carga cognitiva asumible para el perfil de la muestra.

La muestra piloto estuvo compuesta por 35 participantes, predominando el género femenino (71,4%), profesorado de instituciones privadas (88,6%) y niveles elevados de experiencia docente (45 % con más de 15 años). El conocimiento de herramientas de IA fue generalizado, dado que el 100 % conocía herramientas como Chat GPT, Copilot o Gemini, aunque la formación específica en IA mostró heterogeneidad.

Con el fin de garantizar la calidad psicométrica del instrumento, se llevó a cabo un proceso de validación que incluye el análisis de fiabilidad mediante el coeficiente alfa de Cronbach y la validación de constructo mediante análisis factorial exploratorio (AFE). El instrumento incluye 49 ítems tipo Likert que miden diferentes dimensiones relacionadas con el conocimiento, uso, percepciones, actitudes, barreras y competencias en inteligencia artificial en el ámbito docente.

Tabla 2.

Análisis de fiabilidad interna por bloques del cuestionario.

Bloque	Ítems incluidos	N válido	Nº ítems	Alfa de Cronbach
Bloque 2. Conocimiento y formación en IA	P7-P10	32	5	0,69
Bloque 3. Uso, percepciones y barreras en IA	P11-P16	35	12	0,826
Bloque 4. Actitudes e identificación de riesgos	P19-P21	35	6	0,676
Bloque 5. Práctica docente y desarrollo de competencias.	P22-P26	34	7	0,758
Bloque 6. Competencias avanzadas y aplicación de IA	P28-P40	35	19	0,768

El valor obtenido de alfa de Cronbach global ($\alpha = 0,873$) indica un nivel elevado de consistencia interna, lo que sugiere que el conjunto de ítems presenta una adecuada coherencia global para la medición de los constructos analizados. El análisis de fiabilidad por bloques (Tabla 3) muestra que el Bloque 3 (Uso, percepciones y barreras) presenta la mayor consistencia interna ($\alpha = 0,826$), mientras que los Bloques 5 ($\alpha = 0,758$) y 6 ($\alpha = 0,768$) alcanzan niveles de fiabilidad adecuados. El Bloque 2 muestra una consistencia interna límite ($\alpha = 0,699$), cercana al umbral de aceptabilidad, y el Bloque 4 presenta $\alpha = 0,676$, nivel aceptable en estudios exploratorios. El análisis de "alfa si se elimina el ítem", expuesto en la Tabla 4, identifica que la supresión de P9, P15_4 y P33_3 podría mejorar marginalmente la consistencia interna de sus bloques respectivos.

Tabla 3.

Análisis de fiabilidad interna del cuestionario.

N de casos válidos	31
Nº de ítems	49
Alfa de Cronbach	0,873
Rango de medias de los ítems	1,29-4,39
Rango de desviaciones estándar	0,558-1,350
Correlación ítem-total corregida (min-máx.)	-0,169-0,679
Alpha si el elemento se suprime (mín-máx)	0,865-0,880

Los análisis factoriales exploratorios por bloques (Tabla 5), mostraron que la prueba de Bartlett resultó estadísticamente significativa en todos los bloques ($p < 0,001$). Los Bloques 2 y 3 presentaron estructuras factoriales interpretables con soluciones

bifactoriales y multifactoriales coherentes, respectivamente, y porcentajes de varianza explicada superiores al 60% (70,63% y 61,79%). El Bloque 4 obtuvo una solución factorial válida con dos factores diferenciados (50,14% de varianza explicada), aunque con cierta heterogeneidad. Por el contrario, en los Bloques 5 y 6 no fue posible obtener soluciones factoriales estables e interpretables, lo que puede atribuirse a la complejidad conceptual de los bloques y al tamaño muestral reducido.

Tabla 4. Resultados del análisis factorial exploratorio por bloques.						
Bloque	KMO	Bartlett χ^2 (gl)	p	Resultado de extracción	Factores extraídos	% varianza explicada
Bloque 2	0,615	71,243 (10)	< 0,001	Válida	2	70, 63 %
Bloque 3	0,725	264,633 (66)	< 0,001	Válida	3	61,79%
Bloque 4	0,699	59,432 (15)	< 0,001	Válida	2	50,14%
Bloque 5	0,657	116,464 (21)	< 0,001	No estable	-	-
Bloque 6	0,616	357,014 (171)	< 0,001	No estable	-	-

El análisis factorial confirmatorio (AFC) del modelo global mostró un ajuste insuficiente (CFI = 0,844; TLI = 0,835; RMSEA = 0,203; SRMR = 0,212). El AFC identificó ítems con cargas factoriales bajas, no significativas o negativas: P14_4, P15_4, P20, P24, P25, P33_3, P35_3, P36_4 y P40_4, que se consideran candidatos a revisión.

A partir de la revisión cualitativa de la prueba piloto y los análisis psicométricos, se propusieron los siguientes ajustes menores:

- En el ítem P5 (área de especialización), se incorporó explícitamente la categoría "Psicología", dada su relevancia en el contexto educativo y su infrarrepresentación en la opción abierta (3 participantes la mencionaron).
- Se planteó la eliminación de las opciones abiertas en los ítems P37 y P38, ya que no aportaron información adicional significativa y podrían dificultar la homogeneización de las respuestas.
- En el ítem P39, se reformuló la categoría "gestión ética de datos" hacia "ética y buenas prácticas", con el fin de mejorar su comprensión y abarcar un espectro conceptual más amplio.
- Se recomendó una revisión selectiva de los ítems P9, P15_4 y P33_3 desde el enfoque de fiabilidad, así como de los ítems identificados con cargas factoriales bajas en el AFC (P14_4, P15_4, P20, P24, P25, P33_3, P35_3, P36_4 y P40_4), para mejorar su ajuste teórico y psicométrico.

Discusión y conclusiones

Tomando como referencia el objetivo principal de este estudio, para conocer cuál es el grado de formación del profesorado universitario que imparte formación en los grados de Educación Infantil y Primaria en IA aplicada como recurso de enseñanza-aprendizaje, se puede decir que el cuestionario diseñado para ello, presenta una validez en el contenido, dado que trata las cinco dimensiones fundamentales - la ética de la IA, la IA centrada en el ser humano, fundamentos y aplicaciones de la IA, la pedagogía de la IA y la IA para el aprendizaje profesional- por bloques y en los distintos ítems, propuestas en el Marco de Competencias docentes en materia de IA por la UNESCO (2024) para la actualización competencial y alfabetización del profesorado en Inteligencia Artificial. Así mismo, también se contempla el modelo TPACK o del Conocimiento Tecnológico Pedagógico y Disciplinar propuesto por Koehler y Mishra (2009), para una integración de las TIC en las aulas, basadas en la adquisición de conocimientos instrumentales, disciplinares y metodológicos. Esta base conceptual sobre la que se ha construido el instrumento, aportará una visión clara, específica y holística de la capacitación y alfabetización en IA del profesorado universitario que imparte formación a los futuros

maestros y maestras en el panorama actual nacional, abordando todos y cada uno de los aspectos anteriormente mencionados.

En cuanto a la validez del constructo, el cuestionario presenta unas propiedades psicométricas adecuadas en el panel de expertos y durante la fase de prueba piloto, existiendo una coherencia directa interna entre los ítems planteados y los objetivos que se pretendían medir en cada bloque, mostrando evidencias de niveles satisfactorios de fiabilidad ($\alpha = 0,873$), tal y como se ha recogido tras la medición del Alfa de Cronbach.

Por todo ello, se puede decir que el instrumento se considera comprensible, operativo y adecuado para los objetivos del estudio, siendo especialmente robustos los bloques relacionados con el uso y las competencias en IA. Es por ello, que se considera adecuado para su aplicación en el estudio principal. No obstante, la validación definitiva de la estructura interna se abordará en dicha fase, una vez se disponga de una muestra de mayor tamaño que permita realizar análisis factoriales más robustos y evaluar de forma concluyente la validez de constructo del instrumento.

Referencias

- Alastor, E, Guilén-Gámez, FD, & Ruiz-Palmero, J (2024). Competencia digital del futuro docente de Educación Infantil y Primaria: un estudio por comparaciones múltiples. *RELATEC*, 23(1), 9-24. <https://doi.org/10.17398/1695-288X.23.1.9>
- Alcántara, EEG (2026). *Competencias digitales docentes en la era de la inteligencia artificial*. En K Parra, JC Osuna, JC Castellanos, JR López (2026). Universidad, docencia y transformación digital: Debates contemporáneos en la era de la inteligencia artificial (pp. 40-60). Universidad Autónoma Baja California. <https://doi.org/10.61728/AE26001104>
- Almenara, J (2016). *La formación del profesorado en TIC: Modelo TPACK*. Madrid: Ibérica.
- Arenas, AC (2021). *Métodos mixtos de investigación*. Magisterio.
- Bisquerra, R (2009). *Metodología de la investigación educativa*. La Muralla.
- Berrios Zayas, BI (2026). *Competencias Docentes Para la Implementación de Inteligencia Artificial en la Educación*. [Tesis de doctorado, Universidad Nova Suroeste].
- Cabero, J, Roig-Vila, R., Mengual-Andrés, S (2018). Conocimientos tecnológicos, pedagógicos y disciplinares de los futuros docentes según el modelo TPACK. *Digital Education Rewiev*, 32, 73-84.
- Condori, HS, Quispe, GNV, & Mamani, AQ (2026). Relación entre dependencia a la inteligencia artificial, rendimiento académico y procrastinación académica en universitarios. *Revista Eugenio Espejo*, 20(1), 130-145. <https://doi.org/10.37135/ee.04.25.08>
- Escobar-Pérez, J, & Cuervo-Martínez, Á (2008). Validez de contenido y juicio de expertos: una aproximación a su utilización. *Avances en medición*, 6(1), 27-36.
- Espinoza, BOV, Guaylcel, JVG, ..., & Barahona, MVV (2026). Análisis del uso de inteligencia artificial en educación universitaria y su impacto en el aprendizaje estudiantil. *Revista Latinoamericana de Calidad Educativa*, 3(2), 338-347.
- Galindo Domínguez, H, Delgado de Frutos, N, Campo Carrasco, L, & Sainz de la Maza San José, M (2024). Uso de ChatGpt en educación superior. Un análisis en función del género, rendimiento académico, año y grado universitario del alumnado y degree. *REDU. Revista de Docencia Universitaria*, 22(2).
- García, SA, Díaz, IA, Reche, MPC, & Maldonado, JJV (2025). Relación entre las competencias digitales docentes y la aceptación y uso de la Inteligencia Artificial en el periodo de formación inicial docente. *Revista electrónica interuniversitaria de formación del profesorado: REIFOP*, 28(2), 1-17. <https://doi.org/10.6018/reifop.661551>
- Gil, JAP, Moscoso, SC, & Rodríguez, RM (2000). Validez de constructo: el uso de análisis factorial exploratorio-confirmatorio para obtener evidencias de validez. *Psicothema*, 12, 442-446.
- Giler Herrera, M. K., Guillen Vallejo, S. A., & Carvajal Torres, W. A. (2026). Impacto del uso de la IA generativa en la enseñanza universitaria: riesgo, oportunidades y competencias digitales docentes. *Ciencia Y Educación*, 7(3.1), 486 - 504. <https://doi.org/10.5281/zenodo.19542824>
- Hernández Rodríguez, PR, & Domínguez Marin, MJ (2025). Reimaginar la Pedagogía en la Inteligencia Artificial: Análisis del Marco de Competencias para Docentes de la UNESCO. *Ciencia Latina Revista Científica Multidisciplinar*, 9(5), 17313-17328. https://doi.org/10.37811/cl_rcm.v9i5.21247
- Islas, MS, & Rodríguez-Betanzos, A (2026). Exploración del uso de la inteligencia artificial en estudiantes universitarios. *Cuaderno de Pedagogía Universitaria*, 23, e713. <https://doi.org/10.29197/cpu.v23i.739>
- Jaramillo, RAC, Palacios, GMM, Palacios, IAM, & Zúñiga, PIV (2026). Impacto de la inteligencia artificial en el aprendizaje universitario. *Polo del Conocimiento*, 11(5), 2118-2137. <https://doi.org/10.23857/pc.v11i5.11719>
- Miao, F, & Cukurova, M (2024). *Marco de competencias para docentes en materia de IA*. UNESCO. <https://doi.org/10.54675/AQKZ9414>
- Mondragón Marín, PC, Flores-Pacheco, JA, & Sambola, DM (2025). Competencias digitales docentes y su influencia en la integración de Inteligencia Artificial en educación superior. *Wani*. <https://doi.org/10.5377/wani.v1i1.21250>

Montero Zamora, EG, Segura Cueva, KA, Montero Cadena, OG, & Montero Cadena, SB (2024). Evaluación de la validez y fiabilidad en estudios científicos: Revisión sistemática de métodos y buenas prácticas. *Ciencia Y Educación*, 365 - 387. <https://doi.org/10.5281/zenodo.14207509>

Ochoa, J, & Yunkor, Y (2019). El estudio descriptivo en la investigación científica. *ACTA JURÍDICA PERUANA*, 2(2).

Oviedo, HC, & Campo-Arias, A (2005). Aproximación al uso del coeficiente alfa de Cronbach. *Revista colombiana de psiquiatría*, 34(4), 572-580.

Pereira, MMB, & Idoiaga, N (2026). Percepciones negativas del alumnado sobre el uso de la IA en la escritura académica: implicaciones didácticas para la Educación Superior. *Educación XX1*, 29(1), 351-372. <https://doi.org/10.5944/educxx1.43943>

Pérez, ER, & Medrano, LA (2010). Análisis factorial exploratorio: bases conceptuales y metodológicas. *Revista Argentina de Ciencias del Comportamiento (RACC)*, 2(1), 58-66. <https://doi.org/10.32348/1852.4206.v2.n1.15924>

Rendón, OHP (2001). La matriz de congruencia: Una herramienta para realizar investigaciones sociales. *Economía y sociedad*, 6(10), 311-316.

Riquelme, I (2022). *Análisis de las competencias en docentes universitarios en las tecnologías de la información y la comunicación (TIC)*. [Tesis de doctorado, Universidad de Córdoba].

Rosa de Sáa, SDL (2012). *Análisis estadístico comparativo de tres escalas de valoración: Likert, fuzzy-Likert y fuzzy de respuesta libre* [Tesis de doctorado, Universidad de Oviedo].

Urrutia Egaña, M, Barrios Araya, S, Gutiérrez Núñez, M, & Mayorga Camus, M (2014). Métodos óptimos para determinar validez de contenido. *Educación médica superior*, 28(3), 547-558.

Vásquez Fernández, BJ (2025). *Desafíos educativos en la enseñanza superior para integrar la inteligencia artificial en procesos de alfabetización digital*. Yachana, 14(2). <https://doi.org/10.62325/10.62325/yachana.v14.n2.2025.968>

ANEXOS

Anexo 1: Cuestionario

Pregunta de consentimiento informado: ¿Acepta participar voluntariamente en el presente estudio y ha comprendido el tratamiento de los datos que proporcione para los fines anteriormente mencionados, en relación con la gestión y ejecución de la investigación? *

(Marca solo una opción)

- ☐ Sí, acepto participar en el estudio y consiento el tratamiento de los datos
- ☐ No acepto participar en el estudio ni consiento el tratamiento de los datos

Bloque 1. Datos sociodemográficos

P1. Género*

(Marca solo una opción)

- ☐ Hombre
- ☐ Mujer
- ☐ Otro / Prefiero no decirlo

P2. Edad*

(Marca solo una opción)

- ☐ Menos de 30 años
- ☐ De 30 a 40 años
- ☐ De 41 a 50 años

- ☐ De 51 a 60 años
- ☐ Más de 60 años

P3. Años de experiencia docente en la universidad*

(Marca solo una opción)

- ☐ Menos de 5 años
- ☐ De 5 a 10 años
- ☐ De 11 a 15 años
- ☐ Más de 15 años

P4. Carácter de la institución educativa*

(Marca solo una opción)

- ☐ Carácter público
- ☐ Carácter privado

P5. Área de especialización principal*

(Marca solo una opción)

- ☐ Métodos de investigación y diagnóstico en educación
- ☐ Teoría de la educación y pedagogía social
- ☐ Didáctica
- ☐ Psicología
- ☐ Otro (especificar)

Bloque 2. Conocimiento general y formación en Inteligencia Artificial

P6. ¿Cómo definiría la inteligencia artificial (IA)?*

(Marca solo una opción)

- ☐ Máquinas que imitan la inteligencia humana
- ☐ Programas que aprenden automáticamente de datos
- ☐ Robots autónomos
- ☐ No estoy seguro/a

P7. ¿Qué nivel de conocimiento cree tener sobre IA?*

(Escala Likert de 1 a 5: Nulo – Experto)

P8. ¿Conoce herramientas de IA como ChatGPT, Copilot, Gemini, etc.?*

(Marca solo una opción)

- ☐ Sí
- ☐ No

P9. ¿Con qué frecuencia usa este tipo de herramientas de IA?*

(Escala Likert de 1 a 5: Nunca – Siempre)

P10. Formación recibida en IA*

(Para cada ítem, marque una opción)

Escala de respuesta: Nada / Poca / Moderada / Bastante / Mucha / No aplica

- ☐ ¿Ha recibido formación específica en IA?
- ☐ ¿Ha recibido formación en IA aplicada a su práctica docente?
- ☐ ¿La formación recibida le ha resultado útil para aplicar IA en su docencia?

Bloque 3. Uso de la Inteligencia Artificial en la docencia universitaria

P11. ¿Ha utilizado IA para preparar sus clases?*

(Escala Likert de 1 a 5: Nunca – Siempre)

P12. Considero que el uso de la IA en la docencia es pedagógicamente adecuado*

(Escala Likert de 1 a 5: Totalmente en desacuerdo – Totalmente de acuerdo)

P13. ¿Recomendaría a sus colegas usar IA en su práctica docente?*

(Escala Likert de 1 a 5: Nunca – Siempre)

P14. ¿Con qué frecuencia ha utilizado la IA en los siguientes aspectos de su docencia?*

(Para cada ítem: Nunca / Rara vez / Ocasionalmente / Frecuentemente / Siempre)

- ☐ Diseño de actividades didácticas
- ☐ Generación de contenidos (ejercicios, casos prácticos)
- ☐ Corrección automática de trabajos
- ☐ Tutoría virtual (chatbots)

P15. ¿Con qué frecuencia encuentra las siguientes barreras para integrar la IA en su docencia?*

(Para cada ítem: Nunca / Rara vez / Ocasionalmente / Frecuentemente / Siempre)

- ☐ Falta de formación
- ☐ Falta de recursos tecnológicos
- ☐ Desconocimiento de herramientas

☐ Resistencia al cambio

P16. Considero que la IA tendrá un papel cada vez más relevante en la educación durante los próximos 5 años*

(Escala Likert de 1 a 5: Totalmente en desacuerdo – Totalmente de acuerdo)

Bloque 4. Políticas institucionales e impacto de la IA en educación

P17. ¿En su universidad existen planes formativos sobre IA para el profesorado?*

(Marca solo una opción)

- ☐ Sí, y son accesibles
- ☐ Sí, pero son limitados
- ☐ No, no existen
- ☐ No estoy informado/a

P18. ¿Existe un departamento o equipo que apoye en el uso de IA para docencia?*

(Marca solo una opción)

- ☐ Sí, lo conozco y he recurrido a él
- ☐ Sí, pero no lo he utilizado
- ☐ No, no existe
- ☐ No lo sé

P19. ¿Cree que su universidad debería invertir más en formación en IA para el profesorado?*

(Escala Likert de 1 a 5: Totalmente en desacuerdo – Totalmente de acuerdo)

P20. ¿Cree que la IA puede reemplazar parcialmente el rol del docente?*

(Escala Likert de 1 a 5: Totalmente en desacuerdo – Totalmente de acuerdo)

P21. ¿Con qué frecuencia encuentra los siguientes riesgos del uso de IA en educación?*

(Para cada ítem: Nunca / Rara vez / Ocasionalmente / Frecuentemente / Siempre)

- ☐ Pérdida de pensamiento crítico en estudiantes
- ☐ Dependencia excesiva de la tecnología
- ☐ Problemas de privacidad de datos
- ☐ Sesgos en los algoritmos
- ☐ Otros riesgos del uso de IA en educación (Respuesta abierta)

Bloque 5. Pensamiento crítico y ético del profesorado y alumnado

P22. ¿Incluye debates éticos sobre IA en sus clases?*

(Escala Likert de 1 a 5: Nunca – Siempre)

P23. ¿Con qué frecuencia fomenta en sus clases el análisis crítico sobre los límites de la IA?*

(Escala Likert de 1 a 5: Nunca – Siempre)

P24. ¿Con qué frecuencia ha observado en su alumnado un uso acrítico de la IA (p. ej., copiar respuestas sin análisis)?*

(Escala Likert de 1 a 5: Nunca – Siempre)

P25. Considero que mi alumnado tiene una buena capacidad para distinguir entre información generada por IA y contenidos elaborados por personas*

(Escala Likert de 1 a 5: Totalmente en desacuerdo – Totalmente de acuerdo)

P26. Indique su grado de acuerdo con las siguientes estrategias para desarrollar en el alumnado habilidades que la IA no pueda reemplazar*

(Para cada ítem: Totalmente en desacuerdo – Totalmente de acuerdo / No aplica)

- ☐ Proyectos colaborativos con componente humano
- ☐ Evaluación basada en procesos (no solo resultados)
- ☐ Diálogos socráticos y debates

Bloque 6. Conocimiento avanzado en Inteligencia Artificial**P27. ¿Puede explicar la diferencia entre Machine Learning y Deep Learning?***

(Marca solo una opción)

- ☐ Sí, con ejemplos concretos
- ☐ Sé la diferencia teórica, pero no cómo aplicarla
- ☐ No, los confundo
- ☐ No sé qué son

P28. ¿Conoce el marco TPACK para integrar IA en educación?*

(Escala Likert de 1 a 5: Nada – Mucho)

P29. ¿Con qué frecuencia ha usado APIs de IA (p. ej., OpenAI o Gemini) para fines pedagógicos?*

(Escala Likert de 1 a 5: Nada – Mucho)

P30. ¿Ha empleado IA para analizar datos cualitativos (p. ej., respuestas abiertas de estudiantes)?*

(Escala Likert de 1 a 5: Nunca – Siempre)

P31. ¿Ha usado IA para revisar literatura académica o análisis de datos cualitativos?*

(Escala Likert de 1 a 5: Nunca – Siempre)

P32. ¿Cómo implementaría un sistema de IA para adaptar ejercicios al ritmo de aprendizaje individual?*

(Marca solo una opción)

- ☐ Con modelos de recommender systems basados en desempeño previo
- ☐ Usando chatbots que ajusten dificultad en tiempo real
- ☐ No es posible sin violar la privacidad estudiantil
- ☐ No tengo los conocimientos técnicos
- ☐ Otro (Respuesta abierta)

P33. Grado de acuerdo con formas de adaptar contenidos mediante IA*

(Para cada ítem: Totalmente en desacuerdo – Totalmente de acuerdo)

- ☐ Utilizar sistemas de recomendación basados en learning analytics
- ☐ Generar versiones alternativas de los materiales
- ☐ Considero que esta aplicación no es posible

P34. Considero que podría diseñar una rúbrica automatizada con IA para evaluar el pensamiento crítico en ensayos*

(Escala Likert de 1 a 5: Totalmente en desacuerdo – Totalmente de acuerdo)

P35. Detección de favoritismos o desigualdades en sistemas de IA educativos*

(Para cada ítem: Totalmente en desacuerdo – Totalmente de acuerdo)

- ☐ Revisar si los datos de entrenamiento representan a todos los tipos de estudiantes
- ☐ Comparar si las recomendaciones funcionan igual para distintos grupos
- ☐ Considero que los sistemas de IA no presentan este tipo de problemas

P36. Criterios para evaluar la calidad de contenidos generados por IA*

(Para cada ítem: Totalmente en desacuerdo – Totalmente de acuerdo)

- ☐ Coherencia científica verificada mediante fuentes primarias
- ☐ Ausencia de sesgos discriminatorios
- ☐ Adaptación del contenido al nivel cognitivo del alumnado
- ☐ No evaluaría la calidad; confiaría en la IA

P37. ¿Qué estrategia emplearía para reducir alucinaciones en textos generados por IA?*

(Marca solo una opción)

- ☐ Usar prompts específicos con fuentes académicas

- ☐ Combinar IA con verificación humana cruzada
- ☐ Limitar su uso a contenidos no críticos
- ☐ No sé qué son “alucinaciones” en IA
- ☐ Otro (Respuesta abierta)

P38. ¿Qué protocolo seguiría al utilizar sistemas de IA que procesen datos del alumnado?*

(Marca solo una opción)

- ☐ Anonimización estricta y consentimiento informado
- ☐ Solo plataformas con cumplimiento del RGPD
- ☐ No usaría IA con datos sensibles
- ☐ No he considerado este problema
- ☐ Otro (Respuesta abierta)

P39. Si tuviera que formarse en un área técnica de IA, ¿cuál elegiría?*

(Marca solo una opción)

- ☐ Programación de modelos educativos
- ☐ Interpretabilidad de algoritmos (explicabilidad IA)
- ☐ Ética y buenas prácticas
- ☐ Ninguna, no me interesa la formación técnica
- ☐ Otro (Respuesta abierta)

P40. Habilidades técnicas que deberían priorizarse en un curso de IA para docentes*

(Para cada ítem: Totalmente en desacuerdo – Totalmente de acuerdo)

- ☐ Interpretar resultados de IA (explicabilidad)
- ☐ Diseñar prompts efectivos para pedagogía
- ☐ Programación básica para ajustar modelos
- ☐ Ninguna; el curso debería centrarse solo en aspectos didácticos