

Article

From AI Tool Use to Instructional Design: Development and Validation of the AID-CTQ in Higher Education

Natalia Lara Nieto-Márquez ^{1,*}, Rubén Madrigal-Cerezo ¹, Laura Ramos-Marcos ¹, Nicolás Rueda-Díaz ¹, Tomás García-Martín ² and Francisco López-Muñoz ^{3,4,5}

¹ Faculty of Education, Camilo José Cela University, Castillo de Alarcón 49, Villafranca del Castillo, 28692 Madrid, Spain

² Escuela Politécnica Superior de Tecnología y Ciencia, Camilo José Cela University, C/de Juan Hurtado de Mendoza, 4, 28036 Madrid, Spain

³ Faculty of Health Sciences–HM Hospitals, Camilo José Cela University, Castillo de Alarcón 49, Villafranca del Castillo, 28692 Madrid, Spain

⁴ HM Hospitals Health Research Institute, Plaza del Conde del Valle de Suchil 16, 28015 Madrid, Spain

⁵ Neuropsychopharmacology Unit, “Hospital 12 de Octubre” Research Institute, Avenida de Córdoba s/n, 28041 Madrid, Spain

* Correspondence: natalia.lara@ucjc.edu

Abstract

Artificial intelligence (AI) is transforming higher education, although most research addresses its integration in terms of frequency of use or technological acceptance, without examining how it translates into specific curricular and instructional decisions. That is why this study has a dual aim: to develop and validate the AI Instructional Design Questionnaire for Critical Thinking (AID-CTQ) and to analyze how university faculty integrate AI into instructional design practices in higher education. The sample included 144 faculty members from a university in Madrid, selected by convenience. Exploratory and confirmatory factor analyses of the questionnaire supported a three-factor structure: Activity Design (F1), Critical Thinking Assessment (F2), and Self-Regulation and Reflection (F3). The final 12-item model shows good model fit (CFI = 0.98, TLI = 0.98, RMSEA = 0.05, SRMR = 0.05) and adequate overall reliability ($\alpha = 0.86$). At the item level, responses related to assessment and reflective practices showed consistently high agreement, whereas items linked to activity design displayed greater variability. Faculty members with more than 10 years of experience obtained significantly higher scores, indicating that the educational value of AI depends less on the tools used and more on the quality of instructional decisions. Reported use of AI was high, with ChatGPT and Copilot being the most frequently used tools. Overall, the findings indicate that the integration of AI in higher education is evolving from predominantly instrumental uses toward more pedagogical and curriculum-oriented forms of implementation. Accordingly, the educational value of AI lies less in the tool itself than in the quality of the instructional decisions through which it is meaningfully embedded in the curriculum.

Keywords: artificial intelligence; higher education; instructional design; critical thinking; self-regulated learning; curriculum implementation



Academic Editors: Kelum Gamage,
Maithree Wickramasinghe and
Neil Gordon

Received: 30 April 2026

Revised: 15 June 2026

Accepted: 16 June 2026

Published: 20 June 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Artificial Intelligence (AI) is redefining human interaction and behavior in numerous ways, including how faculty members are incorporating it into their teaching practices and

lesson planning. Therefore, the quality of educational programs is also being shaped by the evolution of AI and its potential applications. One area of focus is how generative AI is being integrated into the university curriculum. In this context, the key question is no longer simply whether AI is being used in higher education, but how that use translates into specific curricular and pedagogical decisions.

As with other technological innovations, the educational value of AI depends not only on its features but also on the instructional approach used to incorporate it into the curriculum and teaching practices (Zawacki-Richter et al., 2019; Chen et al., 2020; Cabero-Almenara et al., 2024; Ng et al., 2025). Models such as TPACK, UTAUT, and SAMR have sought to establish this pedagogically focused connection between understanding technological tools and their integration into teaching activities and instructional design (Cabero-Almenara et al., 2024; Drugova et al., 2021; Hamilton et al., 2016; Jiménez-García et al., 2023). From this perspective, the key challenge in higher education is not whether to use AI but how it is being implemented at the curricular level. This curricular implementation involves combining the use of AI tools with decisions regarding learning outcomes, the design of tasks that require observable performance, consistent assessment of results, and feedback that supports students' process improvement.

Although a growing body of literature has examined AI adoption, acceptance, and digital competence among educators, far less attention has been paid to how AI use is translated into concrete instructional design decisions. Existing frameworks explain why teachers adopt AI, but they do not provide evidence regarding how faculty design activities, assessment processes, and reflective learning experiences supported by AI. Furthermore, validated instruments specifically designed to assess AI-mediated instructional design remain scarce. This study addresses this gap by examining faculty instructional design practices and validating a measurement instrument focused on curriculum implementation through AI.

1.1. Frameworks for Technology and AI Adoption: Contributions and Limitations

One of the challenges that has arisen with the emergence of AI and specifically, Generative AI (GenAI) tools, is how to implement them in an educational context. Among the most widely used models for analyzing these challenges are TPACK, UTAUT, and SAMR (Cabero-Almenara et al., 2024; Drugova et al., 2021; Hamilton et al., 2016; Jiménez-García et al., 2023). These models have contributed to our understanding of faculty needs and preparation, as well as the acceptance of technological tools and levels of integration in the educational context. Despite the existence of these models for implementing technology in the classroom by faculty, they do not include an instructional framework to guide faculty in the curricular implementation of AI.

Specifically, the TPACK model focuses on conceptualizing the pedagogical knowledge required to teach using technology based on three interconnected pillars: technological knowledge, pedagogical knowledge, and subject-matter knowledge (León Naranjo, 2024). However, this model does not propose a sequential curriculum design (Fragouli, 2025). Likewise, technology acceptance models such as UTAUT help explain technology adoption in terms of performance expectations, perceived effort, social influence, and facilitating conditions. The main focus of this model is thus on teachers' usage behavior (Cabero-Almenara et al., 2024). Other models, such as SAMR, have been highly influential in practice for categorizing levels of integration: substitution, augmentation, modification, and redefinition. However, it is necessary to consider the context in which it is implemented, as well as to use it without a rigid hierarchy and not as a guarantee of pedagogical improvement (Hamilton et al., 2016; Jiménez-García et al., 2023). On the other hand, adaptations of these models are beginning to emerge. These are based on strategies for training and hybrid implementation

of AI and professional literacy training for educators, as well as the need to establish clear guidelines for application or use and access to AI by these agents (Chan & Tang, 2024; Daher, 2025; Ng et al., 2025). Taken together, these models are useful for understanding technology adoption and levels of integration, but they are less precise when the aim is to analyze the alignment between AI use, task design, assessment, and feedback in curricular implementation (Cabero-Almenara et al., 2024; Drugova et al., 2021; Hamilton et al., 2016; Jiménez-García et al., 2023).

Regulatory frameworks exist at both European and Spanish levels that outline how educators can implement AI in the classroom to enhance student learning and the acquisition of skills. In Europe, the Digital Competence Framework for Educators (DigCompEdu) outlines levels for educators to progress in competencies related to the teaching and learning process, assessment and feedback, and digital content, among others (Redecker, 2017). In addition to plans and strategies related to teachers' digital competence, the European regulation sets out clear guidelines for educators regarding AI-based assessment processes, emphasizing the necessary training in critical thinking and media literacy (Regulation (EU) 2024/1689, 2024). Regarding the use of AI in Spain, in addition to adhering to these European frameworks, legislation addresses teaching practices and the competencies students must acquire. Thus, the National Institute of Educational Technologies and Teacher Training (Instituto Nacional de Tecnologías Educativas y de Formación del profesorado, INTEF) has published guidelines and resources for the educational use of AI, with the aim of clarifying conceptual understanding, proposing preventive strategies, and addressing ethical and data protection considerations. However, its focus is on pre-university levels (INTEF, 2024). It is evident in compulsory education, where it is stated that students must acquire a set of digital competencies, which are aligned with the use of GenAI in the creation of digital content or the development of technological solutions from an ethical standpoint, among others (Real Decreto 217/2022, 2022).

1.2. Instructional Design and AI Integration

The frameworks reviewed above explain why and to what extent faculty adopt technology, but they do not specify how AI is translated into concrete design decisions. From the perspective of instructional design, the integration of AI should be approached as a pedagogical and curricular decision that affects the connection between objectives, tasks, assessment, and feedback. Thus, instructional design is understood as a structured process for planning effective, coherent, and assessable learning activities and experiences. In this regard, the ADDIE model (Analysis, Design, Development, Implementation, Evaluation) serves as a guide for conceptualizing design as a structured and iterative cycle (Branch, 2009; Abuhassna et al., 2024). When applied to AI, this approach shifts the focus from the tool to the educational function it fulfills within a broader pedagogical sequence. This work involves, first, clearly defining which learning outcomes are intended to be developed; second, designing tasks that require observable cognitive performance, such as searching for information, comparing perspectives, analyzing evidence, or arguing; third, establishing assessment procedures that evaluate reasoning and not just the final product; and, finally, supporting the process with feedback, self-regulation, and ethical reflection. Recent work in higher education also suggests that the educational value of AI depends on how it is embedded within task design, assessment, and feedback structures, particularly when institutions seek to move beyond efficiency-oriented uses toward more meaningful curricular integration (Xia et al., 2024; Liang et al., 2025; Rutecka et al., 2025; Sapawi & Yusoff, 2025).

1.3. Curriculum Implementation and AI Integration

While instructional design concerns the planning of coherent and assessable learning experiences, curriculum implementation refers to how those designs are enacted within authentic teaching contexts. It encompasses the decisions teachers make when delivering planned tasks, mediating students' use of AI, and adjusting assessment and feedback in practice, that is, the point at which a design becomes an observable teaching–learning process. The two dimensions are connected through the principle of constructive alignment (Biggs & Tang, 2011), which requires that intended learning outcomes, AI-supported tasks, and assessment criteria remain mutually consistent throughout enactment, and not only at the planning stage. Recent work in higher education emphasizes that the educational value of AI is determined less by the tools themselves than by the degree to which their use remains aligned with intended learning outcomes, task demands, and assessment criteria throughout enactment, and not only at the planning stage (Xia et al., 2024; Liang et al., 2025; Rutecka et al., 2025). From this standpoint, AI integration becomes pedagogically meaningful only when the use of these tools preserves the coherence between what students are expected to learn, what they are asked to do, and how their reasoning is evaluated, which in turn calls for innovative curricular models capable of accommodating AI-mediated tasks (Chu & Ashraf, 2025; Sapawi & Yusoff, 2025). Consequently, the analytical focus shifts from whether AI is adopted to how AI reshapes the enactment of tasks, feedback, and assessment at the curricular level, an aspect that the adoption and acceptance frameworks discussed in Section 1.1 do not fully capture and that motivates the instrument developed in the present study.

1.4. AI Tools Most Commonly Used by Teachers

When it comes to the use of AI tools, one of the earliest studies conducted was the Carrington Wheel study, which identified the most widely used and popular tools for each educational activity based on Bloom's Taxonomy and the SAMR model of technology integration (Jiménez-García et al., 2023). Recent studies on AI use in higher education suggest that faculty members tend to rely primarily on general-purpose generative tools and information-support applications, while more specialized tools are adopted less consistently and often for specific tasks such as automated feedback, assessment support, or content generation (Perezchica-Vega et al., 2024; Ojeda et al., 2023; Karataş et al., 2024; Ng et al., 2025). However, AI is evolving on a monthly, and even weekly, basis, and today we can highlight ChatGPT, Grammarly, Utah Compose, and Gradescope among the most prominent in the scientific literature (García-López et al., 2025; Hansel et al., 2024; Wei et al., 2023; Xia et al., 2024). These tools enable the creation of materials and activities or the provision of feedback and assessments that yield outstanding results when accompanied by appropriate teacher supervision.

One of the most ambitious challenges has been identified in the creation of digital materials using generative AI (GenAI) tools. These tools enable the efficient incorporation of a range of elements of differentiation, creativity, personalization, and inclusion, among others, which can enrich the teaching–learning process (Ejjami, 2024; Sapawi & Yusoff, 2025). However, it is necessary to consider a series of cross-cutting elements such as critical thinking, self-regulation, and reflection so that the creation of these materials or their use in the classroom incorporates all the characteristics mentioned (Chun et al., 2025; Jiménez-García et al., 2023).

At present, it is necessary to highlight ChatGPT once again as a tool that enables self-regulation in learning (Abdelhalim, 2024); GenAI applications developed in-house by educational institutions for fostering critical thinking (Tsopra et al., 2023) or more specialized

tools in addition to ChatGPT, such as EduAid or Magic School, designed for creating materials and personalizing learning (Cabañuz & García-García, 2024; ElSayary et al., 2025).

In all of the studies mentioned previously, ChatGPT is the most prominent tool in all searches conducted regarding its use by educators across the various scientific search engines consulted. Similarly, studies such as those by Karataş et al. (2024) and Ng et al. (2025) also identify ChatGPT as the tool most frequently used by teachers.

1.5. Faculty Perceptions, AI Adoption, and Institutional Policies

At the institutional level, policies developed regarding the use of AI focus on establishing frameworks for use and action, dividing these responsibilities between faculty and students to promote ethical, critical, and responsible use centered on improving the teaching–learning process (INTEF, 2024; Fengchun & Cukurova, 2025). The report by the Organization for Economic Cooperation and Development (OECD) indicates that 35% of Spanish secondary school teachers (a percentage similar to the European average) use AI in their teaching practices (OECD, 2025). In other countries, such as Chile or Australia, the figure exceeds 50%. Beyond Europe, studies conducted in Australia, North America, Latin America, and Asia have similarly reported increasing levels of AI adoption among higher education faculty, although significant differences remain regarding training opportunities, institutional support, and pedagogical integration strategies (Zawacki-Richter et al., 2019; Abbasi et al., 2024; Perezchica-Vega et al., 2024; Liang et al., 2025; Sapawi & Yusoff, 2025).

There are several scenarios in which AI is applied in the educational field, as we can observe its use across all areas of knowledge. Thus, educators' perceptions regarding its use indicate a high willingness to adopt AI, a more positive attitude toward it when teachers possess strong digital competence, a perception of reduced workload when it is used, and clear evidence of the need for ongoing teacher training (Batubara et al., 2025; Espejo Aubá, 2024; Machado et al., 2025). We are facing a process of adaptation to the use of AI or GenAI in the education system, which is undergoing constant change and adaptation, now driven by advancements and the application of AI.

Building on the gap outlined above, the general objective of this study is to analyze the instructional design of university faculty, as well as the integration and use of AI tools by faculty for the instructional design of teaching activities that foster processes of critical thinking, self-regulation, and reflection. The following specific objectives of the study are established:

- Design, develop, and validate an ad hoc questionnaire to study the instructional design approach in teaching activities. Analyze how faculty integrate AI into instructional design and curriculum implementation.
- Contextualize the level of AI adoption in university teaching, identifying which tools are used and for what purpose. As a methodological objective supporting this purpose, develop and validate the AI Instructional Design for Critical Thinking Questionnaire (AID-CTQ).
- Describe the use of AI in the creation of academic activities according to faculty members' years of teaching experience and their fields of expertise. Describe faculty AI use, identifying which tools are employed and for what instructional purposes.
- Examine differences in AI-mediated instructional design according to faculty teaching experience and disciplinary area.

2. Materials and Methods

2.1. Research Design

This study employs a mixed-methods design (quantitative and qualitative) that combines the development and validation of an instrument with descriptive and comparative

analysis. The study is conducted within a single private institution in the Madrid region, where the primary objective is to analyze teachers' instructional design practices, while the development and psychometric validation of the AID-CTQ serve as a methodological objective that supports this analysis. The design integrates exploratory and confirmatory factor analyses, reliability tests, descriptive statistics, nonparametric comparisons between groups, and a complementary qualitative thematic analysis of the open-ended responses.

2.2. Participants

The sample group consists of 144 faculty members from a private university in Madrid, selected through non-probabilistic convenience sampling. These faculty members teach at the undergraduate level (42.36%; $N = 61$), the master's level (6.94%; $N = 10$), or both (50.70%; $N = 73$) across the various schools that make up the university. Regarding demographic characteristics, 53.47% were women ($N = 77$) and 46.53% were men ($N = 67$). The teaching experience reported by the participants ranges from 1 to 50 years. A total of 88.90% of participants ($N = 128$) indicated having prior experience with AI tools, while the remaining 11.10% ($N = 16$) had not used them.

2.3. Instrument of Measurement

The integration of AI into higher education cannot be considered as a mere adoption of technology, but rather as a curricular and instructional decision that reshapes what is considered evidence of learning and how tasks, assessment, and feedback are designed. To analyze how faculty incorporate the use of AI into instructional design decisions focused on critical thinking, a faculty self-report questionnaire was developed, structured around three dimensions: (1) Activity Design, (2) Assessment of Critical Thinking, and (3) Promotion of Self-Regulation and Reflection. This questionnaire is grounded in the frameworks and reference models presented in the introduction and supported by studies such as that of [Chu and Ashraf \(2025\)](#), which highlights the importance of creating innovative curriculum models. Thus, the instrument named the "AI Instructional Design for Critical Thinking Questionnaire" [AID-CTQ] was developed to measure how university professors integrate AI tools into instructional design decisions aimed at creating curricular activities that help foster critical thinking and students' self-regulated learning processes.

Furthermore, these dimensions emerge from the definition of critical thinking as a reasoned and self-regulated process involving skills such as interpretation, analysis, evaluation, and inference, among others, as well as associated dispositions. In other words, self-regulation is considered a constitutive component of critical thinking ([Castillejos López, 2022](#); [Fragouli, 2025](#); [Karataş et al., 2024](#); [Sapawi & Yusoff, 2025](#)). In the presence of generative AI, self-regulation and ethical reflection are prerequisites for quality: it is not enough to obtain a result; one must critically evaluate its reliability, recognize biases (both personal and systemic), and consider ethical and social implications. UNESCO guidelines ([Fengchun & Cukurova, 2025](#)) and competency frameworks for teachers ([Redecker, 2017](#); [INTEF, 2024](#)) emphasize these ethical and human agency dimensions for informed and responsible use.

The development of the questionnaire followed a systematic design process, beginning with a literature review, consultation with experts, and a review of previous experiences with the adoption of educational technologies to define the relevant constructs associated with the use of AI in university courses and the work of university faculty. Thus, an initial pool of items was developed to cover the three proposed dimensions (Appendix A, Table A1). Subsequently, the content was reviewed by a panel of three experts to ensure the validity and clarity of the instrument's content, as well as the comprehensibility of the instructions. Following these reviews, the items were rewritten to improve the clarity and

consistency of the wording, while preserving the underlying constructs and the distribution of items by dimension. Additionally, the initial proposal of five open-ended questions was simplified to two to prioritize contextual information about the AI tools used and general comments.

The end result is a faculty self-assessment tool (Appendix A, Table A2) that evaluates faculty instructional design decisions aimed at integrating AI into the university curriculum to foster critical thinking among students. It consists of 13 Likert-scale items distributed across three theoretical dimensions. Each item is rated on a scale of 1 to 5, where 1 corresponds to “Strongly Disagree” and 5 to “Strongly Agree”. To complement this information, an initial section contextualizing the faculty sample is included, collecting the following information: gender, years of teaching experience, the level at which teaching is conducted at the university (undergraduate or graduate), and field of study. Two open-ended questions were added to contextualize the use of AI or GenAI (tools used and purpose). As mentioned, the questionnaire is structured around three dimensions, aligned with the pedagogical processes to be analyzed in the design of academic activities in the AI era, based on the following theoretical constructs:

- **Dimension 1 (Activity Design):** Activity design by faculty (4 items). This dimension captures decisions regarding the design of curricular teaching activities. This dimension evaluates teaching design decisions that aim to activate processes characteristic of critical thinking in academic activities. This aligns with a constructivist pedagogical foundation (Fragouli, 2025; Kotsis, 2025), positing that learning outcomes, activities, and assessment must be designed coherently to promote meaningful learning. Furthermore, this alignment in the context of AI use requires that the task also promote cognitive processes (evaluating evidence, arguing, or justifying) (Abbasnejad et al., 2025; Castillejos López, 2022).
- **Dimension 2 (Critical Thinking Assessment):** Assessment of critical thinking by educators (4 items; 3 items in the final version). This dimension encompasses assessment practices that evaluate reasoning processes, provide feedback to improve analysis, utilize open-ended questions, and employ authentic assessment through case studies or problems. It is designed to describe assessment design decisions that make critical thinking visible in academic tasks (Kotsis, 2025). In formative assessment using AI, the focus must be on evidence of the processes carried out (criteria used, verification, justification) and not solely on the final product. In line with this, recent reviews on assessment in higher education using generative AI emphasize that assessment must be transformed to foster self-regulation, responsible learning, and integrity, which reinforces the importance of operationalizing assessment practices as a specific component of instructional design (Xia et al., 2024).
- **Dimension 3 (Self-Regulation & Reflection):** Fostering self-regulation and reflection in activities designed by educators (5 items). This dimension encompasses design and facilitation practices aimed at promoting the reassessment of decisions in light of new evidence, the identification of biases, ethical and social reflection, and intellectual autonomy. Thus, it arises from the perspective of using scaffolding through instructional design that incorporates supports to sustain more complex cognitive processes while students develop these competencies (Jiménez-García et al., 2023). This dimension is also grounded in research on metacognition and self-regulated learning, which emphasizes planning, monitoring, reflection, and adjustment as core processes in complex academic reasoning, particularly in environments mediated by digital and AI-supported tools (Drigas & Mitsea, 2020; Loksa et al., 2022; Castillejos López, 2022). In AI-mediated learning contexts, these processes are closely connected to ethical reflection, bias awareness, and responsible judgment, all of which are increasingly

recognized in recent work on assessment and educational AI governance (Xia et al., 2024; Flores-Vivar & García-Peñalvo, 2023; Fengchun & Cukurova, 2025).

2.4. Procedure

This study was conducted during the 2024–2025 academic year as part of the RUCAIE project, within the WISE-III university consortium (to explore the implementation of artificial intelligence in a university setting. To this end, an ad hoc measurement instrument was developed and sent to faculty members via various internal communication channels, along with the study's objectives, assuring them that their responses would be voluntary, anonymous, and confidential. The instrument was administered online using a survey tool provided by the institution. After giving their informed consent, participants accessed the form and answered the set of items. To minimize bias, it was emphasized that there were no right or wrong answers and that the information would be used solely for educational research purposes. Once the questionnaire distribution phase was completed in the second semester, the responses were exported to Excel for data curation and analysis.

2.5. Data Analysis

To assess the validity of the instrument, we began by conducting Exploratory Factor Analysis (EFA) and subsequently Confirmatory Factor Analysis (CFA) to validate the instrument's structure. EFA identified three clearly distinct factors that were confirmed by CFA, as shown in Section 3 of the results. Cronbach's alpha was used for the reliability analysis. The final version of the questionnaire based on these analyses is presented in Appendix A, Table A3. A self-report questionnaire was selected as the most efficient means of capturing faculty instructional design decisions across a heterogeneous sample. The EFA was conducted using principal axis factoring with promax (oblique) rotation, given the expected correlation among the theoretical dimensions, and the number of factors was determined by combining the Kaiser criterion, the scree plot, and theoretical interpretability. The CFA then tested the hypothesized three-factor structure using maximum likelihood estimation and standard goodness-of-fit indices (CFI, TLI, IFI, RMSEA, SRMR).

Descriptive statistics for the questionnaire items were calculated, as well as comparisons by years of teaching experience and fields of knowledge, to determine whether differences existed. The open-ended questions regarding AI tools and their uses were analyzed using a deductive-inductive thematic coding approach, taking into account that a single response could include several simultaneous uses of AI. All statistical analyses were performed using the specialized statistical software JASP (version 0.16.3). For the analysis of qualitative data, the specialized statistical software ATLAS.TI (version 23) was used. Additionally, support tools such as Excel were used to create the graphs, and DeepL Translator (free web version) was used to assist with and review the translation.

3. Results

3.1. Validation of the AID-CTQ Measurement Instrument

3.1.1. Exploratory Factor Analysis of the AID-CTQ

Exploratory factor analysis (EFA) was conducted to examine the underlying dimensional structure of the AI Instructional Design Questionnaire for Critical Thinking (AID-CTQ). The overall KMO coefficient obtained was 0.84. The item-level MSA values ranged from 0.71 to 0.91, supporting the suitability for factor extraction. Bartlett's sphericity test was statistically significant, $\chi^2(78) = 863.755$, $p < 0.001$, indicating sufficient correlations among the items to apply factor analysis. For factor extraction, the minimum residual method with oblique rotation (Promax) was applied, given that a correlation between the instrument's dimensions was theoretically expected (Table 1).

Table 1. Factor Loadings and Kaiser-Meyer-Olkin test by items.

Item	Factor 1 (F1)	Factor 2 (F2)	Factor 3 (F3)	Uniqueness	MSA
Q1	0.62			0.65	0.81
Q2	0.83			0.21	0.81
Q3	0.83			0.27	0.82
Q4	0.81			0.33	0.81
Q5		0.73		0.54	0.83
Q6		0.51		0.59	0.84
Q7		0.68		0.60	0.85
Q8		0.46		0.75	0.71
Q9			0.59	0.48	0.89
Q10			0.63	0.40	0.87
Q11			0.88	0.35	0.87
Q12			0.98	0.23	0.86
Q13			0.57	0.41	0.91

Applied rotation method is promax.

As a result, the items were grouped into the three factors corresponding to the theoretical dimensions: F1—Activity Design, F2—Critical Thinking Assessment, F3—Self-Regulation and Reflection. Items Q1–Q4 loaded onto F1 with values ranging from 0.62 to 0.83, confirming the consistency of the instructional activity design dimension. Items Q5–Q8 loaded onto F2 with loadings ranging from 0.46 to 0.73, representing the dimension of critical thinking assessment. Although Q8 had a comparatively lower loading and higher uniqueness it remained above the minimum threshold of 0.40. Items Q9–Q13 loaded onto F3 with loadings ranging from 0.57 to 0.98, indicating strong saturation of the self-regulation construct. The three-factor solution explains 55.4% of the total variance (Table 2).

Table 2. Factor characteristics and explained variance of the exploratory factor analysis.

	Unrotated Solution			Rotated Solution		
	SumSq. Loadings	Proportion Var.	Cumulative	SumSq. Loadings	Proportion Var.	Cumulative
Factor 1	4.64	0.36	0.36	2.92	0.22	0.22
Factor 2	1.87	0.14	0.50	2.49	0.19	0.42
Factor 3	0.70	0.05	0.55	1.80	0.14	0.55

The factor correlation matrix revealed moderate associations between the dimensions (ranging from 0.32 to 0.66), which supports the theoretical relationship among the constructs and confirms their empirical distinction (Table 3). Based on these results, the study proceeded with confirmatory factor analysis (CFA).

Table 3. Factor correlations in the exploratory factor analysis of the AID-CTQ.

	Factor 1	Factor 2	Factor 3
Factor 1	1	0.32	0.66
Factor 2	0.32	1	0.33
Factor 3	0.66	0.33	1

3.1.2. Confirmatory Factor Analysis of the AID-CTQ

Continuing with the confirmatory factor analysis (CFA), it should be noted that the initial model specified 13 items organized into three correlated latent factors corresponding to the structure identified in the EFA (Appendix A, Table A2). The initial 13-item model showed an acceptable but not optimal fit. The analyses were performed using the robust

ML estimator. Although some indices were close to the recommended cutoff points, the RMSEA above 0.08 and the TLI below 0.95 indicated localized areas of misfit. A joint review of the modification indices and the AFE results revealed that item Q8 had a comparatively lower loading (0.46), high uniqueness (0.75), and a limited contribution to the factor's internal consistency. From a statistical and theoretical perspective, the item was removed, resulting in a substantial improvement in model fit. Following its removal, the refined 12-item model showed a substantial improvement in overall fit, comfortably meeting the recommended cutoff criteria ($CFI/TLI \geq 0.95$; $RMSEA \leq 0.06$; $SRMR \leq 0.08$). The non-significance of the χ^2 statistic in the final model supports adequate overall fit. However, given the sensitivity of the χ^2 statistic to sample size, the model evaluation was based primarily on incremental and absolute fit indices. This objective improvement, together with the preserved conceptual coherence across the three theoretical dimensions, supports the decision to refine the instrument and strengthens the structural validity of the AID-CTQ (Table 4).

Table 4. Goodness-of-fit indices for the 13-item and 12-item CFA models.

Index	13-Items	12-Items
CFI	0.94	0.98
TLI	0.92	0.98
RMSEA	0.08	0.05
SRMR	0.06	0.05
$\chi^2 p$	0.001	0.14

The CFA of the final model, consisting of 12 items and three factors (Appendix A, Table A3), yields the following indices: $\chi^2(51) = 62.029$, $p = 0.139$; $CFI = 0.98$; $TLI = 0.98$; $IFI = 0.98$; $RMSEA = 0.05$ (90% CI: 0.000–0.085; p -close = 0.51); $SRMR = 0.05$.

Furthermore, all standardized factor loadings were statistically significant ($p < 0.001$) and ranged from moderate to strong values (Table 5). The items associated with F1 and F3 showed strong loadings, while F2 showed acceptable moderate loadings. As for the multiple correlations (R^2), they indicated that a substantial proportion of the variance in each item was explained by its corresponding construct. Thus, the factor structure obtained in the CFA confirmed the assignment of the items to the three factors identified in the EFA, which supports the validity in the study population. The standardized CFA model is presented in Figure 1.

Table 5. Standardized factor loadings and explained variance (R^2) for the final 12-item AID-CTQ model. F1—Activity Design, F2—Critical Thinking Assessment, F3—Self-Regulation and Reflection.

Factor	Indicator	Estimate	Std. Error	z-Value	p	Lower	Upper	R^2
Factor 1	Q2	1.06	0.11	10.05	<0.001	0.86	1.27	0.75
	Q3	1.01	0.09	10.95	<0.001	0.83	1.19	0.74
	Q4	0.91	0.10	9.13	<0.001	0.71	1.10	0.58
Factor 2	Q1	0.58	0.11	5.31	<0.001	0.37	0.79	0.40
	Q5	0.45	0.07	6.20	<0.001	0.31	0.59	0.45
	Q6	0.61	0.12	5.24	<0.001	0.38	0.84	0.62
Factor 3	Q7	0.64	0.13	4.82	<0.001	0.38	0.90	0.43
	Q9	0.51	0.06	8.02	<0.001	0.38	0.63	0.57
	Q10	0.70	0.08	8.83	<0.001	0.55	0.86	0.67
	Q11	0.68	0.08	9.03	<0.001	0.54	0.83	0.59
	Q12	0.65	0.08	8.66	<0.001	0.51	0.80	0.68
	Q13	0.59	0.06	9.90	<0.001	0.48	0.71	0.67

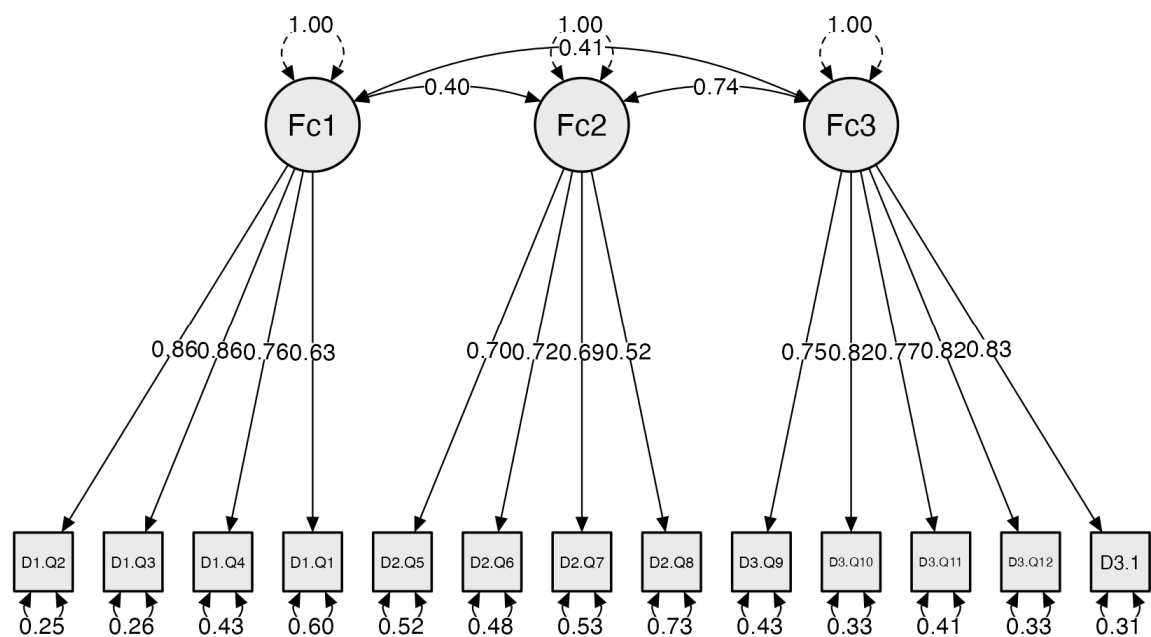


Figure 1. Standardized confirmatory factor analysis model of the final 12-item AID-CTQ.

3.1.3. Reliability Analysis of the AID-CTQ

Internal consistency reliability was analyzed using Cronbach’s alpha. For the overall AID-CTQ, reliability was good ($\alpha = 0.86$, 95% CI [0.82–0.89]), with an average inter-item correlation of 0.32, indicating adequate internal consistency without redundancy among items (Table 6). For the specific dimensions:

- Factor 1 (Activity Design) showed good reliability ($\alpha = 0.86$, 95% CI [0.81–0.89]). The correlations between the items and the rest of the scale ranged from 0.54 to 0.77, indicating strong internal consistency.
- Factor 2 (Critical Thinking Assessment) showed moderate reliability ($\alpha = 0.67$, 95% CI [0.56–0.75]). Although lower than that of the other dimensions, the coefficient remains within acceptable limits for short scales in applied educational research contexts.
- Factor 3 (Self-regulation and reflection) exhibits high reliability ($\alpha = 0.88$, 95% CI [0.84–0.90]), with correlations between items and the rest ranging from 0.65 to 0.77.

Table 6. Internal consistency of the final AID-CTQ model.

Dimension	No. of Items	Cronbach’s α	95% CI
Factor 1—Activity Design	4	0.86	[0.81–0.89]
Factor 2—Critical Thinking Assessment	3	0.67	[0.56–0.75]
Factor 3—Self-Regulation & Reflection	5	0.88	[0.84–0.90]
Total Scale (12 items)	12	0.86	[0.82–0.89]

3.2. Results of the Descriptive Analysis of the AID-CTQ Results

When analyzing the descriptive statistics for the 12 items in the final questionnaire, the following means, standard deviations, and score ranges were obtained, as shown in Table 7. Table 7 also presents the full wording of the final AID-CTQ items, grouped by their three dimensions, so that the instrument can be read directly within the main text; the complete bilingual questionnaire is additionally provided in Appendix A.

Table 7. Final 12-item AID-CTQ instrument, grouped by dimension, with item statements and descriptive statistics. Items are rated on a five-point Likert scale (1 = strongly disagree to 5 = strongly agree). Item Q8 was removed following validation. The full bilingual questionnaire, including earlier versions, is provided in Appendix A (Table A3).

Item	Statement	Mean	SD	Minimum	Maximum
Activity Design (F1)					
Q1	The activities I design using AI encourage students to seek out and critically reflect on information.	4.08	0.97	1	5
Q2	I often incorporate AI-based exercises that require students to compare different perspectives on a topic.	3.51	1.26	1	5
Q3	My activities promote the analysis of the topics studied using AI.	3.59	1.26	1	5
Q4	I encourage discussion and debate so that students practice evidence-based argumentation using AI.	3.57	1.25	1	5
Critical Thinking Assessment (F2)					
Q5	I assess students not only on their answers but also on the reasoning process they use.	4.62	0.66	3	5
Q6	I provide feedback that helps students improve their analytical skills.	4.53	0.76	1	5
Q7	I use open-ended questions (on exams or assignments) to have students demonstrate inference and critical judgment.	4.40	0.91	1	5
Self-Regulation & Reflection (F3)					
Q9	I encourage students to review and adjust their decisions from different perspectives.	4.37	0.71	2	5
Q10	I encourage students to identify their own biases and work to ensure they do not limit their performance.	4.24	0.92	1	5
Q11	I create activities that promote reflection on the ethical and social implications of decisions.	4.12	0.97	1	5
Q12	I facilitate self-reflection so students identify areas for improvement in their thought processes.	4.26	0.85	1	5
Q13	I promote intellectual autonomy, encouraging students to question information and make informed decisions.	4.50	0.70	3	5

As shown in the table, the means ranged from 3.51 to 4.62 on the five-point Likert scale, indicating a high level of agreement among faculty regarding AI-mediated instructional design practices aimed at fostering critical thinking. The item with the highest mean was Q5 (*I evaluate students not only on their answers but also on the reasoning process they use*) ($M = 4.62$; $SD = 0.66$), suggesting a strong consolidation of practices related to the assessment of critical thinking. In contrast, item Q2 (*I usually incorporate AI-based exercises into my course that require students to compare different viewpoints on a topic*) had the lowest mean ($M = 3.51$; $SD = 1.26$), reflecting greater heterogeneity in the implementation of this practice. Standard deviations ranged from 0.66 to 1.26, indicating moderate dispersion of responses. Items Q2 and Q3 showed greater variability, which could be interpreted as greater diversity in teaching approaches regarding the practices analyzed in these items.

To investigate the use of AI in instructional design, the analysis was expanded upon based on the contextualization of the sample to determine whether faculty experience or the faculty members' field of knowledge had any effect on its use.

3.2.1. Analysis of Results by Faculty Experience

To explore potential differences in AI-mediated instructional design practices based on teaching experience, three ad hoc groups were formed: G1 (<10 years) $N = 52$; G2

(10–20 years) N = 47; G3 (>20 years) N = 45. The results of this analysis are shown in Figure 2.

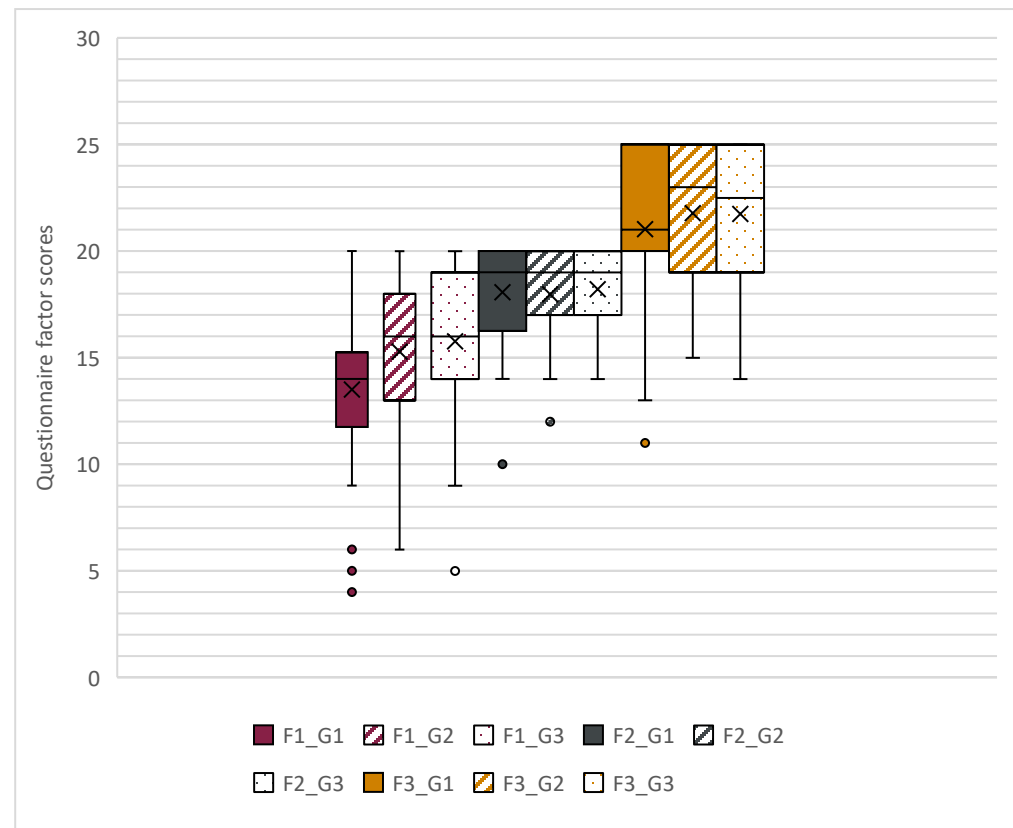


Figure 2. Distribution of AID-CTQ factor scores by teaching experience group. F1—Activity Design, F2—Critical Thinking Assessment, F3—Self-Regulation and Reflection.

As shown in F1 (Activity Design), a gradual increase is observed as teaching experience increases. Thus, the group with the most experience (G3, >20 years) has the highest mean score ($M = 15.77$; $SD = 3.78$) and greater homogeneity, as the standard deviation is slightly lower (G1: $M = 13.11$; $SD = 4.36$ /G2: $M = 14.86$; $SD = 4.50$). This would suggest a greater consolidation of instructional design strategies. In F2 (Critical Thinking Assessment), the differences between groups are minimal (G1: $M = 13.49$; $SD = 1.97$ /G2: $M = 13.51$; $SD = 1.85$ /G3: $M = 13.65$; $SD = 1.63$), indicating that assessment practices related to critical thinking exhibit cross-sectional stability. Finally, F3 (Self-regulation and reflection) shows moderate differences across groups based on years of teaching experience (G1: $M = 21.02$; $SD = 3.68$ /G2: $M = 21.79$; $SD = 3.21$ /G3: $M = 21.75$; $SD = 3.37$). The group with 10–20 years of experience stands out, with a mean slightly higher than G3. High scores on this factor indicate a well-established integration of practices that foster critical reflection and self-regulation through the use of AI.

Since the normality tests (Shapiro–Wilk) indicated significant deviations across the three groups of teaching experience, the nonparametric Kruskal–Wallis test was used to analyze differences between these groups and delve deeper into the analysis. Thus, in F1, the analysis revealed statistically significant differences between the groups ($H(2) = 8.034$, $p = 0.018$), with a medium effect size ($\epsilon^2 \approx 0.06$). This pattern was consistent with the parametric ANOVA: $F(2, 103) = 3.575$, $p = 0.032$, $\eta^2 = 0.065$. This test was supplemented with post hoc analysis using the Dunn test with Holm correction (Table 8).

Table 8. Dunn’s post hoc comparisons by teaching experience for Factor 1 (Activity Design).

Comparison	z	W _i	W _j	p	p _{bonf}	p _{holm}
1–2	−1.97	42.00	56.28	0.025 *	0.074	0.049 *
1–3	−2.75	42.00	62.14	0.003 **	0.009 **	0.009 **
2–3	−0.81	56.29	62.14	0.210	0.629	0.210

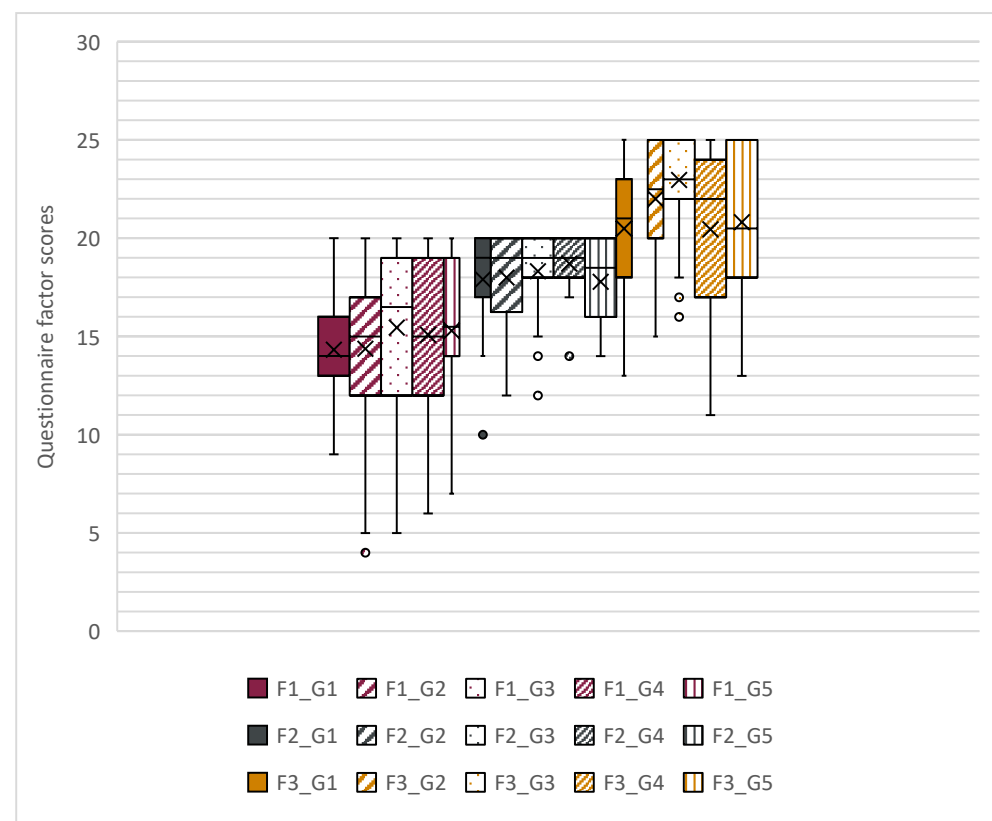
* $p < 0.05$, ** $p < 0.01$.

Significant differences were found between G1 and G2, as well as between G1 and G3. The effect size was greater in the comparison between teachers with less than 10 years of experience (G1) and teachers with more than 20 years of experience (G3). This suggests that teaching experience exceeding 10 years favors the design of activities that promote critical thinking through the implementation of AI, possibly due to greater pedagogical mastery or the progressive integration of active methodologies.

The analyses of F2 and F3 did not reveal statistically significant differences; therefore, the level of teaching experience does not influence the development of these areas (critical thinking, self-regulation, and reflection) through the activities proposed by teachers.

3.2.2. Analysis of Results by Field of Knowledge

To explore whether the field of study influences the instructional design of AI-based activities, four ad hoc groups were created corresponding to the academic fields of the participating faculties, as well as an additional group for those interdisciplinary profiles that could be classified under more than one faculty. G1 (Health Sciences), N = 32; G2 (Education), N = 30; G3 (Communication and Humanities), N = 32; G4 (Technology and Science), N = 15; G5 (Interdisciplinary), N = 35. The results of the descriptive analysis are shown in Figure 3.

**Figure 3.** Distribution of AID-CTQ factor scores by field of knowledge. F1—Activity Design, F2—Critical Thinking Assessment, F3—Self-Regulation and Reflection.

Moderate differences were observed in the mean scores across groups in F1, by area: Health Sciences: $M = 14.32$, $SD = 2.80$; Education: $M = 14.38$, $SD = 4.15$; Communication and Humanities: $M = 15.46$, $SD = 4.39$; Technology and Science: $M = 15.09$, $SD = 4.28$; Interdisciplinary: $M = 13.91$, $SD = 5.87$. It can be observed that Communication and Humanities and Technology and Science have the highest means in AI-based instructional design, while the Interdisciplinary group shows the lowest mean and greatest dispersion ($SD = 5.87$), suggesting internal heterogeneity.

Across all areas in F2, the scores are very consistent. The differences between means are minimal (≈ 1 point at most), and the standard deviations are relatively low (between 1.35 and 2.28), indicating stability in assessment practices across the board.

Finally, in F3, greater differences are observed. The Communication and Humanities area has the highest mean ($M = 22.97$), followed by Education ($M = 22.00$). Technology and Science and Health Sciences show more moderate scores. The standard deviations are relatively contained (≈ 2.58 – 4.20), indicating internal consistency within the groups.

To investigate differences across knowledge areas, the nonparametric Kruskal–Wallis test was applied, as the assumptions of normality were not met. In F1 (Activity Design), no significant differences were observed across knowledge areas ($H(4) = 2.834$, $p = 0.586$). Thus, the variability in this dimension cannot be attributed to the area of knowledge. Nor were statistically significant differences observed in the assessment of critical thinking by area of knowledge (F2; $H(4) = 4.455$, $p = 0.348$). However, significant differences were found between academic areas in F3 (promotion of self-regulation and reflection), with a medium effect size ($\epsilon^2 \approx 0.056$). Post hoc analyses using the Dunn test, with Holm correction, revealed differences between the Communication and Humanities area (G3) and the Health Sciences group (G1).

Post hoc analyses using the Dunn test with Holm correction indicated that only the comparison between the Health Sciences area and the Communication and Humanities area remained significant after adjustment ($p_{\text{holm}} = 0.024$). Although some additional comparisons were significant without correction, they did not hold up after controlling for multiplicity (Table 9).

Table 9. Dunn’s post hoc comparisons by disciplinary area for Factor 3 (Self-Regulation and Reflection).

Comparison	z	W_i	W_j	p	p_{bonf}	p_{holm}
1–2	−1.67	52.28	68.18	0.047 *	0.474	0.331
1–3	−2.82	52.28	79.08	0.002 **	0.024 *	0.024 *
1–4	−0.056	52.28	53.00	0.478	1.00	0.968
1–5	−0.460	52.28	56.73	0.323	1.00	0.968
2–3	−1.18	68.18	79.08	0.120	1.00	0.674
2–4	1.20	68.18	53.00	0.115	1.00	0.674
2–5	1.22	68.18	56.73	0.112	1.00	0.674
3–4	2.06	79.08	53.00	0.020 *	0.196	0.157
3–5	2.37	79.08	56.73	0.009 **	0.089	0.080
4–5	−0.29	53.00	56.73	0.385	1.00	0.968

* $p < 0.05$, ** $p < 0.01$.

Although some comparisons showed significant differences without correction for multiplicity ($p < 0.05$), after applying Holm’s correction, only the difference between the Health Sciences and Communication and Humanities fields remained significant ($p_{\text{holm}} = 0.024$). The remaining comparisons did not reach statistical significance after adjustment.

3.3. Outcomes of AI Tools Used by Teachers for Activity Design

In addition to the questionnaire, teachers were asked about their use of AI. Thus, 88.9% reported that they were using AI at the time the questionnaire was completed. Regarding the tools most frequently used by teachers, there is a clear concentration on a small number of applications, along with others that are more specific to particular tasks or areas of expertise. The most prominent tool was ChatGPT (N = 64; 44.4%), followed by Copilot (N = 59; 41%), indicating a predominance of generative AI use in teaching practice. These are followed by Perplexity (N = 10; 6.9%), Consensus (N = 9; 6.3%), and Gemini (N = 9; 6.3%), suggesting significant interest in tools geared toward searching for, synthesizing, and comparing information. Finally, Claude (N = 7; 4.9%), DALL-E (N = 6; 4.2%), and Gamma (N = 6; 4.2%) appear, while other tools such as Research Rabbit (N = 5; 3.5%), Connected Papers (N = 4; 2.8%), Elicit (N = 4; 2.8%), and Canva (N = 4; 2.8%) were used less frequently, though still identifiable within the teaching repertoire.

Likewise, there was a lower use of specific tools related to multimedia creation, writing support, content organization, or academic research, such as Fliki (N = 3; 2.1%), Suno (N = 3; 2.1%), Tome (N = 3; 2.1%), SciSpace (N = 3; 2.1%), DeepL (N = 3; 2.1%), and MagicSchool (N = 3; 2.1%). A wide range of tools with only a single mention was also identified, such as Paper Pal, Midjourney, Quizlet AI, QuestionWell, and Genially, among others. This pattern reflects a high degree of dispersion in more specialized uses and a clear dominance of a small group of general-purpose platforms. Thus, the adoption of specific AI tools is much more limited. Only two teachers reported not using any AI tools.

Regarding the uses of these tools, thematic coding indicated a primary use in supporting the creation and adaptation of materials (N = 42; 29.2%), the search, synthesis, and management of academic information (N = 38; 26.4%), assessment and evaluative support (N = 25; 17.4%), and use with students and AI literacy (N = 25; 17.4%). To a lesser extent, uses related to instructional planning and design (N = 21; 14.6%), writing, editing, and translation (N = 18; 12.5%), research support (N = 18; 12.5%), and programming or technical analysis (N = 16; 11.1%). Finally, there was a smaller number of responses related to critical supervision, plagiarism, and precautions (N = 9; 6.3%) and to non-use or limited use (N = 4; 2.8%).

In terms of pedagogical depth, the corpus primarily reflects instrumental or productivity-oriented uses, as well as uses involving direct pedagogical support, while more clearly transformative or reflective uses appear less frequently. In the coding performed, responses associated with an instrumental/productive logic were present in 67 cases (46.5%), those linked to pedagogical support in 41 (28.5%), and those associated with a transformative/reflective use in 30 responses (20.8%). This pattern suggests that teachers' integration of AI is in a transitional phase, still closely linked to operational efficiency, but with clear signs of progress toward uses closer to instructional design and critical reflection.

4. Discussion

The overall objective of this study is to analyze university instructors' instructional design practices and their integration and use of AI tools to design activities that foster critical thinking, self-regulation, and reflection. Thus, the study seeks to address how the integration of these emerging technologies can be described in terms of decision-making in instructional design, based on an ad hoc questionnaire. Consequently, the findings make a twofold contribution. First, the AI Instructional Design for Critical Thinking Questionnaire (AID-CTQ) is psychometrically validated, and its internal consistency is analyzed. Second, the study addresses the remaining sub-objectives to provide a description of how faculty integrate AI into instructional design. The results support the idea that the central challenge in higher education lies not only in determining whether AI is used, but

in understanding how it translates into specific curricular and pedagogical decisions. This would complement the limitations noted regarding models such as TPACK, UTAUT, or SAMR, which are useful for understanding technological integration or acceptance but less precise when the goal is to analyze how tasks, assessment, learning outcomes, and feedback align in AI-mediated contexts (Cabero-Almenara et al., 2024; Drugova et al., 2021; Hamilton et al., 2016; Jiménez-García et al., 2023).

4.1. Psychometric Validation and Reliability of the AID-CTQ

From a methodological perspective, the AID-CTQ is confirmed to have a three-factor structure centered on three dimensions: activity design (F1), critical thinking assessment (F2), and self-regulation and reflection (F3). This organization is consistent with the conceptualization of instructional design as a process for translating the pedagogical proposal through which faculty can organize their tasks in the era of AI (e.g., Kotsis, 2025; Liang et al., 2025; Rutecka et al., 2025; Xia et al., 2024).

The final 12-item model showed satisfactory fit and good overall reliability, supporting its usefulness for future research on the curricular integration of AI in higher education. This instrument allows us to move beyond models focused on technological adoption or acceptance to focus the analysis on the pedagogical decisions through which AI acquires educational meaning. Thus, it aligns with studies such as that of Xia et al. (2024), which highlights the importance of explicitly addressing how assessment and teaching practices are being transformed based on evidence.

4.2. Usage of AI Tools by Faculty in Teaching Activities

When interpreting the results of the AID-CTQ, it is important to note that, in the context of this study, 88.9% of instructors reported using AI tools. However, this high rate of use contrasts with the heterogeneity observed in the AID-CTQ indicators, which reinforces the conceptual distinction between using AI and integrating it pedagogically (Jiménez-García et al., 2023; Ojeda et al., 2023). As outlined in Section 3.3 of the results, there is a notable increase in the use of ChatGPT, consistent with other studies in the context of higher education, with the tool being used by faculty to expand, revise, and adapt their teaching materials rather than to replace design work (Karataş et al., 2024). Furthermore, a strong concentration of responses was observed in a group of general-purpose tools, with ChatGPT and Copilot standing out, followed by applications such as Perplexity, Consensus, Gemini, Claude, DALL·E, and Gamma.

This pattern suggests that faculty members rely primarily on accessible, versatile, and general-purpose tools, particularly those based on text generation, conversational assistance, and information synthesis. In contrast, more specialized tools for multimedia creation, knowledge organization, specialized academic research, or assessment show much more sporadic adoption. This predominance of general-purpose tools is consistent with a stage of integration in which AI is first incorporated as cross-cutting support for common academic and teaching tasks, before expanding into more specialized ecosystems of the possibilities offered by AI in the educational context (Karataş et al., 2024; Ojeda et al., 2023). This trend confirms the importance of understanding AI as a co-designer and decision-making support tool, rather than as an autonomous designer of curricula, similar to what has been suggested in other previous studies: Abbasi et al. (2024), Ullah et al. (2025), Zawacki-Richter et al. (2019) and Karataş et al. (2024). In this vein, the design of curricular activities supported by large language models (LLMs) is pedagogically useful when the teacher maintains control of the design process and does not delegate it entirely to AI (Rutecka et al., 2025).

Furthermore, the thematic coding of the open-ended responses shows that the primary uses are concentrated in the creation and adaptation of materials, the search for, synthesis, and management of academic information, assessment and evaluative support, and use with students and AI literacy. Less common are uses related to lesson planning, writing and translation, research support, and programming or technical analysis. This overview confirms that AI is primarily being used to address core university work needs: preparing lessons, producing resources, searching for information, structuring assessments, and supporting complex academic tasks. This aligns with the TALIS report (OECD, 2025) and studies such as those by Perezchica-Vega et al. (2024) and Ng et al. (2025), which highlight a primary use in the generation of lesson plans, teaching activities, or specific instructional resources. The use of AI for academic information management or student assessment processes (creation of rubrics, question banks, and exams) is also noted. However, it also suggests that the adoption of these tools remains concentrated in areas of high practical and immediate utility, rather than in a profound redesign of the curriculum.

These results are also consistent with the levels of technology integration described in models such as SAMR (Hamilton et al., 2016) or with the stages of critical adoption of AI identified in recent literature (Fengchun & Cukurova, 2025; Xia et al., 2024): most teachers would be at the substitution or modification stages, while a significant minority show signs of pedagogical redefinition. This distribution supports the idea that teachers' integration of AI is in a transitional phase, still closely tied to operational efficiency, but with clear signs of progress toward uses more closely aligned with instructional design and critical reflection.

The gap between the high rate of reported use and the most transformative integration profiles suggests that the pedagogical challenge is no longer one of access or adoption, but rather one of qualifying the use: what kinds of curricular and pedagogical decisions are triggered by engagement with these tools. Previous studies indicate that digital competence acts as a lever for pedagogical innovation and that specific competence in AI can contribute to improving both the quality of instructional design and the ability to foster critical thinking skills in students (Ojeda et al., 2023; Perezchica-Vega et al., 2024). Similarly, the INTEF guide for the educational use of AI explicitly places student autonomy and critical thinking among the priority goals of its integration into teaching practice (INTEF, 2024). Therefore, familiarity with AI should not be interpreted solely as a matter of technical mastery, but as a potential facilitator of pedagogical redesign. This distinction, between using AI as a tool for efficiency and using it as a means to scaffold students' thinking, constitutes precisely the central axis that the AID-CTQ seeks to operationalize.

The descriptive results provide a revealing picture of the current state of AI integration at the participating institution. In general terms, faculty reported relatively high levels across the three dimensions assessed, suggesting a favorable attitude toward the pedagogical use of AI. The highest levels of agreement were observed in items related to the assessment of student reasoning and to practices aimed at promoting reflection, ethical awareness, and intellectual autonomy. In contrast, the items associated with activity design showed greater variability, particularly those related to comparing perspectives and structuring evidence-based discussion. This pattern suggests that faculty members appear to have embraced the discourse on the need to critically evaluate AI more clearly than they have the practical application of AI in complex academic tasks. That is, the results point to an ongoing transition from an instrumental use of AI toward a more mature curricular integration. This finding is consistent with recent literature in higher education (e.g., Jiménez-García et al., 2023; Perezchica-Vega et al., 2024), which emphasizes that the educational potential of AI depends less on the tool itself than on faculty supervision, pedagogical intent, and the redesign of task and assessment structures. This observed "operational gap" suggests that institutional training should pivot from basic tool-use in-

struction toward “relational engineering”, focusing on the design of prompts and activities that scaffold higher-order cognitive processes.

4.3. Differences in Instructional Design Facilitated by AI Tools: AID-CTQ

One of the key findings emerged when analyzing differences based on teaching experience. Significant differences were observed only in the dimension of activity design (F1), where teachers with less than 10 years of experience scored significantly lower than those with 10–20 years and those with more than 20 years of experience. No differences were observed in the assessment of critical thinking (F2) or in self-regulation and reflection (F3). This pattern suggests that teaching experience appears to be specifically linked to the ability to translate pedagogical principles into concrete task-design decisions. It is possible that AI-mediated instructional design requires a higher level of curricular judgment, methodological repertoire, and pedagogical maturity, elements that are progressively consolidated through professional practice. This finding does not necessarily imply that more experienced faculty members are more receptive to AI, but rather that they seem to integrate it more effectively from an instructional design perspective. In this sense, the results expand upon studies focused on technological acceptance, where variables such as age can influence willingness to use (Cabero-Almenara et al., 2024), by showing that professional experience may play a different role when analyzing the quality of instructional design.

In contrast, the absence of differences based on experience in dimensions F2 and F3 suggests that these may be more closely linked to shared pedagogical discourses, institutional frameworks, or cross-cutting training processes than to accumulated years of teaching experience. In other words, the importance of assessing reasoning and promoting reflective processes appears to be widely recognized among faculty, although the specific ways of implementing these practices remain inconsistent. This interpretation is particularly relevant when assessing critical thinking (F2). Although this dimension received high scores and showed consistency across groups, it also exhibited more moderate internal consistency. This may be interpreted as an indication that, while this component is widely accepted in theory, it remains challenging to implement in a stable and consistent manner in AI-mediated contexts. This complexity has been highlighted by recent reviews such as those by Fowler (2023) and Xia et al. (2024), which emphasize the need to shift assessment from the final product toward processes of reasoning, verification, justification, and academic integrity.

The results by field of study did not identify significant differences in activity design (F1) or in the assessment of critical thinking (F2), suggesting that these two dimensions are relatively cross-cutting within the university context. However, the differences observed across disciplinary areas were concentrated in the dimension of self-regulation and reflection, suggesting that this component of AI-mediated instructional design may be shaped more strongly by disciplinary culture than activity design or assessment practices. One possible interpretation is that fields such as Communication and Humanities more naturally incorporate tasks centered on discourse analysis, perspective-taking, interpretation, and critical engagement with meaning, which may facilitate the explicit integration of reflective and ethical processes when working with AI. By contrast, in Health Sciences and other applied fields, AI is often embedded in more procedural, structured, or decision-oriented learning contexts, where reflection is present but may be operationalized through protocol-based reasoning, clinical judgment, or risk-aware practice rather than through explicitly dialogic or discursive tasks (Rincón et al., 2025; Chun et al., 2025; Babacan et al., 2025; Zhang & Tang, 2025). This interpretation is broadly consistent with the literature suggesting that AI integration takes different curricular forms across disciplinary contexts. In applied and health-related fields, AI tends to be associated with structured, case-based,

or decision-support tasks (Chun et al., 2025), whereas in Communication and Humanities it is more often linked to interpretive, discursive, and creative uses that may make reflective and ethical processes more explicit within teaching practice (Babacan et al., 2025; Zhang & Tang, 2025).

This disciplinary finding ties into a broader debate in the literature: the need to prevent students from uncritically delegating their judgment to AI systems, and to promote the evaluation of the reliability, biases, and implications of the generated content. Previous studies had already highlighted the importance of establishing critical initiatives for the use of AI, as well as concerns regarding the reliability and quality of its results (e.g., Fowler, 2023; Flores-Vivar & García-Peñalvo, 2023). In higher education, this challenge intensifies when attempting to discern to what extent academic outcomes reflect students' critical and creative thinking or, conversely, the instrumental use of AI tools (Castillejos López, 2022; Garro Mena, 2024). From this perspective, the fact that disciplinary differences emerge precisely in the dimension of self-regulation and reflection reinforces the importance of this dimension as the pedagogical and ethical core of AI integration.

4.4. Integrating AI into Higher Education Curricula: Evidence, Trends, and Practical Implications

The findings of this study are part of a broader landscape of AI-driven curricular transformation in higher education. Recent works support these evolving changes to adapt to emerging technologies, highlighting three major areas of challenge: curriculum design and integration, pedagogical practice and teacher training, and strategic planning with innovation (Chu & Ashraf, 2025; Sapawi & Yusoff, 2025). Within this framework, understanding AI as a tool to be integrated into the curriculum in a transdisciplinary manner involves addressing how such curricular integration should be structured to support pedagogical practices and the changes it entails for the teaching role (Zheng, 2025; Kotsis, 2025).

Regarding specific areas of AI integration into the curriculum, the literature identifies applications in learning analytics, personalized learning, and recommendation systems; the use of large language models (LLMs) for curriculum design; and specific platforms that integrate AI in a structured manner (e.g., Chu & Ashraf, 2025; Chen et al., 2020; Zawacki-Richter et al., 2019; Rutecka et al., 2025; Chun et al., 2025; Zhang & Tang, 2025). These initiatives confirm that integration is occurring both in the form of transdisciplinary activities aimed at acquiring knowledge and practical skills, and in the form of specialized teaching by field: in technological knowledge, focused on technical tools and content (Abbasnejad et al., 2025; Ullah et al., 2025; Kotsis, 2025); in Health sciences, oriented toward clinical case resolution, assisted diagnosis, AI laboratories, and research, with explicit attention to biases and risks (Rincón et al., 2025); in the humanities and social sciences, implemented primarily through pedagogical strategies centered on critical thinking (Zhang & Tang, 2025; Zheng, 2025); and in business, arts, and design, combining the teaching of generative tools, data analysis, and creative experimentation to reformulate ideation processes.

Across the board, the integration of AI into the curriculum tends to rely on constructivist pedagogical models or active methodologies such as problem-based learning (Abbasnejad et al., 2025; Hao et al., 2025). Systematic reviews emphasize that teachers' knowledge, institutional support, and AI literacy are decisive factors in determining whether AI tools enhance or merely complicate curricular processes (Liang et al., 2025; Zawacki-Richter et al., 2019). Investing in teacher training on AI, ethical governance, and the necessary infrastructure is therefore a structural prerequisite for AI tools to truly improve the quality and equity of curricula (Abbasi et al., 2024; Ullah et al., 2025; Sapawi & Yusoff, 2025).

In practical terms, the study's findings point to several significant implications for higher education. First, teacher training in AI should shift from a tool-centered approach to one focused on pedagogical skills: designing authentic tasks, assessing reasoning, scaffolding self-regulation, and fostering ethical reflection. The profile of tools used by teachers, predominantly generalist platforms such as ChatGPT and Copilot, indicates that the entry point is accessible, but that training should aim to pedagogically refine that initial use, progressing from instrumental approaches toward transformative applications. Second, higher education institutions should promote assessment models that value processes of thinking, critical review, and justification, rather than just final products. Third, the results underscore the need for clearer institutional frameworks to guide a responsible, critical, and pedagogically grounded use of AI. These implications are consistent with DigCompEdu, UNESCO's guidelines on generative AI, the INTEF guide, and the European AI Act, all of which focus on the ethical, informed, and human-supervised integration of these technologies (Redecker, 2017; Fengchun & Cukurova, 2025; INTEF, 2024; Regulation (EU) 2024/1689, 2024).

4.5. Limitations and Future Research Opportunities

This study has several limitations that should be taken into account. First, the sample comes from a single private university in Madrid and was selected using a non-probabilistic convenience sampling procedure, which limits the generalizability of the results. Second, this is a study based on teacher self-reports; therefore, the data reflect reported practices rather than those observed directly. In addition, both the exploratory and confirmatory factor analyses were conducted on the same sample of 144 faculty members. Although this approach is acceptable in an initial psychometric study, it limits the strength of the structural validation and suggests that future research should replicate the factor structure in an independent sample. Third, some subgroups in the comparative analysis were small, particularly in the case of the Technology and Science area, which calls for caution in inferential interpretation. Finally, although the overall psychometric performance of the instrument was satisfactory, the critical thinking assessment dimension (F2) showed more moderate internal consistency and may require further refinement in future studies. Additionally, the wide range of teaching experience (1–50 years) and the inclusion of faculty without prior AI experience (11.10%) introduced heterogeneity that may affect response patterns. This heterogeneity was retained intentionally to reflect the diversity of instructional profiles currently engaged in AI adoption, but it may also limit the comparability of subgroups and should be considered when interpreting the between-group analyses.

Consequently, future research could: (a) validate the AID-CTQ in broader institutional and disciplinary contexts; (b) triangulate faculty self-reports with analyses of actual student activities and products; (c) examine the relationship between faculty design decisions and students' actual performance on AI-mediated academic tasks; and (d) explore the link between the type of tools used and the pedagogical depth of the identified uses, to understand which conditions favor the transition from instrumental uses to transformative uses.

5. Conclusions

This study presents an analysis of the use of AI in higher education from an instructional design perspective to gain an understanding of teaching practices, thereby expanding upon existing studies that address this analysis solely from the perspective of adoption or frequency of use. The AID-CTQ demonstrated a valid three-factor structure and adequate overall reliability, supporting its utility in research on the curricular integration of AI in

university settings. Thus, in line with the objectives of this study, the following conclusions are drawn:

- The AID-CTQ, developed ad hoc, provides a specific and operational tool for measuring instructional design involving the use of AI, aimed at fostering critical thinking in students.
- Greater educational experience is associated with higher scores in activity design. For the knowledge domain, some differences are observed in the reflective dimension.
- The results suggest that the primary need is not to learn more tools but to strengthen task design, reasoning assessment, and critical self-regulation.

Thus, the study's findings suggest that the educational value of implementing AI in higher education will depend on faculty members' ability to translate it into instructional design decisions that foster critical thinking, self-regulation, and responsible reflection. The high reported usage rate (88.9%) and the predominance of general-purpose tools such as ChatGPT and Copilot confirm that AI is already part of the standard teaching repertoire, although most of the identified uses still respond to the logic of productivity and instrumentalism rather than the logic of curricular transformation. In this sense, the shift from "tool use" to "curricular implementation" constitutes not only a conceptual shift but also an empirically observable pedagogical challenge. This aligns with international frameworks that emphasize ethical integration, centered on human agency and pedagogically validated.

Author Contributions: Conceptualization, N.L.N.-M., T.G.-M. and F.L.-M.; methodology, N.L.N.-M.; validation, N.L.N.-M., T.G.-M. and F.L.-M.; formal analysis, N.L.N.-M. and L.R.-M.; investigation, N.L.N.-M., R.M.-C., L.R.-M., N.R.-D., T.G.-M. and F.L.-M.; resources, R.M.-C. and N.R.-D.; data curation, N.L.N.-M., R.M.-C. and L.R.-M.; writing—original draft preparation, N.L.N.-M. and R.M.-C.; writing—review and editing, N.L.N.-M., R.M.-C., L.R.-M. and N.R.-D.; visualization, N.L.N.-M., R.M.-C., L.R.-M. and N.R.-D.; supervision, N.L.N.-M., T.G.-M. and F.L.-M.; project administration, N.L.N.-M., T.G.-M. and F.L.-M.; funding acquisition, T.G.-M. and F.L.-M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was commissioned by the World Innovation Summit for Education (WISE), an initiative of Qatar Foundation, in collaboration with its lead research partner, the Institute of International Education (IIE), as part of the Artificial Intelligence and Higher Education Research Consortium Initiative. The views, findings, and conclusions expressed herein are those of the authors and do not necessarily reflect those of WISE, Qatar Foundation, or IIE.

Institutional Review Board Statement: The study was approved by the Research Ethics Committee of Camilo José Cela University (CEI-UCJC, Madrid, Spain) under code 19_24_WISE, corresponding to the research project Critical Thinking and Workforce Transformation: Redefining University Curricula for the AI Era (RUCAIE). Data collection, storage, processing, and transfer procedures complied with Regulation (EU) 2016/679 (General Data Protection Regulation, GDPR) and applicable Spanish legislation governing data protection, documentation, and archives, including Ley 16/1985, de 25 de junio, del Patrimonio Histórico Español [Law 16/1985 on Spanish Historical Heritage], Ley 4/1993, de 21 de abril, de Archivos y Patrimonio Documental de la Comunidad de Madrid [Law 4/1993 on Archives and Documentary Heritage of the Community of Madrid], and Real Decreto 1708/2011, de 18 de noviembre, por el que se establece el Sistema Español de Archivos [Royal Decree 1708/2011 establishing the Spanish Archive System]. Participants provided informed consent prior to participation, and all data were anonymized before analysis, dissemination, and storage in accordance with the approved protocol.

Informed Consent Statement: The volunteers declared their consent to participate in this study.

Data Availability Statement: Data can be requested by writing to the corresponding author.

Acknowledgments: First, we would like to thank and acknowledge the contributions of the students who participated in the research for this project: Alejandra Bedmar Arrogante; Mar Esteve García; Sandra Martín Aguirre; and Claudia Manuela Portillo. We would also like to thank the Corporate Intelligence team for their support and collaboration in distributing the online questionnaire for this project: Giuseppe Auricchio; Guillermo Delso Segovia; Begoña Sopena Egusquiza; and César Valentín Gómez de la Cal. During the preparation of this manuscript, the authors used DeepL Translator (free web version) for translation support and language refinement. The authors also used ChatGPT Plus (paid version) to assist with the clarification of specific questions and the final review of the manuscript before submission. No AI tool was used to generate, interpret, or analyze the research data, nor to make scientific decisions. All outputs were critically reviewed, verified, and edited by the authors, who take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Appendix A

This appendix includes the various versions of the questionnaire designed (the “AI Instructional Design for Critical Thinking Questionnaire” [AID-CTQ]), which is explained in Section 2 of this work.

Table A1. Initial AID-CTQ developed based on the literature. For clarity, the items used in Spanish are included in [brackets].

Dimension	Item
Activity Design (F1) [Diseño de Actividades]	Q1 The activities I design using AI encourage students to seek out and critically evaluate information. [Las actividades que diseño con IA fomentan la búsqueda y evaluación crítica de información por parte de los estudiantes].
	Q2 I often incorporate exercises into my course that require students to compare different perspectives on a topic. [Suelo incorporar en mi asignatura ejercicios que requieran que los estudiantes comparen diferentes puntos de vista sobre un tema].
	Q3 My activities promote reflection and in-depth analysis of the topics studied. [Mis actividades promueven la reflexión y el análisis profundo de los temas estudiados].
	Q4 I encourage discussion and debate so that students can practice evidence-based argumentation. [Fomento la discusión y el debate para que los estudiantes practiquen la argumentación basada en evidencia].
Critical Thinking Assessment (F2) [Evaluación del Pensamiento Crítico]	Q5 I evaluate students not only on their answers, but also on the reasoning process they use. [Evalúo a los estudiantes no solo por sus respuestas, sino también por el proceso de razonamiento que utilizan].
	Q6 I provide feedback that helps students improve their analytical and evaluative skills. [Proporciono retroalimentación que ayuda a los estudiantes a mejorar su capacidad de análisis y evaluación].
	Q7 I use open-ended questions on exams or assignments to help students demonstrate their ability to draw inferences and exercise critical judgment. [Uso preguntas abiertas en exámenes o tareas para que los estudiantes demuestren su capacidad de inferencia y juicio crítico].
	Q8 I design case studies or real-world problems that students must solve using critical thinking. [Diseño casos de estudio o problemas reales que los estudiantes deben resolver utilizando pensamiento crítico].

Table A1. *Cont.*

Dimension	Item	
Self-Regulation & Reflection (F3) [Fomento de la Autorregulación y Reflexión]	Q9	I encourage students to review and adjust their decisions in light of new information. [Fomento que los estudiantes revisen y ajusten sus decisiones a la luz de nueva información].
	Q10	I encourage students to identify their own biases and work to overcome them. [Animo a los estudiantes a identificar sus propios sesgos y trabajar para superarlos].
	Q11	I design activities that encourage reflection on the ethical and social implications of decisions. [Creo actividades que promuevan la reflexión sobre las implicaciones éticas y sociales de las decisiones].
	Q12	I facilitate self-reflection among my students so they can identify areas for improvement in their thought processes. [Facilito la autorreflexión en mis estudiantes para que identifiquen áreas de mejora en su proceso de pensamiento].
	Q13	I promote intellectual autonomy by encouraging students to question information and make informed decisions. [Promuevo la autonomía intelectual, alentando a los estudiantes a cuestionar la información y a tomar decisiones informadas].
Context of AI usage [Contextualización del uso]	Q14	What AI tools have you used in your teaching so far? [¿Qué herramientas de IA has utilizado en tu docencia hasta la fecha?]
	Q15	What has been your goal in using AI in your teaching activities? [¿Con qué objetivo has utilizado la IA en tus actividades docentes?]
	Q16	Describe an activity you have implemented that you consider particularly effective for developing critical thinking in your students. What results did you observe? [Describe una actividad que hayas implementado que consideres especialmente efectiva para desarrollar el pensamiento crítico en tus estudiantes. ¿Qué resultados observaste?]
	Q17	How have you implemented or used AI in your classes to foster critical thinking? [¿Cómo has implementado o utilizado la IA en tus clases para trabajar y fomentar el pensamiento crítico?]
	Q18	Do you have any other comments regarding the use of AI and the development of critical thinking in students? [¿Tienes algún otro comentario en relación al uso de la IA y desarrollo del pensamiento crítico en los estudiantes?]

Table A2. AID-CTQ with modifications following expert review; the wording and clarity of the items have been improved, and the number of open-ended questions has been reduced. For clarity, the items used in Spanish are included in [brackets].

Dimension	Item	
Activity Design (F1) [Diseño de Actividades]	Q1	The activities I design using AI encourage students to seek out and critically reflect on information. [Las actividades que diseño con IA fomentan la búsqueda y reflexión crítica de la información por parte de los estudiantes.]
	Q2	I often incorporate AI-based exercises into my course that require students to compare different perspectives on a topic. [Suelo incorporar en mi asignatura ejercicios con IA que requieran que los estudiantes comparen diferentes puntos de vista sobre un tema.]
	Q3	My activities promote the analysis of the topics studied using AI. [Mis actividades promueven el análisis de los temas estudiados con la IA.]
	Q4	I encourage discussion and debate so that students practice evidence-based argumentation using AI. [Fomento la discusión y el debate para que los estudiantes practiquen la argumentación basada en evidencia a partir del uso de la IA.]

Table A2. Cont.

Dimension	Item
Critical Thinking Assessment (F2) [Evaluación del Pensamiento Crítico]	Q5 I assess students not only on their answers but also on the reasoning process they use. [Evalúo a los estudiantes no solo por sus respuestas, sino también por el proceso de razonamiento que utilizan.]
	Q6 I provide feedback that helps students improve their analytical skills. [Proporcione retroalimentación que ayuda a los estudiantes a mejorar su capacidad de análisis.]
	Q7 I use open-ended questions (on exams or assignments) to have students demonstrate their ability to make inferences and exercise critical judgment. [Uso preguntas abiertas (en exámenes o tareas) para que los estudiantes demuestren su capacidad de inferencia y juicio crítico.]
	Q8 I design case studies or real-world problems that students must solve using critical thinking. [Diseño casos de estudio o problemas reales que los estudiantes deben resolver utilizando pensamiento crítico.]
Self-Regulation & Reflection (F3) [Fomento de la Autorregulación y Reflexión]	Q9 I encourage students to review and adjust their decisions from different perspectives. [Fomento que los estudiantes revisen y ajusten sus decisiones a partir de diferentes perspectivas.]
	Q10 I encourage students to identify their own biases and work to ensure they do not limit their performance. [Animo a los estudiantes a identificar sus propios sesgos y trabajar para que no supongan una limitación en su desempeño.]
	Q11 I create activities that promote reflection on the ethical and social implications of decisions. [Creo actividades que promuevan la reflexión sobre las implicaciones éticas y sociales de las decisiones.]
	Q12 I facilitate self-reflection among my students so they can identify areas for improvement in their thought processes. [Facilito la autorreflexión en mis estudiantes para que identifiquen áreas de mejora en su proceso de pensamiento.]
	Q13 I promote intellectual autonomy by encouraging students to question information and make informed decisions. [Promuevo la autonomía intelectual, alentando a los estudiantes a cuestionar la información y a tomar decisiones informadas.]
Context of AI usage [Contextualización del uso]	Qa1 What AI tools have you currently used in your teaching? [¿Qué herramientas de IA has utilizado en tu docencia hasta la fecha?]
	Qa2 What is the purpose for which you have used AI in your educational activities? [¿Con qué objetivo has utilizado la IA en tu docencia?]

Table A3. AID-CTQ with adjustments following statistical validation. Item Q8 has been removed. For clarity, the items used in Spanish are included in [brackets].

Dimension	Item
Activity Design (F1) [Diseño de Actividades]	Q1 The activities I design using AI encourage students to seek out and critically reflect on information. [Las actividades que diseño con IA fomentan la búsqueda y reflexión crítica de la información por parte de los estudiantes.]
	Q2 I often incorporate AI-based exercises into my course that require students to compare different perspectives on a topic. [Suelo incorporar en mi asignatura ejercicios con IA que requieran que los estudiantes comparen diferentes puntos de vista sobre un tema.]
	Q3 My activities promote the analysis of the topics studied using AI. [Mis actividades promueven el análisis de los temas estudiados con la IA.]
	Q4 I encourage discussion and debate so that students practice evidence-based argumentation using AI. [Fomento la discusión y el debate para que los estudiantes practiquen la argumentación basada en evidencia a partir del uso de la IA.]

Table A3. Cont.

Dimension		Item
Critical Thinking Assessment (F2) [Evaluación del Pensamiento Crítico]	Q5	I assess students not only on their answers but also on the reasoning process they use. [Evalúo a los estudiantes no solo por sus respuestas, sino también por el proceso de razonamiento que utilizan.]
	Q6	I provide feedback that helps students improve their analytical skills. [Proporciono retroalimentación que ayuda a los estudiantes a mejorar su capacidad de análisis.]
	Q7	I use open-ended questions (on exams or assignments) to have students demonstrate their ability to make inferences and exercise critical judgment. [Uso preguntas abiertas (en exámenes o tareas) para que los estudiantes demuestren su capacidad de inferencia y juicio crítico.]
Self-Regulation & Reflection (F3) [Fomento de la Autorregulación y Reflexión]	Q9	I encourage students to review and adjust their decisions from different perspectives. [Fomento que los estudiantes revisen y ajusten sus decisiones a partir de diferentes perspectivas.]
	Q10	I encourage students to identify their own biases and work to ensure they do not limit their performance. [Animo a los estudiantes a identificar sus propios sesgos y trabajar para que no supongan una limitación en su desempeño.]
	Q11	I create activities that promote reflection on the ethical and social implications of decisions. [Creo actividades que promuevan la reflexión sobre las implicaciones éticas y sociales de las decisiones.]
	Q12	I facilitate self-reflection among my students so they can identify areas for improvement in their thought processes. [Facilito la autorreflexión en mis estudiantes para que identifiquen áreas de mejora en su proceso de pensamiento.]
	Q13	I promote intellectual autonomy by encouraging students to question information and make informed decisions. [Promuevo la autonomía intelectual, alentando a los estudiantes a cuestionar la información y a tomar decisiones informadas.]
Context of AI usage [Contextualización del uso]	Qa1	What AI tools have you currently used in your teaching? [¿Qué herramientas de IA has utilizado en tu docencia hasta la fecha?]
	Qa2	What is the purpose for which you have used AI in your educational activities? [¿Con qué objetivo has utilizado la IA en tu docencia?]

References

- Abbasi, B., Wu, Y., & Luo, Z. (2024). Exploring the impact of artificial intelligence on curriculum development in global higher education institutions. *Education and Information Technologies*, 30(1), 547–581. [\[CrossRef\]](#)
- Abbasnejad, B., Soltani, S., Taghizadeh, F., & Zare, A. (2025). Developing a multilevel framework for AI integration in technical and engineering higher education: Insights from bibliometric analysis and ethnographic research. *Interactive Technology and Smart Education*, 23(1), 49–79. [\[CrossRef\]](#)
- Abdelhalim, S. M. (2024). Using CHATGPT to promote research competency: English as a foreign language undergraduates' perceptions and practices across varied metacognitive awareness levels. *Journal of Computer Assisted Learning*, 40(3), 1261–1275. [\[CrossRef\]](#)
- Abuhassna, H., Alnawajha, S., Awae, F., Adnan, M. A. B. M., & Edwards, B. I. (2024). Synthesizing technology integration within the Addie model for instructional design: A comprehensive systematic literature review. *Journal of Autonomous Intelligence*, 7(5), 1–28. [\[CrossRef\]](#)
- Babacan, H., Arık, E., Bilişli, Y., Akgün, H., & Özkara, Y. (2025). Artificial intelligence and journalism education in higher education: Digital transformation in undergraduate and graduate curricula in Türkiye. *Journalism and Media*, 6(2), 52. [\[CrossRef\]](#)
- Batubara, H. H., Syukur, F., Junaedi, M., Purwanti, K. L., & Shanie, A. (2025). Teachers' perceptions and readiness for mobile artificial intelligence integration in elementary education. *International Journal of Interactive Mobile Technologies*, 19(16), 148–162. [\[CrossRef\]](#)
- Biggs, J., & Tang, C. (2011). *Teaching for quality learning at university* (4th ed.). McGraw-Hill Education.
- Branch, R. M. (2009). *Instructional design: The ADDIE approach*. Springer. [\[CrossRef\]](#)
- Cabañuz, C., & García-García, M. (2024). ChatGPT es un estudiante por encima de la media en la Facultad de Medicina de la Universidad de Zaragoza y un colaborador excelente en la elaboración de material docente. *Revista Española de Patología*, 57(2), 91–96. [\[CrossRef\]](#) [\[PubMed\]](#)

- Cabero-Almenara, J., Palacios-Rodríguez, A., Loaiza-Aguirre, M. I., & Rivas-Manzano, M. D. R. D. (2024). Acceptance of educational artificial intelligence by teachers and its relationship with some variables and pedagogical beliefs. *Education Sciences*, 14(7), 740. [CrossRef]
- Castillejos López, B. (2022). Inteligencia artificial y los entornos personales de aprendizaje: Atentos al uso adecuado de los recursos tecnológicos de los estudiantes universitarios. *Educación*, 31(60), 9–24. [CrossRef]
- Chan, K. K.-W., & Tang, W. K.-W. (2024). Evaluating English teachers' artificial intelligence readiness and training needs with a TPACK-based model. *World Journal of English Language*, 15(1), 129–145. [CrossRef]
- Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE Access*, 8, 75264–75278. [CrossRef]
- Chu, T., & Ashraf, M. (2025). Artificial intelligence in curriculum design: A data-driven approach to higher education innovation. *Knowledge*, 5(3), 14. [CrossRef]
- Chun, J., Kim, J., Kim, H., Lee, G., Cho, S., Kim, C., Chung, Y., & Heo, S. (2025). A comparative analysis of on-device AI-driven, self-regulated learning and traditional pedagogy in university health sciences education. *Applied Sciences*, 15(4), 1815. [CrossRef]
- Daher, R. (2025). Integrating AI literacy into teacher education: A critical perspective paper. *Discover Artificial Intelligence*, 5(1), 217. [CrossRef]
- Drigas, A., & Mitsea, E. (2020). The 8 pillars of metacognition. *International Journal of Emerging Technologies in Learning (iJET)*, 15(21), 162–178. [CrossRef]
- Drugova, E., Zhuravleva, I., Aiusheeva, M., & Grits, D. (2021). Toward a model of learning innovation integration: TPACK-SAMR based analysis of the introduction of a digital learning environment in three Russian universities. *Education and Information Technologies*, 26(4), 4925–4942. [CrossRef] [PubMed]
- Ejjami, R. (2024). The future of learning: AI-based curriculum development. *International Journal for Multidisciplinary Research*, 6, 1–31. [CrossRef]
- ElSayary, A., Kuhail, M. A., & Hojeij, Z. (2025). Examining the role of prompt engineering in utilizing generative AI tools for lesson planning: Insights from teachers' experiences and perceptions. *Human Behavior and Emerging Technologies*, 2025(1), 9986139. [CrossRef]
- Espejo Aubá, P. C. (2024). La inteligencia artificial en educación: Percepciones y saberes de los docentes. *European Public & Social Innovation Review*, 9, 1–19. [CrossRef]
- Fengchun, M., & Cukurova, M. (2025). Marco de competencias para docentes en materia de IA. UNESCO. [CrossRef]
- Flores-Vivar, J. M., & García-Peñalvo, F. J. (2023). Reflexiones sobre la ética, potencialidades y retos de la inteligencia artificial en el marco de la educación de calidad (ODS4). *Comunicar: Revista Científica de Comunicación y Educación*, 31(74), 37–47. [CrossRef]
- Fowler, D. S. (2023). AI in higher education: Academic integrity, harmony of insights, and recommendations. *Journal of Ethics in Higher Education*, (3), 127–143. [CrossRef]
- Fragouli, E. (2025). Digital transformations & innovations in business PG education: Rethinking curriculum development and teaching practices. *International Journal of Higher Education Management*, 11(2), 37–48. [CrossRef]
- García-López, C., Tabuenca-Cuevas, M., & Navarro-Soria, I. (2025). A systematic review of the use of AI in EFL and EL classrooms for gifted students. *Trends in Higher Education*, 4(3), 33. [CrossRef]
- Garro Mena, L. (2024). Optimizando el aprendizaje mediante IA: La eficacia de flujos de trabajo estructurados en la educación superior. *Revista de Docencia Universitaria*, 22(2), 105–121. [CrossRef]
- Hamilton, E. R., Rosenberg, J. M., & Akcaoglu, M. (2016). The substitution augmentation modification redefinition (SAMR) model: A critical review and suggestions for its use. *TechTrends*, 60, 433–441. [CrossRef]
- Hansel, C. A., Ottenbreit-Leftwich, A., Quick, J. D., Greene, A. H., & Ricci, M. (2024). Gradescope in large lecture classrooms: A case study at Indiana University: How an online grading platform enhanced student learning and instructor feedback in large-scale courses. *Journal of Teaching and Learning with Technology*, 13(1), 33–48. [CrossRef]
- Hao, Y., Liu, Y., Liu, B., Amarantidis, G., & Ghannam, R. (2025). Integrating AI in engineering education: A comprehensive review and student-informed module design for U.K. students. *IEEE Transactions on Education*, 68, 173–185. [CrossRef]
- Instituto Nacional de Tecnologías Educativas y de Formación del Profesorado [INTEF]. (2024). Guía sobre el uso de la inteligencia artificial en el ámbito educativo. Available online: <https://intef.es/Noticias/guia-sobre-el-uso-de-la-inteligencia-artificial-en-el-ambito-educativo/> (accessed on 26 March 2026).
- Jiménez-García, E., Orenes-Martínez, N., & López-Fraile, L. A. (2023). Rueda de la pedagogía para la inteligencia artificial: Adaptación de la Rueda de Carrington. *RIED-Revista Iberoamericana de Educación a Distancia*, 27(1), 87–113. [CrossRef]
- Karataş, F., Eriçok, B., & Tanrikulu, L. (2024). Reshaping curriculum adaptation in the age of artificial intelligence: Mapping teachers' AI-driven curriculum adaptation patterns. *British Educational Research Journal*, 51, 154–180. [CrossRef]
- Kotsis, K. (2025). Integrating artificial intelligence into higher education curricula: Challenges and opportunities for science-based programs. *International Journal of Artificial Intelligence Engineering and Transformation*, 6(1), 4–12. [CrossRef]

- León Naranjo, J. (2024). El modelo conocimiento tecnológico pedagógico y de contenido (TPACK): Una estrategia para potenciar las competencias digitales de los docentes: The technological pedagogical and content knowledge (TPACK) model: A strategy to enhance teachers' digital skills. *LATAM Revista Latinoamericana de Ciencias Sociales y Humanidades*, 5(4), 2079–2094. [CrossRef]
- Liang, J., Stephens, J., & Brown, G. (2025). A systematic review of the early impact of artificial intelligence on higher education curriculum, instruction, and assessment. *Frontiers in Education*, 10, 1522841. [CrossRef]
- Loksa, D., Margulieux, L., Becker, B. A., Craig, M., Denny, P., Pettit, R., & Prather, J. (2022). Metacognition and self-regulation in programming education: Theories and exemplars of use. *ACM Transactions on Computing Education*, 22(4), 1–31. [CrossRef]
- Machado, A., Tenório, K., Santos, M. M., Barros, A. P., Rodrigues, L., Mello, R. F., Paiva, R., & Dermeval, D. (2025). Workload perception in educational resource recommendation supported by artificial intelligence: A controlled experiment with teachers. *Smart Learning Environments*, 12(1), 20. [CrossRef]
- Ng, D. T. K., Chan, E. K. C., & Lo, C. K. (2025). Opportunities, challenges and school strategies for integrating generative AI in education. *Computers and Education: Artificial Intelligence*, 8, 100373. [CrossRef]
- Ojeda, A. D., Solano-Barliza, A. D., Álvarez, D. O., & Cárcamo, E. B. (2023). Análisis del impacto de la inteligencia artificial ChatGPT en los procesos de enseñanza y aprendizaje en la educación universitaria. *Formación Universitaria*, 16(6), 61–70. [CrossRef]
- Organization for Economic Co-Operation and Development [OECD]. (2025). TALIS 2024. *Estudio internacional de la enseñanza y del aprendizaje*. Available online: https://www.libreria.educacion.gob.es/libro/talis-2024-estudio-internacional-de-la-ensenanza-y-el-aprendizaje-informe-espanol_186331/ (accessed on 26 March 2026).
- Perezchica-Vega, J. E., Sepúlveda-Rodríguez, J. A., & Román-Méndez, A. D. (2024). Inteligencia artificial generativa en la educación superior: Usos y opiniones de los profesores [Generative artificial intelligence in higher education: Uses and opinions of teachers]. *European Public & Social Innovation Review*, 9, 1–20. [CrossRef]
- Real Decreto 217/2022, de 29 de marzo, por el que se establece la ordenación y las enseñanzas mínimas de la Educación Secundaria Obligatoria, 198. (2022). Available online: <https://www.boe.es/eli/es/rd/2022/03/29/217/con> (accessed on 26 March 2026).
- Redecker, C. (2017). *European framework for the digital competence of educators: DigCompEdu* (Y. Punie, Ed.). European Commission. Joint Research Centre. Publications Office. Available online: <https://data.europa.eu/doi/10.2760/178382> (accessed on 26 March 2026).
- Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828, 2024/1689. (2024). Available online: <https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng> (accessed on 26 March 2026).
- Rincón, E., Jiménez, D., Aguilar, L., Flórez, J., Tapia, Á., & Peñuela, C. (2025). Mapping the use of artificial intelligence in medical education: A scoping review. *BMC Medical Education*, 25, 526. [CrossRef] [PubMed]
- Rutecka, P., Cicha, K., Rizun, M., & Strzelecki, A. (2025). Generative AI in curriculum design: Empirical insights into model performance and educational constraints. *IEEE Transactions on Learning Technologies*, 18, 757–768. [CrossRef]
- Sapawi, M. S. M., & Yusoff, N. M. R. (2025). Artificial intelligence in curriculum development: A global systematic review of trends, challenges, and strategic directions. *Journal of Curriculum Studies Research*, 7(2), 466–497. [CrossRef]
- Tsopra, R., Peiffer-Smadja, N., Charlier, C., Campeotto, F., Lemogne, C., Ruszniewski, P., Vivien, B., & Burgun, A. (2023). Putting undergraduate medical students in AI-CDSS designers' shoes: An innovative teaching method to develop digital health critical thinking. *International Journal of Medical Informatics*, 171, 104980. [CrossRef] [PubMed]
- Ullah, R., Hashim, F., Bandiali, M., & Akbar, A. (2025). Artificial intelligence in curriculum design a roadmap for adaptive and personalized learning in higher education. *The Critical Review of Social Sciences Studies*, 3(3), 304–322. [CrossRef]
- Wei, P., Wang, X., & Dong, H. (2023). The impact of automated writing evaluation on second language writing skills of Chinese EFL learners: A randomized controlled trial. *Frontiers in Psychology*, 14, 1249991. [CrossRef] [PubMed]
- 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(1), 40. [CrossRef]
- Zawacki-Richter, O., Marín, V., 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(1), 39. [CrossRef]
- Zhang, Y., & Tang, Q. (2025). Integrating AI-generated content tools in higher education: A comparative analysis of interdisciplinary learning outcomes. *Scientific Reports*, 15, 25802. [CrossRef] [PubMed]
- Zheng, T. (2025). AI education in higher education: A taxonomy for curriculum reform and the mission of knowledge. *arXiv*, arXiv:2509.23363. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.