e-ISSN: 2171-7966
REVISTA DE MEDIOS Y EDUCACIÓN
MEDIA & EDUCATION JOURNAL

Assessing teachers' perceptions of AI and Universal Design for Learning: design and validation

Evaluación de las percepciones de los docentes sobre la IA y el Diseño Universal para el Aprendizaje: diseño y validación

Dra. Antonia Cascales-Martínez
Profesora Asociada. Universidad de Murcia. España

D. Ludovico Vespasiani
University of Salerno. Italia

Received: 2025/03/24; **Revised:** 2025/04/30; **Accepted:** 2025/07/29; **Online First:** 2025/08/23; **Published:** 2025/09/01

ABSTRACT
The integration of Artificial Intelligence (AI) in education offers potential for inclusive practices aligned with Universal Design for Learning (UDL). However, teachers' perceptions of AI remain underexplored. This non-experimental, descriptive, and cross-sectional study developed and validated a questionnaire to assess teachers' views on AI in inclusive education. The instrument was constructed through a literature review and expert evaluation using the Aggregated Judgments Technique. Eight experts assessed content validity, and a pilot study was conducted with 55 in-service teachers from public schools in four Italian regions. The instrument demonstrated strong psychometric properties: high internal consistency (α = .90–.96), excellent content validity (V ≥ .90), and acceptable structural adequacy (KMO = .74). Most teachers reported basic digital skills, limited AI use, and minimal formal training in UDL, though they acknowledged AI's inclusive potential. The validated instrument is reliable and suitable for evaluating teachers' perceptions of AI as a tool for inclusive education. Findings highlight the need for targeted training and support to enhance AI integration in schools.

RESUMEN
La incorporación de la Inteligencia Artificial (IA) en la educación representa una oportunidad para fomentar prácticas inclusivas, en línea con el Diseño Universal para el Aprendizaje (DUA). Sin embargo, las percepciones docentes sobre su uso siguen siendo poco exploradas. Este estudio no experimental, descriptivo y transversal desarrolló y validó un cuestionario para evaluar dichas percepciones. Se utilizó una revisión bibliográfica y la Técnica de Juicios Agregados. Ocho personas expertas participaron en la validación de contenido, y se realizó una prueba piloto con 55 docentes en servicio de escuelas públicas en cuatro regiones italianas. El instrumento mostró alta consistencia interna (α = .90–.96), excelente validez de contenido (V ≥ .90) y adecuación estructural aceptable (KMO = .74). La mayoría del profesorado reportó competencias digitales básicas, uso limitado de IA y escasa formación en DUA, aunque reconoció el potencial inclusivo de la IA. El cuestionario validado es fiable y pertinente para evaluar las percepciones docentes sobre la IA como herramienta para una educación inclusiva. Los resultados subrayan la necesidad de formación específica y apoyo institucional para una integración efectiva de la IA en contextos escolares.

KEYWORDS · PALABRAS CLAVES
artificial intelligence; inclusive education; Universal Design for Learning; teachers.
inteligencia artificial; educación inclusiva; diseño universal de aprendizaje; docente.

1. Introduction

Universal Design for Learning (UDL) is an educational model based on neuroscience that aims to build a learning context without barriers and accessibility to all students (CAST, 2018). In recent years, the theoretical and practical framework of UDL has evolved, which has evolved into three main versions: UDL 1.0, UDL 2.0, and UDL 3.0, each characterized by an increasing level of integration of digital technologies and Artificial Intelligence (AI) into teaching processes. UDL 3.0 (CAST, 2024), the most recent version, explicitly integrates AI and data analysis, enabling real-time adaptation of content and teaching strategies in response to the individual needs of students. This evolution has led to learning systems capable of providing immediate feedback, supporting self-regulation, and ensuring a higher level of inclusivity and personalization (Al-Azawei et al., 2016; López-Ibáñez et al., 2023; Zhang et al., 2024).

This approach promotes proactive planning, which considers students' differences and promotes the idea that all students can learn successfully when provided with suitable resources and supports. The principles of UDL focus on three neural networks: the "what" of learning (representation), the "how" (expression) and the "why" (motivation), permitting different forms of access, participation and demonstration of knowledge (CAST, 2018). Furthermore, the UDL approach considers students' differences as a common situation and not an exception, providing strategies that guarantee participation and inclusion for every type of learner (CAST, 2018, 2024; Rusconi & Squillaci, 2023; Zhang et al., 2024). Integrating digital technologies and AI, especially within the framework of UDL 3.0, serves as a key lever for creating truly inclusive, flexible, and evidence-based learning environments. Furthermore, involving technologies represents a viable solution for implementing UDL principles, as it allows teachers to personalize content, promote self-regulation, and monitor student progress in a formative manner (Rose & Meyer, 2002; Tarconish et al., 2023). It enables the adaptation of content by offering different types of representation (Hall et al., 2012), facilitates the expression of knowledge through tools such as recordings, presentations, and text editors (Okolo & Diedrich, 2014), and fosters motivation through instant feedback, choice options, and gamification (Katz, 2013).

Moreover, UDL is effective when its adoption is driven by a clear pedagogical aim (Edyburn, 2013; López-Ibáñez et al., 2023), allowing educators to anticipate students' needs and eliminate barriers before they arise (Espada-Chavarría et al., 2023),

Through natural language processing algorithms and models calibrated to account for difficulty, AI-based learning systems can provide real-time adaptation in response to textual complexity or cognitive load (Fadel et al., 2019). Similarly, intelligent tutoring systems offer instant and differentiated feedback, boosting motivation and reducing metacognitive barriers (Luckin et al., 2016). Empirical evidence also shows that these systems enhance persistence and performance among students with special educational needs (Al-Azawein et al., 2016; Almeqdad et al., 2023).

AI makes it possible enables multimodal learning analytics, such as long pauses, irregular clicks, or lack of responses and informing teachers about moments of demotivation. This supports real-time interventions aligned with the principle of multimodal engagement (Hyatt & Owenz, 2024; UNESCO, 2021). In terms of accessibility, advances in computer vision enable automatic captioning, image descriptions, and sign language translation (Saborío-Taylor & Rojas-Ramírez, 2024).

However, it is crucial to emphasize that the use of AI in UDL must be grounded in transparent algorithmic criteria, data protection, and human oversight, to avoid reinforcing structural inequalities (European Commission, 2023).

In summary, AI should not be seen as a replacement for teaching, but a tool to enhance it. It can help realize the ideal of a truly responsive, inclusive, and evidence-based curriculum by offering granular adjustments, timely feedback, and universal accessibility (Cascales et al., 2024; Van Leeuwen & Rummel, 2022). Therefore, integrating AI into universal learning practices is not a technological luxury but an ethical and pedagogical imperative for achieving an educational system that ensures both excellence and equity.

For this reason, understanding teachers' perceptions of AI integration in educational contexts is essential for identifying its potential to foster more inclusive pedagogical practices and to derive educational insights that enhance the quality of the system.

The objective of the present research is to design and validate an *ad hoc* instrument to assess teachers' perceptions and experiences regarding the implementation of AI. The central focus is to determine whether this advanced technology can significantly contribute to a more inclusive teaching approach, ensuring internal coherence and content validity.

2. Methodology

This study adopts a non-experimental, descriptive, and cross-sectional design (Hernández-Sampieri et al., 2014), as the variables were neither manipulated nor randomly assigned to subjects. Instead, data were collected simultaneously, with a particular focus on expert judgments, to define and document precisely the instrument's properties in terms of validity and internal consistency (Cubo Delgado et al., 2011; Montero & León, 2007; Torrado, 2014).

2.1 Participant

During the content validation phase of the questionnaire, eight experts participated. They were selected based on their demonstrated expertise in the fields of AI and UDL, as well as their qualifications to assess the relevance, clarity, and applicability of the questionnaire items in non-university educational contexts. The panel was composed predominantly of female(5) experts.

Regarding academic qualifications, 25% of the experts held a PhD, 50% held either a Master's degree or a relevant specialization (in inclusive education, educational technologies, or statistics), and the remaining 25% held a Bachelor's degree in Education or Pedagogy.

Professionally, most experts had direct teaching experience in both primary and secondary education, with specializations in areas such as linguistics, logic-mathematics, anthropology, and inclusive education. Others were involved in academic research, particularly in experimental pedagogy with a focus on data analysis, motivation, and educational well-being.

The pilot study population comprised in-service teachers (excluding university faculty) working in public educational institutions across four Italian regions: Campania, Lazio, Piemonte, and Toscana. The sample was obtained through non-probability convenience sampling based on availability. A total of 55 participants completed the initial version of the instrument.

The sample was predominantly female (85.5%), while 14.5% were male. The average age was 46.1 years, and the average teaching experience was 16.7 years. Academically, 63.6% held a Master's degree, 21.8% held a high school diploma, 10.9% held a Bachelor's degree, and 3.6% held a PhD.

In terms of teaching levels, most participants taught at the primary school level (63.6%), followed by upper secondary school (30.9%), kindergarten (3.6%), and lower secondary school (1.8%). The vast majority were employed in public schools (94.5%), while 5.5% worked in private institutions.

Regarding training in UDL, 72.7% reported not having received any specific training related to UDL in the past five years, whereas 27.3% had. Nonetheless, 92.7% indicated the use of digital support tools in their teaching. Specifically, 70.9% used them several times per week, 10.9% used them once per week, 12.7% used them occasionally, and 5.5% rarely used them. The application of UDL principles and related technologies was reported by 52.7% of participants. In 18.2% of cases, they were never applied, while in the remaining instances, they were used in specific contexts: learning difficulties (12.7%), special educational needs (7.3%), vulnerable situations (3.6%), or other circumstances (5.5%).

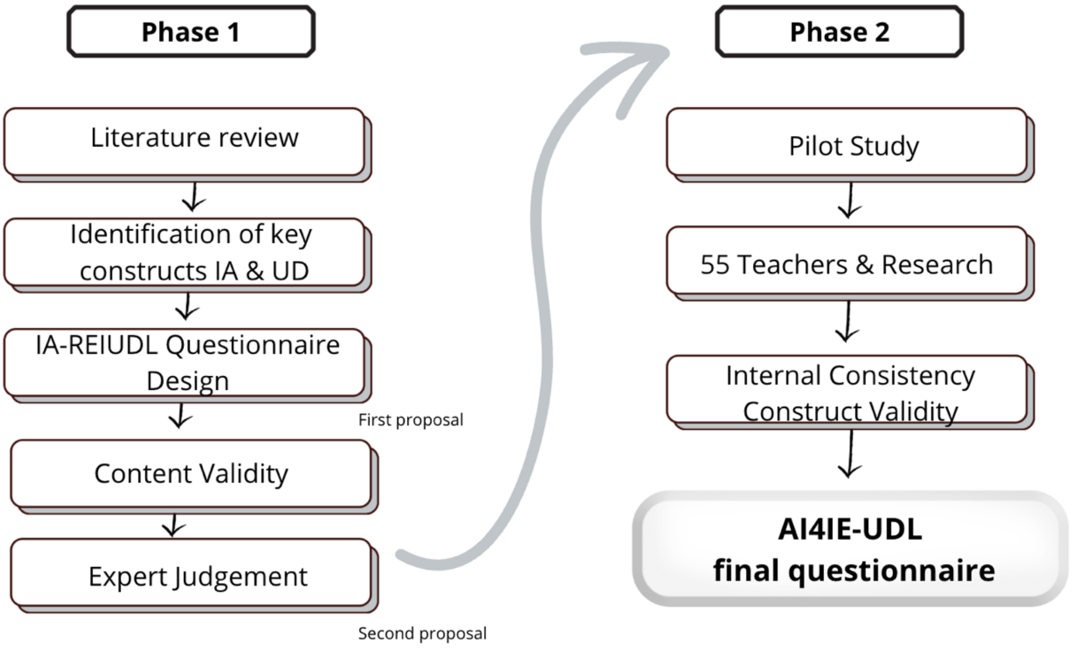
Regarding self-perceived digital competence, 47.3% rated themselves as beginners, 45.5% as intermediate, and 7.5% as advanced. Regarding familiarity with AI in educational contexts, 29.1% considered it non-existent or very limited, 25.5% had a basic level, 10.9% had an intermediate or advanced level, and 34.5% indicated that their knowledge had been acquired autonomously.

2.2 Instrument

The development of the questionnaire was approached using reference instruments developed for the use of AI in the educational context (Lopez Ros et al., 2024; Ng et al., 2024; Perkins et al., 2024). After this initial literature review, the instrument was constructed using the Aggregated Judgments Technique (Cabero & Llorente, 2013; Escobar & Cuervo, 2008). This process is illustrated in Figure 1.

Figure 1

Phases of instrumental research



Source: Own elaboration

Subsequently, content validation was initiated, along with an assessment of the instrument's reliability in terms of both stability and internal consistency.

The selected experts were asked to assign a score on a 5-point scale (quantitative evaluation) for each of the following dimensions: clarity of language, practical coherence, theoretical relevance, and sufficiency of content (Escobar & Cuervo, 2008).

Additionally, they were invited to provide further suggestions, if they deemed it necessary (qualitative evaluation). Responses were received within fifteen days. All evaluations and proposed improvements related to each item were compiled into a single document to facilitate the organization and implementation of revisions.

Once the content validation process and the revisions based on expert feedback were completed, the final version of the instrument was distributed to 55 in-service teachers who had agreed to participate in the study. This phase aimed to assess the instrument's reliability.

Finally, following the evaluation carried out by the experts, the final version of the questionnaire was established. It consists of four sections: the first gathers sociodemographic data, while the second comprises 29 items grouped into three dimensions, assessed using a five-point Likert scale, as shown in Table 1.

Table 1

The final questionnaire

Dimension	Content	Type of item and scale	Items number
I. Socio-demographic data	Collect basic teaching information -age, gender, title, experience, job context, type of job context, UDL training, ICT use, level of digital competence and AI familiarity	Close answers with multiple choice	12
II. AI impact	Assess the degree of agreement with statements about how AI contributes to equity, the removal of barriers, and the personalization of learning	5-point Likert scale (1 = Totally disagree… 5 = Tottally agree)	10
III. AI and UDL	Examines the practical integration of AI within the UDL framework: accessibility, adaptation, collaboration, and continuous improvement		9
IV. Teacher Satisfaction	It measures satisfaction, perception of institutional support, and the intention to delve deeper into AI for inclusive education		10

Notes: Own elaboration

The questionnaire was administered digitally, ensuring anonymity and confidentiality. Teachers received an institutional invitation with study information and completion instructions. Participation was voluntary and took approximately 15–20 minutes. Informed consent was obtained prior to participation.

2.3 Data analysis

Data analysis was conducted using IBM SPSS Statistics 29, following the steps outlined below:

- Descriptive analysis of the item scores
- Content Validity Index (CVI): This index was calculated for each item to assess its content validity. In addition, measures of stability and internal consistency were applied to evaluate the reliability of the instrument.
- Concordance analysis: An analysis of expert agreement was conducted for each dimension, using a correlational approach based on Kendall's W coefficient.
- Construct validity and reliability analysis: The construct validity and overall reliability of the instrument were examined. For all analyses, a significance level of .05 was adopted ($p < .05$).
-

3. Analysis and results

First, a comparative, descriptive, and interpretive analysis was conducted based on the evaluations provided by the expert judges for each aspect of the evaluation scale. This approach made it possible to identify all questionnaire items with an average score below the threshold value, which were subsequently revised.

The final version of the questionnaire was renamed IA-REIUDL from Artificial Intelligence to Strengthen Inclusive Education through Universal Design for Learning (AI4IE-UDL). Following the review, the items were refined to improve clarity, inclusivity, and alignment with UDL terminology. In the sociodemographic section, age was changed to an open-response field, an "Other" option was added under gender, equivalent academic levels were consolidated, and educational contexts (early childhood, primary, lower, and upper secondary) were specified. The scales evaluating UDL training, digital competence, and familiarity with AI were also simplified.

The question regarding technological resources was reformulated into a direct inquiry about the actual availability of technological support tools within the school. The section on the impact of AI was rewritten to emphasize how AI, when integrated with UDL, contributes to reducing barriers and addressing learner diversity.

In the teaching practices section, a redundant item was removed, and the remaining items were revised to highlight collaboration among colleagues, ongoing professional development, and active student involvement, all framed within an inclusive educational approach. Finally, the teacher satisfaction items were simplified to focus on the adequacy of support received, the creativity fostered by AI, and the willingness to share experiences with peers.

The reformulated items met the final criteria of clarity, coherence, sufficiency, and relevance. This refinement process resulted in the final version of the questionnaire, ensuring its precision and effectiveness for data collection within the scope of the research.

Tables 2, 3, and 4 present the content validity results, including Aiken's *V* coefficient and the 95% confidence interval for each item, in terms of clarity, coherence, theoretical relevance, and sufficiency. According to Hernández-Nieto (2011) and as supported by recent methodological studies (Maldonado-Suárez & Santoyo-Telles, 2024), values of Aiken's *V* equal to or greater than .90 are considered indicative of excellent content validity and expert agreementThe analysis of content validity for the "AI impact" dimension (Table 2), based on Aiken's V coefficient, demonstrated consistently high levels of expert agreement across the four evaluated criteria: sufficiency, relevance, clarity, and coherence. On average, the items scored .89 for sufficiency, .90 for relevance, .92 for clarity, and .89 for coherence, indicating strong content validity. Notably, item 2—focused on AI's role in enhancing student motivation—achieved the highest relevance score (V = 1.00), while item 1—related to equity through personalization—received the lowest (V = .79). Clarity was rated highest for item 2 (V = .96) and lowest for item 1 (V = .88). Coherence scores ranged from .75 (item 3) to .96 (item 1). These results confirm that the items are well-constructed and conceptually aligned with the intended dimensions, supporting the instrument's validity for assessing perceptions of AI's impact in inclusive educational contexts.

Table 2

AI impact: Aiken V Content Validity Indices and 95% Confidence Intervals

Dimension AI impact	Sufficiency V (CI 95 %)	Relevance V (CI 95 %)	Clarity V (CI 95 %)	Coherence V (CI 95 %)
1. I believe that AI strengthens equity through personalized teaching proposals	.83 (.67 – .96)	.79 (.53 – .98)	.88 (.70 – .99)	.96 (.88 – 1.00)
2. AI facilitates methods centered on student motivation	.96 (.88 – 1.00)	1.00 (1.00 – 1.00)	.96 (.88 – 1.00)	.96 (.88 – 1.00)
3. Integrating AI and UDL principles promotes inclusion for students with different learning needs and styles	.88 (.71 – 1.00)	.88 (.71 – 1.00)	.88 (.71 – 1.00)	.75 (.51 – .94)
4. I perceive advantages in content diversification when I use AI applications (chatbots, generators...)	.88 (.76 – .96)	.92 (.81 – 1.00)	.92 (.81 – 1.00)	.88 (.76 – .96)
5. I observe that AI improves information access when students need additional support, such as format or accessibility adjustments	.88 (.71 – 1.00)	.92 (.81 – 1.00)	.92 (.81 – 1.00)	.92 (.81 – 1.00)
6. AI tools with automatic feedback allow me to more accurately identify my students' needs	.88 (.71 – 1.00)	.92 (.81 – 1.00)	.92 (.81 – 1.00)	.88 (.71 – 1.00)
7. AI helps me offer activities that promote student autonomy, supporting self-regulation and decision-making in their learning	.92 (.81 – 1.00)	.88 (.71 – 1.00)	.92 (.81 – 1.00)	.88 (.71 – 1.00)
8. I value using AI to design learning goals tailored to diverse interests and student abilities	.88 (.71 – 1.00)	.92 (.81 – 1.00)	.92 (.81 – 1.00)	.92 (.81 – 1.00)
9. AI is useful for creating collaborative environments where students actively engage in their learning	.83 (.65 – .96)	.83 (.65 – .96)	.88 (.70 – .99)	.88 (.70 – .99)
10. I believe these technologies contribute to developing students' digital competencies	.96 (.88 – 1.00)	.92 (.81 – 1.00)	.96 (.88 – 1.00)	.92 (.81 – 1.00)

The content validity analysis for the “AI and UDL” dimension (Table 3), based on Aiken’s V coefficients, revealed high consistency in expert evaluations across the four criteria: sufficiency, relevance, clarity, and coherence. Overall, the values ranged from 0.79 to 0.96, indicating strong content validity. Item 15, which addresses the periodic evaluation of AI tools and corrective actions when accessibility or usability criteria are not met, received the highest scores in sufficiency (V = .96) and coherence (V = .96). Similarly, item 14, focused on how AI supports student self-regulation and diversified participation, achieved the highest relevance score (V = .96). On the other hand, items 13 and 18 showed the lowest coherence values (V = .79), though still within acceptable thresholds. All items exceeded the minimum acceptance threshold (V ≥ .70), supporting the conceptual and linguistic adequacy of the statements. These results confirm that the items designed to assess the integration of AI within the UDL framework are valid and appropriate for use in inclusive educational contexts. Table 3 presents the set of items assessing the integration of AI within UDL practices. All items show high Aiken V values, indicating strong expert agreement on the sufficiency, relevance, clarity, and coherence of the items. This suggests that the statements are well-formulated and considered valid representations of the intended constructs.

Table 3

AI and UDL: Aiken V Content Validity Indices and 95% Confidence Intervals

Dimension AI and UDL	Sufficiency V (CI 95 %)	Relevance V (CI 95 %)	Clarity V (CI 95 %)	Coherence V (CI 95 %)
11. I integrate AI into my teaching proposals to create flexible learning paths aligned with UDL principles	.88 (.78 – .96)	.88 (.78 – .96)	.88 (.78 – .96)	.83 (.67 – .96)
12. I analyze data generated by AI to adjust my strategies to different learning rhythms	.92 (.81 – 1.00)	.92 (.81 – 1.00)	.92 (.81 – 1.00)	.83 (.59 – 1.00)
13. I coordinate with other professionals (colleagues, trainers, principals, etc.), sharing resources and experiences to use AI inclusively and in alignment with UDL	.92 (.81 – 1.00)	.92 (0.81 – 1.00)	.92 (0.81 – 1.00)	.79 (.61 – .95)
14. I show students how AI can strengthen their self-regulation and diversify participation methods in the learning process	.88 (.71 – 1.00)	.96 (.88 – 1.00)	.88 (.71 – 1.00)	.96 (.88 – 1.00)
15. I periodically evaluate the effectiveness of the AI tools I use and take corrective action when they do not meet accessibility or usability criteria	.96 (.88 – 1.00)	.92 (.81 – 1.00)	.92 (.81 – 1.00)	.96 (.88 – 1.00)

Dimension AI and UDL	Sufficiency V (CI 95 %)	Relevance V (CI 95 %)	Clarity V (CI 95 %)	Coherence V (CI 95 %)
16. I seek training opportunities to deepen my understanding of applying UDL with AI	.92 (.81 – 1.00)	.92 (.81 – 1.00)	.92 (.81 – 1.00)	.96 (.88 – 1.00)
17. I encourage collaboration and idea exchange with colleagues to improve AI implementation from an inclusive perspective	.92 (.81 – 1.00)	.88 (.71 – 1.00)	.88 (.71 – 1.00)	.92 (.81 – 1.00)
18. I involve students in selecting AI solutions, encouraging participation and adaptation to their diverse needs	.88 (.71 – 1.00)	.88 (.71 – 1.00)	.88 (.71 – 1.00)	.79 (.61 – .95)
19. I verify that AI reduces barriers to accessing information, strengthening motivation and participation for all students	.88 (.71 – 1.00)	.92 (.81 – 1.00)	.88 (.71 – 1.00)	.92 (.81 – 1.00)

The content validity analysis for the “Teacher Satisfaction” dimension (Table 4), based on Aiken’s V coefficients, demonstrated consistently high expert agreement across all four criteria: sufficiency, relevance, clarity, and coherence. The average Aiken’s V values were .91 for sufficiency, .91 for relevance, .89 for clarity, and .92 for coherence, indicating strong content validity.

Item 20 received the highest sufficiency score (V = .96), while item 23 had the lowest (V = .83). In terms of relevance, item 20 also scored highest (V = .96), and item 21 lowest (V = .83). Clarity was rated highest for item 20 (V = .96) and lowest for item 23 (V = .79). Coherence scores ranged from 0.83 (item 23) to .96 (items 20, 24, 28, and 29). These results confirm the strong conceptual and linguistic adequacy of the items, supporting the instrument’s validity for assessing teacher satisfaction with AI integration in inclusive education.

Table 4

Teacher satisfaction: Aiken V Content Validity Indices and 95% Confidence Intervals

Dimension Teacher satisfaction	Sufficiency V (CI 95 %)	Relevance V (CI 95 %)	Clarity V (CI 95 %)	Coherence V (CI 95 %)
20. I I am satisfied with the results obtained by combining AI with inclusive strategies in my teaching practice	.96 (.88 – 1.00)	.96 (.88 – 1.00)	.96 (.88 – 1.00)	.96 (.88 – 1.00)

Dimension Teacher satisfaction	Sufficiency V (CI 95 %)	Relevance V (CI 95 %)	Clarity V (CI 95 %)	Coherence V (CI 95 %)
21. The time I invest in training on AI is offset by improvements in student learning	.88 (.71 – 1.00)	.83 (.59 – 1.00)	.83 (.59 – 1.00)	.88 (.71 – 1.00)
22. I believe that the available support (technical, pedagogical, financial) in my educational context is sufficient for effectively using AI within UDL V	.92 (.81 – 1.00)	.88 (.71 – 1.00)	.88 (.71 – 1.00)	.92 (.81 – 1.00)
23. I manage AI applications with confidence, enabling me to resolve technical or educational issues	.83 (.59 – 1.00)	.88 (.71 – 1.00)	.79 (.61 – .95)	.83 (.59 – 1.00)
24. . I perceive that my workplace values the use of AI to promote participation for all students	.96 (.88 – 1.00)	.92 (.81 – 1.00)	.92 (.81 – 1.00)	.96 (.88 – 1.00)
25. I appreciate AI as a key resource for more personalized teaching tailored to student diversity	.88 (.71 – 1.00)	.92 (.81 – 1.00)	.92 (.81 – 1.00)	.92 (.81 – 1.00)
26. . Introducing AI in the classroom enhances my creativity in designing new educational resources	.92 (.81 – 1.00)	.92 (.81 – 1.00)	.92 (.81 – 1.00)	.92 (.81 – 1.00)
27. I intend to deepen my use of AI to strengthen my inclusive teaching practice	.92 (.81 – 1.00)	.92 (.81 – 1.00)	.92 (.81 – 1.00)	.92 (.81 – 1.00)
28. My educational context and/or teaching team are willing to share experiences related to AI in education	.92 (.81 – 1.00)	.92 (.81 – 1.00)	.88 (.76 – .96)	.96 (.88 – 1.00)
29. I believe it is necessary to define clearer guidelines that consolidate the use of AI as an effective support for inclusive teaching	.92 (.81 – 1.00)	.92 (.81 – 1.00)	.92 (.81 – 1.00)	.96 (.88 – 1.00)

The experts who evaluated the questionnaire demonstrated a high level of agreement regarding the overall adequacy of the items. The intraclass correlation coefficient (ICC) ranged from .90 to .94, and more than 85% of the evaluations fell into the highest category

(4 = very adequate). When analyzing item ranking concordance using Kendall's W coefficient (with correction for ties [m = 8 judges, n = 10 items per dimension]), the resulting values were necessarily more modest. This is due to the limited variability in rankings, as most items received identical scores, leading to numerous ties and reduced rank variance:

- AI's impact on inclusive education: W = .171, χ^2 = 12.31, df = 9, p = .197
- AI and UDL: W = .198, χ^2 = 14.26, df = 9, p = .112
- Teacher satisfaction with AI and UDL: W = .261, χ^2 = 18.79, df = 9, p = .028

Thus, while the W values (ranging from .17 to .26) are not close to 1 and do not all reach statistical significance, this outcome is attributed to the limited variability resulting from the near-unanimous assignment of the highest score. Nevertheless, considering the absolute agreement, as reflected in the high percentage of maximum-category ratings and ICC values, it is reasonable to conclude that the expert judgment reliability is very high and that the items demonstrate excellent content validity, also supported by Aiken's V coefficients ≥ .90.

To assess construct validity and instrument reliability, data from the 55 participants in the pilot test were analyzed. Missing values (less than 3% of cells) were imputed using item means, a standard approach in exploratory studies with small samples. Table 5 summarises the descriptive statistics and the corrected discrimination index for each item in the questionnaire administered during the pilot phase. The mean scores range from 3.24 to 3.58, with standard deviations close to 1. Skewness values (−.59 to −.50) and kurtosis values (−.41 to .08) are moderate, indicating relatively symmetrical and mesokurtic distributions. The percentage of responses in the modal category ranges from 38% to 46%, well below the 80% threshold that would indicate excessive concentration. The discrimination index clearly exceeds the .30 cut-off point, suggesting that the items are effective in distinguishing between participants with varying scores on the overall scale. No item triggered any warnings according to the predefined criteria.

The KMO value was .74, indicating sufficient shared variance for factor analysis; following Kaiser (1974), this falls within the "mediocre to meritorious" range, as defined by Kaiser (1974). Bartlett's test of sphericity yielded χ^2 (406) = 1,470.36, p < .001, rejecting the null hypothesis of identity and confirming the suitability of factor analysis. In terms of internal consistency, Cronbach's alpha coefficients were .96, .92, and .90 for the subscales, and .94 for the global scale. These values indicate excellent internal consistency (Nunnally & Bernstein, 1994), confirming that the items within each dimension consistently measure the same construct and minimize measurement error.

Regarding convergent validity, the Average Variance Extracted (AVE) values were: D1 = .72, D2 = .58, D3 = .56, all exceeding the .50 threshold proposed by Hair et al. (2022). This indicates that each dimension explains more than 50% of the variance in its items. Composite Reliability (CR) ranged from .92 to .96, well above the recommended minimum of .70, confirming that true score variance predominates over error variance, thus supporting convergent validity. For discriminant validity, the square roots of AVE values (√AVE values ranged from .75 to .85) were greater than the absolute correlations between factors (|r| = .39–.48) in all combinations. This satisfies Hair et al.'s (2022) criterion, confirming that each construct is empirically distinct and that factor overlap does not compromise subscale interpretation.

Table 5

Descriptive Statistics and Item Discrimination Index

Items	Mean	Standard Deviation	Skewness	Kurtosis	% in Modal Category	Modal Category	Discrimination Index
1	3.24	1.04	-.59	-.41	43.6	4	.71
2	3.44	.96	-.58	-.01	45.5	4	.74
3	3.58	.99	-.57	.07	41.8	4	.59
4	3.33	1.02	-.58	.08	38.2	4	.70
5	3.47	1.09	-.50	-.34	38.2	4	.59
6	3.36	0.89	-.62	.28	43.6	4	.69
7	3.20	1.06	-.50	-.29	36.4	4	.69
8	3.49	1.10	-.89	.19	49.1	4	.56
9	3.38	1.11	-.71	-.21	45.5	4	.74
10	3.67	1.16	-.94	.07	49.1	4	.65
11	2.22	1.18	.66	-.62	34.5	1	.68
12	2.69	1.48	.20	-1.38	32.7	1	.65
13	1.91	1.01	1.07	.57	41.8	1	.30
14	2.09	1.19	.95	-.02	40.0	1	.57
15	2.33	1.33	.56	-.97	38.2	1	.64
16	2.35	1.28	.78	-.51	38.2	2	.65
17	2.25	1.27	.79	-.49	34.5	1	.60
18	2.16	1.33	.89	-.42	43.6	1	.53
19	2.47	1.36	.44	-1.09	32.7	1	.61
20	3.45	.96	-.63	.56	40.0	4	.85
21	3.18	.98	-.61	.27	45.5	3	.78
22	3.07	1.14	-.37	-.64	32.7	4	.45
23	3.15	1.01	-.51	-.05	41.8	3	.71
24	2.75	1.09	-.08	-.76	36.4	3	.48
25	3.44	1.08	-.58	-.10	38.2	4	.73
26	3.38	1.15	-.41	-.57	34.5	4	.75
27	3.67	1.17	-.88	-.06	45.5	4	.73
28	2.87	1.22	.12	-.88	29.1	3	.47
29	3.75	1.16	-.72	-.25	32.7	4	.56

In summary, the instrument clearly and reliably distinguishes the three theoretical dimensions it was designed to measure, as reported in Table 6.

Table 6

Discriminant validities (Fornell-Larcker criteria)

Couple of dimensions	|*r*|	√AVE minor	Compliant
D1-D2	.44	.76	yes
D1-D3	.48	.75	yes
D2-D3	.39	.75	yes

4. Discussion

Literature highlights how technologies and AI can be integrated with UDL principles (particularly in the latest version, UDL 3.0), promoting personalization, accessibility, and formative assessment in educational contexts (Hall et al., 2012; Katz, 2013; López-Ibáñez et al., 2023; Okolo & Diedrich, 2014; Rose & Meyer, 2002; Tarconish et al., 2023).These technologies enable different modalities of representation and expression, promote motivation through feedback and the possibility of choice, and allow for the proactive removal of barriers (Edyburn, 2013; Espada-Chavarría et al., 2023). Recent developments in AI provide real-time adaptivity, learning analytics, and accessibility tools. However, they also raise regulatory and ethical concerns (European Commission, 2023; Fadel et al., 2019; Hyatt & Owenz, 2024; López-Ibáñez et al., 2023; Luckin et al., 2016; Saborío-Taylor & Rojas-Ramírez, 2024; UNESCO, 2021).

Instrument validation confirmed high reliability and content validity, highlighting its suitability for measuring teachers' perceptions of AI and UDL integration, aligned with the theses of Hyatt & Owenz, 2024 and Saborío-Taylor and Rojas-Ramírez, 2024.

Most teachers report having a basic level of digital skills and limited experience with AI, although they express strong confidence in the inclusive potential of these technologies. Almost all teachers use digital tools regularly in their lessons, particularly with students with special educational needs or learning difficulties. However, there is a lack of teacher training courses specifically focused on UDL and AI, and fewer than one-third of teachers have attended such training. Teachers perceive digital technologies as useful for diversifying content and assessment in educational contexts, aligning with UDL principles. However, the effective integration of AI has not yet begun. Furthermore, the results indicate that teachers who feel more familiar with AI or have developed knowledge independently show greater confidence in its inclusive potential. This suggests that self-directed learning and initiative are key factors in addressing gaps left by institutional training. Despite the generally positive attitude, some challenges emerge. Teachers report a lack of clear guidelines, ethical frameworks, and institutional support for the responsible use of AI. Concerns about data protection, algorithmic transparency, and the risk of reinforcing existing inequalities are frequently cited, echoing findings in the literature. Moreover, the limited availability of AI-based accessibility tools (such as automatic subtitling or multimodal analytics) highlights the need for increased awareness and targeted investments in training and infrastructure.

The results emphasize the importance of continuous training that addresses not only technical skills but also the pedagogical and ethical dimensions of AI integration. Among the main ethical implications of the use of AI in education are the protection of students' personal data, the transparency and explainability of algorithms, the prevention of bias and discrimination, responsibility in automated decision-making, and the preservation of teachers' professional autonomy. It is also essential to ensure that the use of AI does not replace pedagogical judgement or the human connection that underpins the educational process. Teachers' openness to innovation, combined with their call for more structured support, underscores the need for systemic strategies to make AI a true enabler of inclusive and equitable education, avoiding new inequalities.

5. Conclusions

This study has led to the development and validation of a rigorous instrument, named AI4IE-UDL, designed to assess teachers' perceptions regarding the integration ofAI and UDL in educational contexts. The construction of the questionnaire was based on a systematic review of the literature, ensuring its alignment with the most current theoretical frameworks. Content validity was established through expert panel evaluations, which demonstrated high levels of agreement and reliability, supported by Aiken's V and the Intraclass Correlation Coefficient (ICC). Furthermore, statistical analyses confirmed strong internal consistency, validating the coherence of the items in measuring the intended theoretical constructs.

The instrument, which will be made available to the scientific community, represents a significant contribution to a rapidly evolving field. The integration of AI in education is driving profound transformations, and the availability of valid and reliable tools is essential for analyzing its impact from both inclusive and pedagogical perspectives.

From a practical standpoint, the findings highlight the need to design teacher training programs that explicitly incorporate UDL principles alongside the ethical and pedagogical use of AI. These programs should go beyond technical competencies to address critical dimensions such as data-informed decision-making, algorithmic fairness, and privacy protection. Additionally, it is essential to foster a culture of educational innovation that promotes self-directed learning, the exchange of best practices, and the development of professional learning communities focused on digital inclusion.

It is also recommended to strengthen collaboration among schools, universities, and research institutions to co-design technological solutions that address real classroom needs. Such partnerships can support the development of accessible tools, including adaptive feedback systems, automatic subtitling, and multimodal analytics, which accommodate learner diversity.

For future research, it is advisable to expand the application of the instrument to different educational levels and geographical contexts to validate its utility in diverse environments. It is also pertinent to deepen the analysis of its psychometric properties using Structural Equation Modeling (SEM) within the framework of Confirmatory Factor Analysis (CFA), which will allow for the comparison of exploratory validation results and the distribution of items within each evaluated dimension.

Another relevant line of research involves exploring the impact of AI on specific UDL dimensions, such as emotional self-regulation, executive functioning, and intrinsic motivation. Additionally, it is necessary to address the ethical implications of AI in education, particularly regarding system transparency, algorithmic bias, and the risks of digital exclusion.

Finally, this study underscores the importance of having an instrument capable of systematically identifying and analyzing the synergies between AI and UDL. A rigorous mapping of these relationships enables the recognition of the characteristics, motivations, and expectations of an increasingly diverse and connected student population, as well as the anticipation of barriers and opportunities in algorithm-mediated educational environments. This knowledge is essential for continuously adapting teaching practices and support services—from the personalization of AI-driven learning pathways to the provision of multiple means of representation and expression inherent to UDL—ensuring that higher education institutions effectively respond to the challenges posed by democratic societies (Rose & Meyer, 2002; UNESCO, 2023; Tarconish et al., 2023). As a prospective continuation of this work, future research should address its current limitations. This study acknowledges the use of a small convenience sample, mainly composed of women and primary school teachers. Subsequent studies are encouraged to expand validation to other educational contexts and professional profiles, conduct confirmatory analyses with larger and more diverse samples, and include group invariance testing to strengthen the generalizability of the findings.

Author Contributions
Conceptualization: A.C.-M. L.V.; Methodology: A.C.-M.; Validation: A.C.-M., L.V.,; Formal analysis: A.C.-M.; Investigation: A.C.-M. L.V.; Resources: A.C.-M.; Writing – original draft: A.C.-M. L.V.; Writing – review & editing: A.C.-M., Author 2; Supervision: A.C.-M. Project administration: A.C.-M.

Data Availability Statement
The data set used in this study is available at reasonable request to the corresponding author

AI Tools Usage Statement

To ensure the linguistic accuracy of the manuscript, Grammarly (Premium version) was used exclusively for grammar, spelling, and punctuation review. The functions employed were:

- Grammar Check: Detection and correction of grammatical errors.

- Spelling & Punctuation: Suggestions regarding spelling and punctuation.

- Clarity & Conciseness: Style recommendations to improve fluency without altering the conceptual content.

- Grammarly's intervention was strictly limited to the Introduction and Conclusions sections; it was not used in any other part of the manuscript nor for generating new content. All suggestions were manually reviewed and either accepted or rejected by the authors, who assume full responsibility for the final version of the document.

Ethics approval
Ethical approval and consent to participate The project was evaluated and approved by the Research Ethics Committee of the University of Murcia, ensuring that all ethical standards were respected. Informed consent was obtained from all participants included in the study.

Consent for publication
All authors have consented to the publication of the results obtained by means of the corresponding consent forms.

Conflicts of interest
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Al-Azawei, A., Serenelli, F., & Lundqvist, K. (2016). Universal Design for Learning (UDL): A Content Analysis of Peer-Reviewed Journal Papers from 2012 to 2015. *Journal of the Scholarship of Teaching and Learning, 16*, 39-56. https://doi.org/10.14434/josotl.v16i3.19295

Almeqdad, Q. I., Alodat, A. M., Alquraan, M. F., Mohaidat, M. A., & Al-Makhzoomy, A. K. (2023). The effectiveness of universal design for learning: A systematic review of the literature and metaanalysis. *Cogent Education, 10*(1). https://doi.org/10.1080/2331186X.2023.2218191

Cabero, J., & Llorente, M. C. (2013). La aplicación del juicio de experto como técnica de evaluación de las tecnologías de la información y la comunicación (TIC). *Revista de Tecnología de Información y Comunicación en Educación, 7*(2), 11-22. https://revistaeduweb.org/index.php/eduweb/article/view/206

Cascales Martínez, A., López Ros, S. P., & Gomariz Vicente, M. Á. (2024). Entre aulas y algoritmos: Validación de un cuestionario sobre la perspectiva docente ante la Inteligencia Artificial Generativa. En R. Satorre Cuerda (Ed.), *La docencia universitaria en tiempos de IA* (pp. 15–27). Octaedro. https://octaedro.com/libro/la-docencia-universitaria-en-tiempos-de-ia/

CAST. (2018). *Universal Design for Learning Guidelines version 2.2.* CAST. https://udlguidelines.cast.org/

CAST. (2024). *Universal Design for Learning Guidelines version 3.0. CAST.* https://udlguidelines.cast.org/

Cubo Delgado, S., Martín Marín, B., & Ramos Sánchez, J. L. (Eds.). (2011). *Métodos de investigación y análisis de datos en ciencias sociales y de la salud*. Ediciones Pirámide.

Edyburn, D. L. (2013). Critical issues in advancing the special education technology evidence-base. *Exceptional Children, 80*(1), 7-24.

Escobar Pérez, J., & Cuervo Martínez, Á. (2008). Validez de contenido y juicio de expertos: una aproximación a su utilización. *Avances en medición, 6*(1), 27-36.

European Commission**.** (2023, December 9). *Political agreement on the Artificial Intelligence Act reached by the co-legislators.* [Press release]. European Commission. https://ec.europa.eu/commission/presscorner/detail/en/ip_23_6473

Fadel, C., Holmes, W., & Bialik, M. (2019). Artificial Intelligence in Education Promises and Implications for Teaching and Learning. *Encyclopedia of Education and Information Technologies*.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A primer on partial least squares structural equation modeling (PLS-SEM)* (3rd ed.). SAGE.

Hall, T. E., Meyer, A., & Rose, D. H. (2012). *Universal Design for Learning in the Classroom: Practical Applications*. Guilford Press.

Hernández-Nieto, R. A. (2011). *Instrumento de medición de la validez de contenido: diseño y desarrollo de un procedimiento cuantitativo para la validación de contenido por juicio de expertos*. Universidad de Los Andes. Mérida, Venezuela.

Hernández-Sampieri, R., Fernández-Collado, C., & Baptista-Lucio, P. (2014). *Metodología de la investigación* (6.ª ed.). McGraw-Hill.

Espada-Chavarría, R., González-Montesino, R. H., López-Bastías, J. L., & Díaz-Vega, M. (2023). Universal Design for Learning and instruction: Effective strategies for inclusive higher education. *Education Sciences, 13*(6), 620. https://doi.org/10.3390/educsci13060620

Hyatt, S. E., & Owenz, M. B. (2024). Using Universal Design for Learning and artificial intelligence to support students with disabilities. *College Teaching*, 1-8.

Kaiser, H. F. (1974). *An index of factorial simplicity. Psychometrika, 39*(1), 31–36. https://doi.org/10.1007/BF02291575

Katz, J. (2013). The Three-Block Model of Universal Design for Learning (UDL): Engaging students in inclusive education. *Canadian Journal of Education, 36*(1), 153–194. https://doi.org/10.2307/canajeducrevucan.36.1.153

López-Ibáñez, F. J., Cascales Martinez, A., & Martínez-Segura, M. J. (2023). Diseño universal para el aprendizaje y TIC en el área de educación física: diseño y validación de una propuesta de intervención. *PUBLICACIONES*, *53*(3), 135–178. https://doi.org/10.30827/publicaciones.v53i3.23867

López Ros, S. P., Cascales Martínez, A., & Gomariz Vicente, M. A. (2024). Cuestionario sobre el conocimiento, actitud y percepción del profesorado hacia la Inteligencia Artificial Generativa. CAPIAG-P (Profesorado) [Conjunto de datos]. http://hdl.handle.net/10201/141368

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). Intelligence Unleashed: An Argument for AI in Education. Pearson. https://static.googleusercontent.com/media/edu.google.com/en//pdfs/Intelligence-Unleashed-Publication.pdf

Maldonado-Suárez, N., & Santoyo-Telles, F. (2024). *Validez de contenido por juicio de expertos: Integración cuantitativa y cualitativa en la construcción de instrumentos de medición. Revista de Investigación Educativa (REIRE)*. https://www.researchgate.net/publication/381947747

Montero, I., & León, O. G. (2007). *A guide for naming research studies in Psychology.* International Journal of clinical and Health psychology, 7(3), 847-862.

Ng, D. T. K., Wu, W., Leung, J. K. L., Chiu, T. K. F., & Chu, S. K. W. (2024). Design and validation of the AI literacy questionnaire: The affective, behavioural, cognitive and ethical approach. *British Journal of Educational Technology*, *55*(3), 1082-1104. https://doi.org/10.1111/bjet.13411

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.

Okolo, C. M., & Diedrich, J. (2014). Twenty-five years later: How is technology used in the education of students with disabilities? *Journal of Special Education Technology, 29*(1), 1–20. https://doi.org/10.1177/016264341402900101

Perkins, M., Furze, L., Roe, J., & MacVaugh, J. (2024). The Artificial Intelligence Assessment Scale (AIAS): A framework for ethical integration of generative AI in educational assessment. *Journal of University Teaching and Learning Practice*, *21*(6), 49-66. https://doi.org/10.53761/q3azde36

Rose, D. H., & Meyer, A. (2002). *Teaching every student in the digital age: Universal design for learning*. ASCD.

Rusconi, L., & Squillaci, M. (2023). *Effects of a Universal Design for Learning (UDL) Training Course on the Development Teachers' Competences: A Systematic Review*. Education Sciences, 13(5), 466. https://doi.org/10.3390/educsci13050466

Saborío-Taylor, S., & Rojas-Ramírez, F. (2024). Universal Design for Learning and artificial intelligence in the digital era: Fostering inclusion and autonomous learning. *International Journal of Professional Development, Learners and Learning, 6*(2).

Tarconish, E., Scott, S., Banerjee, M., & Lombardi, A. (2023). *Universal Design for Instruction & Learning in Higher Education: Where Have We Been and Where Are We Headed? Journal of*

Postsecondary Education and Disability, 36(3), 207-223. https://files.eric.ed.gov/fulltext/EJ1435143.pdf

Torrado, M. (2014). Estudios de encuesta. En A. Bisquerra (Ed.), *Metodología de la investigación educativa* (4.ª ed., pp. 230-257). Editorial La Muralla.

UNESCO. (2021). *Inteligencia artificial*. En UNESCO**.** Recuperado el 30 de abril de 2025 de https://www.unesco.org/es/artificial-intelligence

UNESCO. (2023). Guidance for generative AI in education and research. United Nations Educational, Scientific and Cultural Organization.

Van Leeuwen, A., & Rummel, N. (2022). The function of teacher dashboards depends on the amount of time pressure in the classroom situation: Results from teacher interviews and an experimental study. *Unterrichtswissenschaft, 50*(4), 561-588. https://doi.org/10.1007/s42010-022-00156-9

Zhang, L., Carter, R. A., Jr., Greene, J. A., & Bernacki, M. L. (2024). Unraveling challenges with the implementation of Universal Design for Learning: A systematic literature review. *Educational Psychology Review, 36*(1), Article 35. https://doi.org/10.1007/s10648-024-09860-7

How to cite:

Cascales-Martínez, A., & Vespasiani, L. (2025). M.T., del Río, D., Teira-Serrano, C., & Cadime, I. (2025). Assessing teachers' perceptions of AI and Universal Design for Learning: design and validation [Evaluación de las percepciones de los docentes sobre la IA y el Diseño Universal para el Aprendizaje: diseño y validación]. *Pixel-Bit, Revista de Medios y Educación*, *74,* art.9 https://doi.org/10.12795/pixelbit.116690