Journal of Inclusive Methodology and Technology in Learning and Teaching
ISSN 2785-5104
Anno 6 n. 4 (2026)

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

Engaging teachers as designers: toward a validated checklist for AI-supported inclusive instructional design

Arianna Marras [1]*, Elena Pizzuti [2] and Angela Pasqualotto [2]

[1] Department of Humanities, Languages and Cultural Heritage, University of Cagliari, Cagliari, Italy; arianna.marras@unica.it
[2] Department of Education and Learning, University of Applied Sciences and Arts of Southern Switzerland (SUPSI), Locarno, Switzerland; elena.pizzuti@edu.ti.ch; angela.pasqualotto@supsi.ch
* Correspondence: arianna.marras@unica.it

Abstract: Generative artificial intelligence (AI) supports inclusive instructional design only when anchored in explicit design intentions and teachers' professional judgment. This study reports a first practitioner appraisal of a 20-indicator checklist that operationalizes the four phases of the AI-Enhanced Design Framework and was derived from a documented action-research diary. Twenty-five teachers and support professionals in Canton Ticino (Switzerland) used the checklist to redesign one of their own activities with generative AI during a professional development workshop, then rated each indicator for clarity and perceived utility on five-point scales, judged the checklist as a whole, and answered open questions. Ratings were generally high (item means 3.52–4.52), and clarity and utility were closely associated across indicators (Spearman's ϱ=.83). Seven indicators, mostly concerning universal design and differentiation, reached 80% agreement on both dimensions, whereas indicators relying on information about students (diagnostic pre-assessment, use of formative data, grouping by need) were among the least endorsed. Indicators of the universal design phase were rated higher than those of the differentiation and resilience phases. Participants endorsed the checklist's support for design intentionality but were more reserved about its length and time sustainability. The findings identify targeted revision priorities rather than a settled instrument.

Keywords: artificial intelligence; inclusive instructional design; universal design for learning; teacher professional development; checklist development

Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction[1]

The rapid diffusion of generative artificial intelligence (AI) in educational contexts has substantially expanded the possibilities available to teachers for producing, adapting, and differentiating instructional materials. Generative systems can support tasks such as summarization, translation, content adaptation, assessment, and the personalization of learning activities (Ray, 2023; Yan et al.,

[1] The authors jointly conceived the overall structure of the article. The work is the result of a collaborative effort. Specifically, Section 1 was written by Elena Pizzuti; Sections 2, 4, and 5 by Arianna Marras; and Section 3 by Angela Pasqualotto. All authors have approved the final version submitted to the journal.

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

2024). These affordances are particularly relevant from an inclusive education perspective, as AI can facilitate multimodal access to content, the generation of differentiated resources, and the adaptation of instructional activities to learner variability (Marras et al., 2025).

At the same time, technological affordances cannot be assumed to translate automatically into improved instructional quality. Concerns related to output accuracy, algorithmic bias, data protection, and overreliance on automated systems require sustained critical and professional mediation by teachers (Kasneci et al., 2023; Yan et al., 2024). Recent research further suggests that the educational legitimacy of AI depends on its alignment with pedagogical goals, professional control, institutional feasibility, and ethical accountability (Marras & Pasqualotto, 2026). This makes teacher professional judgment central to AI-supported instructional design. The relevant question is therefore not simply whether AI can generate pedagogically useful materials, but how teachers translate established pedagogical principles into concrete decisions when interacting with generative systems. The present study addresses this question by operationalizing an AI-supported instructional design framework into observable indicators of teacher decision-making. Specifically, it examines the development of a checklist intended to make teachers' design reasoning more explicit, documentable, and open to reflection.

1.1. Theoretical Framework for AI-Supported Inclusive Instructional Design

The main theoretical reference adopted in the present study is the AI-Enhanced Design Framework developed by Tucker and Novak (2025), which positions AI within a pedagogically structured instructional design process rather than as an autonomous tool or content generator. The framework organizes design into four phases: (1) Identify Desired Results, focused on learning outcomes and assessment alignment; (2) Design Equitable Learning Experiences, aimed at reducing barriers and providing flexible learning opportunities; (3) Adapt and Personalize Instruction and Support, based on evidence-informed differentiation; and (4) Cultivate Resilient Learners, addressing self-awareness, adaptability, perseverance, and social support (Tucker & Novak, 2025). These phases integrate established pedagogical traditions: backward design within Understanding by Design (Wiggins & McTighe, 2005), Universal Design for Learning (UDL; CAST, 2024), Differentiated Instruction (Tomlinson, 2014), and social-emotional and self-regulatory dimensions of learning. Within this architecture, UDL and Differentiated Instruction are especially relevant to inclusive design. UDL treats learner variability as an ordinary condition of teaching and promotes the anticipatory reduction of barriers through multiple means of engagement, representation, and action and expression (CAST, 2024). Differentiated Instruction complements this proactive logic by adapting content, processes, products, and support in response to evidence about learners' needs and progress (Tomlinson, 2014). The two approaches can therefore be understood as complementary: UDL supports the design of flexible environments, while differentiation supports ongoing instructional adaptation.

This pedagogical orientation is also consistent with broader models of teacher professional knowledge. The Technological Pedagogical Content Knowledge (TPACK) framework emphasizes that meaningful technology integration depends on the interaction of technological, pedagogical, and disciplinary knowledge rather

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

than on technical proficiency alone (Mishra & Koehler, 2006; Mishra et al., 2023). Applied to generative AI, this means that competence cannot be reduced to prompt formulation or tool use, but includes the ability to evaluate technological affordances and limitations, align them with instructional aims, and exercise pedagogical and ethical judgment (Miao & Cukurova, 2024; Marras et al., 2025).

From this perspective, AI does not constitute a pedagogical intervention in itself. Its educational value emerges through teachers' professional mediation. Its acceptance by teachers is accordingly conditional (Marras & Pasqualotto, 2026). AI-generated outputs should therefore be treated as provisional resources requiring verification, contextualization, adaptation, or rejection before instructional use.

The theoretical gap addressed in this study lies precisely between pedagogical principles and observable practice. Existing frameworks specify how inclusive and differentiated design should be oriented, but provide less guidance on how these principles become concrete teacher decisions during interaction with generative AI. Accordingly, the relevant unit of analysis is not the AI-generated artifact alone, but the professional decision through which it is requested, evaluated, modified, or integrated. To address this gap, the present study operationalizes the AI-Enhanced Design Framework into observable indicators of teacher decision-making. Building on previous action-research work (Pizzuti, 2026), in which a 45-indicator grid was derived from the framework and its use documented in a design diary, the present study selected 20 indicators to form a checklist structured around the four phases of the framework. The instrument is therefore intended not as a direct transcription of the framework, but as an operational bridge between pedagogical theory and professional action, making teachers' design reasoning more explicit and observable.

1.2. Aim and Research Questions

The study aimed to obtain a first practitioner appraisal of a 20-indicator checklist after authentic use, treating teachers as designers whose judgments inform the revision of the instrument. RQ1: How do teachers rate the clarity and perceived utility of each indicator, and do ratings differ across phases? RQ2: Do ratings vary by role, experience, prior AI use, and UDL familiarity? RQ3: How do teachers judge the checklist as a whole? RQ4: Which indicators do teachers find most useful or problematic, what do they see as missing, and how do they view a dedicated family of critical-ethical indicators?

2. Materials and Methods

2.1. Setting and Participants

The study took place in a workshop of the official professional development program of Canton Ticino (Atelier DECS 2026), for teachers of the second and third cycles of compulsory schooling. The workshop was held in two sessions (n=12 and n=13). Enrollment was self-selected and participation in the study voluntary. Twenty-five participants completed the questionnaire, with no missing responses on closed items: 17 classroom or subject teachers, seven special-education and support professionals, and one school leader (excluded from role comparisons). Twelve had more than 15 years of experience, nine 6–15, and four 0–5; five had never used AI to prepare materials, 14 occasionally, and six regularly; 13 used UDL in their design, seven knew it by name only, and five did not know it.

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

2.2. Checklist and Procedure

For each phase, five indicators were selected from the grid, prioritizing those activated in the diary (by frequency) and completing phases with theoretically central indicators where activations were insufficient. The only student-facing indicator activated in the diary was excluded, consistent with the greater legitimacy of teacher-facing uses (Marras & Pasqualotto, 2026). Indicators are listed in Table 1 (translated from Italian). The two-hour workshop combined a short theoretical framing, a guided demonstration of one documented action (an easy-to-read version of a history text, indicator I29, from prompt to verified and integrated material; Pizzuti, 2026), and about 45 minutes of work, in pairs or individually, in which participants redesigned one of their own activities. A worksheet required an unchanged learning objective, anonymous functional descriptions of learner profiles, foreseeable barriers, the choice of two to three indicators, prompt drafting, and verification of outputs for accuracy, accessibility, and data protection. No personal student data were entered.

2.3. Questionnaire and Analysis

The anonymous questionnaire collected background variables and asked participants to rate each indicator on clarity (1=not clear at all; 5=completely clear) and perceived utility for orienting AI-supported design (1=not useful at all; 5=very useful), a dual scale meant to separate poorly worded indicators from marginal ones. All participants rated all 20 indicators. Six statements on the checklist as a whole (C1–C6) were rated from 1 (strongly disagree) to 5 (strongly agree), and open questions addressed useful (D1) and problematic indicators (D2), missing dimensions (D3), a transversal family of critical-ethical indicators (D4, yes/no/don't know), and conditions of use (D5). Submission constituted consent; data were processed in aggregate form.

As agreement criterion we adopted at least 80% of ratings of 4–5 (20 of 25), in line with the .78 threshold of the item-level content validity index (Lynn, 1986; Polit & Beck, 2006); since raters were practitioners rather than content experts, the criterion sets revision priorities rather than accepting or rejecting items. Clarity and utility were compared with Wilcoxon signed-rank tests and Spearman's correlation; phase means with Friedman tests (Kendall's W) and Holm-corrected Wilcoxon tests; and groups with Mann–Whitney and Kruskal–Wallis tests, with Holm correction across 25 tests. Open responses were coded inductively by indicator, phase, and theme; respondents are identified as P1–P25.

3. Results

3.1. Indicator Ratings

Overall, 77.2% of ratings were 4 or 5 and 6.5% were 1 or 2; item means ranged from 3.52 to 4.52 for clarity and from 3.64 to 4.56 for utility (Table 1). Clarity and utility were closely associated across indicators ($\rho = .83$, $p < .001$), and no within-indicator difference was significant (all $p \geq .07$). Seven indicators met the criterion on both dimensions (I22, I25, I29, I40, I62, I78, I101; Figure 1); two were clear but less useful (I34, I108) and one useful but less clear (I5). Ten fell below the criterion on both: four of five Phase 1 indicators, three in Phase 3 (I56, I61, I71), three in Phase 4 (I93, I95, I105), and none in Phase 2. Utility agreement was lowest for

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

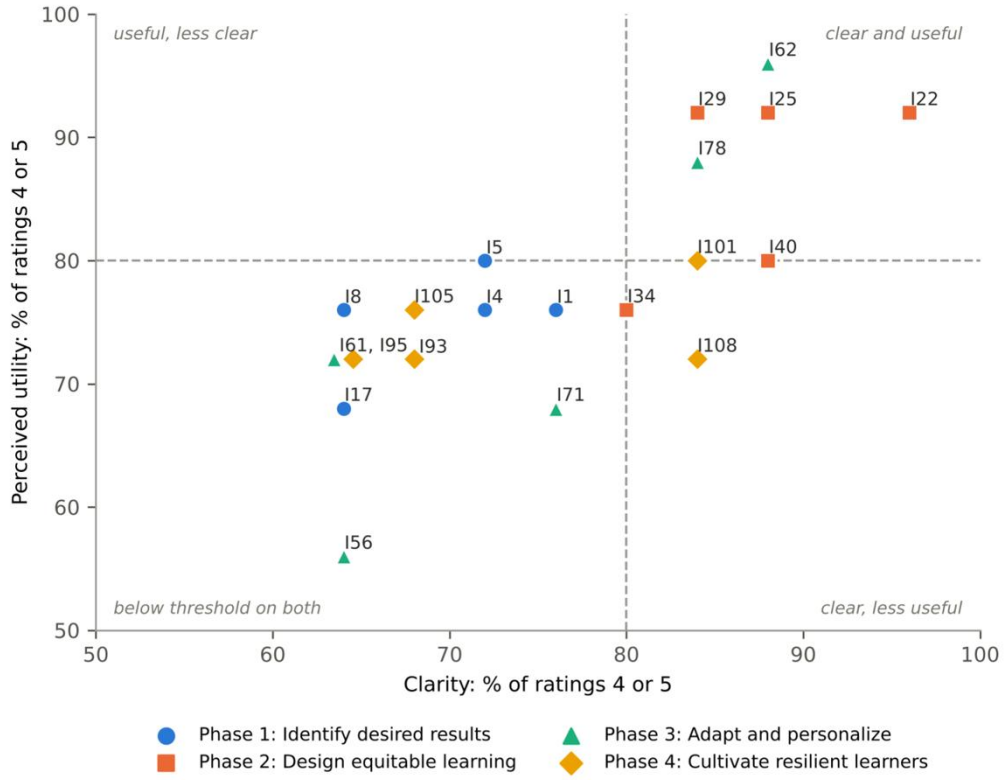
diagnostic pre-assessment (I56: 56%; clarity 64%), and mean clarity was lowest for the use of formative data (I61: M = 3.52).

Table 1. Indicators and ratings of clarity and utility (N = 25).

Code	Indicator	Clarity *M* (*SD*)	Clarity %	Utility *M* (*SD*)	Utility %
Phase 1: Identify Desired Results					
I1	Rewording learning objectives in accessible language and sharing them with students	4.16 (0.90)	76	4.24 (0.93)	76
I4	Generating or revising assessment rubrics directly linked to the stated learning outcomes	4.00 (0.96)	72	4.04 (0.84)	76
I5	Checking internal coherence among objectives, activities, and assessment criteria during design	4.04 (0.89)	72	4.20 (0.76)	**80**
I8	Structuring the objectives, evidence, and activities sequence of the unit (backward design)	3.88 (1.05)	64	4.16 (0.90)	76
I17	Documenting the use of AI across design phases	4.00 (0.96)	64	3.96 (1.06)	68
Phase 2: Design Equitable Learning Experiences					
I22	Designing materials in different formats (text, audio, images) with semantic coherence and accessibility	4.52 (0.59)	**96**	4.56 (0.77)	**92**
I25	Generating opening prompts and questions that activate prior knowledge consistent with the objectives	4.44 (0.82)	**88**	4.44 (0.65)	**92**
I29	Designing simplified or easy-to-read versions of texts	4.44 (0.77)	**84**	4.48 (0.77)	**92**
I34	Generating maps or outline drafts as a basis for guided activities (scaffolding)	4.20 (0.76)	**80**	4.12 (0.97)	76
I40	Generating planners, checklists, and reminders that pace the phases of a task	4.28 (0.89)	**88**	4.16 (0.94)	**80**
Phase 3: Adapt and Personalize Instruction and Support					
I56	Generating initial diagnostic tools (pre-assessment) targeted to the unit objectives	3.68 (0.99)	64	3.64 (0.86)	56
I61	Using formative data to adapt timing, materials, and supports	3.52 (1.05)	64	3.84 (1.07)	72
I62	Generating variants of materials and activities graded by mastery level and complexity	4.24 (0.78)	**88**	4.48 (0.59)	**96**
I71	Identifying groups with similar needs and planning targeted instruction	4.00 (1.04)	76	3.76 (0.93)	68
I78	Rewording texts, examples, and resources to ensure plurality of perspectives and accessibility	4.16 (0.80)	**84**	4.32 (0.80)	**88**
Phase 4: Cultivate Resilient Learners					
I93	Co-designing adaptations or variants of content to include plural viewpoints	3.88 (1.01)	68	4.04 (0.89)	72
I95	Generating scenarios or case studies that foster empathic perspective-taking and cooperation	3.76 (1.13)	64	3.92 (1.00)	72
I101	Revising language and materials from an inclusive and positive perspective	4.24 (0.83)	**84**	4.28 (0.89)	**80**
I105	Co-designing self-assessment tools that integrate metacognitive and decision-making dimensions	3.96 (0.89)	68	4.04 (0.84)	76
I108	Co-designing rubrics or shared rules from students' contributions	4.12 (0.88)	**84**	3.96 (0.93)	72

Note. Codes refer to the 45-indicator grid (Pizzuti, 2026). Ratings on five-point scales; % = proportion of ratings of 4 or 5; values of 80% or above are in bold.

Figure 1. Perceived clarity and perceived utility of the 20 indicators (N = 25). Each point is one indicator: the horizontal axis shows the percentage of respondents who rated its clarity 4 or 5, the vertical axis the percentage who rated its utility 4 or 5. Dashed lines mark the 80% agreement criterion (at least 20 of 25 respondents) and divide the plane into four profiles: clear and useful (upper right), useful but less clear (upper left), clear but less useful (lower right), and below criterion on both (lower left). Marker shape and color indicate the Design Path phase. I61 and I95 have identical values and overlap.



Phase means differed for clarity, $\chi^2(3) = 12.73$, $p = .005$, $W = .17$, and utility, $\chi^2(3) = 13.61$, $p = .003$, $W = .18$. Phase 2 was rated clearer ($M = 4.38$) than Phases 3 (3.92; adjusted $p = .014$) and 4 (3.99; $p = .035$), and more useful (4.35) than Phases 3 (4.01; $p = .012$) and 4 (4.05; $p = .030$); Phase 1 (4.02 and 4.12) did not differ from other phases after correction.

3.2. Global Judgments and Group Comparisons

Agreement was highest for support for more intentional design choices (C3: 88%; $M = 4.32$, $SD = 0.69$), followed by self-efficacy in evaluating AI outputs (C6: 76%; 4.12, 0.88), applicability (C1: 72%; 3.92, 1.04), and intention to reuse (C5: 72%; 4.00, 1.00), and was lower for time sustainability (C2: 60%; 3.64, 1.04) and the adequacy of 20 indicators (C4: 48%; 3.48, 1.26). No group difference by role, experience, prior AI use, or UDL familiarity survived Holm correction. Uncorrected differences emerged for experience (utility higher among teachers with 6–15 years: 4.47 vs. 3.81 and 3.98; $p = .035$) and session (clarity higher in the second session: 4.32 vs. 3.82; $p = .015$).

3.3. Open Responses

Among the eight respondents who named specific indicators (D1), I22 was cited most often (five), followed by I4, I62, and I78 (three each). Reasons included time saving (P19), accessibility (P22), and differentiating upward (P23). Eleven respondents found no indicator problematic (D2); others found Phase 1 indicators hard to translate into practice (P16, P18, P19), "formative data" unclear (P5), I56

and I61 too specific for regular classes (P21), and indicators requiring class information to be shared with the AI problematic (P11). Eighteen identified no missing dimension or could not yet answer (D3). On a critical-ethical family (D4), eight answered yes, five no, and 12 were undecided; opponents located this dimension in professional judgment ("already inherent in what we do as teachers," P19), and one supporter proposed placing privacy guidance within prompt formulation (P25). Time was the most cited condition of use (D5; seven respondents).

4. Discussion

This study provides a first practitioner appraisal of a checklist designed to operationalize the AI-Enhanced Design Framework into observable decisions for AI-supported inclusive instructional design. Overall, the instrument was perceived as understandable and useful, particularly as a support for more intentional design choices. At the same time, the findings indicate that the current version should be regarded as developmental rather than definitive: some dimensions of AI-supported design appear more readily translatable into actionable teacher practices than others, and usability depends not only on theoretical relevance but also on the cognitive and practical demands placed on teachers.

The strongest results concerned Phase 2, Design Equitable Learning Experiences. This is theoretically consistent with the UDL-oriented emphasis of the phase, which frames learner variability as an expected condition of teaching and promotes anticipatory design through multiple means of representation, engagement, and action and expression (CAST, 2024). Highly rated indicators such as multimodal materials, text simplification, activation of prior knowledge, and task planners correspond to concrete design operations and to affordances of generative AI that teachers can immediately recognize and enact (Marras et al., 2025). This may explain why accessibility-oriented uses were more readily endorsed than practices requiring more complex pedagogical inference. This pattern also converges with the action-research diary, in which UDL was the most represented area of activation (Pizzuti, 2026), although the comparison concerns areas rather than individual indicators.

By contrast, some of the weakest ratings concerned Phase 3, particularly diagnostic pre-assessment, formative data, and grouping by need. These findings should not be interpreted as questioning the relevance of Differentiated Instruction, which explicitly relies on evidence about learner readiness, progress, and needs (Tomlinson, 2014). Rather, they suggest that data-informed adaptation requires additional professional and contextual conditions, including access to appropriate evidence, confidence in interpreting it, and clarity about what information can responsibly be used with AI. Generative AI may make differentiation technically easier, but this increases rather than reduces the importance of professional judgment. This interpretation is consistent with the AI competency perspective proposed by Miao and Cukurova (2024) and with the conditional integration of AI reported by Marras and Pasqualotto (2026).

A similar pattern emerged for Phase 4, Cultivate Resilient Learners, whose indicators were rated positively overall but endorsed less consistently than those related to universal design.

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

Phase 1 also revealed a tension between theoretical relevance and immediate usability. Indicators related to backward design and the documentation of AI use were less strongly endorsed, and coherence checking (I5), although judged useful, was judged less clear, possibly because these indicators require higher-level design reasoning and are less directly associated with generative production. However, their role remains central: the logic of Understanding by Design requires learning activities and assessment to derive from clearly defined outcomes rather than from the availability of attractive resources (Wiggins & McTighe, 2005). Lower endorsement therefore suggests a need for clearer wording or additional scaffolding rather than theoretical removal.

The global judgments highlight a broader design issue. Participants valued the checklist primarily for supporting intentional design, while expressing more reservations about time sustainability and the number of indicators. This suggests that the instrument may be more useful as a reflective design scaffold than as a prescriptive sequence to be applied in full to every lesson. A future version could preserve the four-phase structure while allowing more selective or modular use. Opinions on a dedicated critical-ethical family were mixed, with most respondents undecided. Given the reservations about length, we propose treating critical and ethical judgment as a transversal dimension embedded in the verification steps of AI-supported design rather than as additional indicators. This choice is consistent with the UNESCO AI Competency Framework for Teachers (Miao & Cukurova, 2024) and with the emphasis placed by Marras et al. (2025) on pedagogical sensitivity, critical evaluation, and ethical awareness.

The dual rating scale discriminated only partially. Clarity and utility were strongly associated across indicators (ϱ = .83), and the two judgments diverged for only three indicators (I5, I34, and I108). Its discriminative value should therefore be re-examined after prolonged use.

Finally, the absence of statistically significant differences by professional role, experience, prior AI use, or UDL familiarity should be interpreted cautiously given the modest sample size and exploratory design. More broadly, however, the findings support the theoretical premise underlying the checklist: the relevant object of AI-supported instructional design is not the generated output itself, but the professional decision through which that output is requested, evaluated, adapted, and integrated. In this sense, the checklist may be most valuable not because it prescribes what teachers should ask AI to produce, but because it prompts them to articulate why a particular use of AI is pedagogically justified.

5. Conclusions

This study provides a first practitioner-based appraisal of a 20-indicator checklist designed to translate the AI-Enhanced Design Framework into observable decisions for AI-supported inclusive instructional design. The findings suggest that teachers perceive value in a tool that makes pedagogical intentions explicit, particularly when AI is used to design accessible, flexible, and differentiated learning resources. At the same time, the uneven performance of indicators across phases and participants' concerns about time and length indicate that the instrument requires further refinement rather than simple confirmation in its current form.

The findings reinforce a central premise of the study: generative AI does not itself produce inclusive instructional design. Its educational value depends on the

professional judgment through which teachers align technological affordances with learning objectives, learner variability, pedagogical principles, and ethical constraints (Tucker & Novak, 2025; Miao & Cukurova, 2024; Marras & Pasqualotto, 2026). The checklist should therefore be understood as a developing reflective scaffold, whose main contribution lies in making teacher decision-making around AI more explicit and discussable, rather than as a prescriptive protocol.

These findings should nevertheless be interpreted in light of the exploratory nature of the study, which involved a small, self-selected sample in a single professional development context and assessed perceived clarity and utility immediately after guided use. Moreover, each participant worked actively with two to three indicators but rated all 20, so that most ratings reflect reading and demonstration rather than use. Phase comparisons are confounded with the workshop content, since the guided demonstration used a Phase 2 indicator, and with the fixed order in which phases were presented in the questionnaire. Because the workshop was led by the developers of the checklist, social desirability may have inflated ratings; participants who worked in pairs may have provided non-independent judgments; and school level was not recorded. The results therefore cannot be considered a formal psychometric validation of the instrument (Lynn, 1986; Polit & Beck, 2006). Consistent with the iterative logic of educational design research (McKenney & Reeves, 2019), further iterations should test a shorter, modular version across school levels and under more naturalistic conditions of use. Future research should also examine the sustainability of the instrument, the feasibility of indicators requiring formative or diagnostic information, and the potential role of critical-ethical considerations as a transversal dimension.

References

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

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, Article 102274. https://doi.org/10.1016/j.lindif.2023.102274

Lynn, M. R. (1986). Determination and quantification of content validity. Nursing Research, 35(6), 382–385. https://doi.org/10.1097/00006199-198611000-00017

Marras, A., Fornara, F., & Pasqualotto, A. (2025). Harnessing AI for inclusive teaching: Professional development strategies for early educators" In S. Papadakis (Ed.) *AI in early education: Integrating artificial intelligence for inclusive and effective Learning.* (pp. 267-290). Wiley. https://doi.org/10.1002/9781394352821.ch16

Marras, A., & Pasqualotto, A. (2026). Negotiating assistive technologies and AI in inclusive education: Professional agency in neurodivergent contexts. *Frontiers in Child and Adolescent Psychiatry*, 5, Article 1820276. https://doi.org/10.3389/frcha.2026.1820276

McKenney, S., & Reeves, T. C. (2019). Conducting educational design research (2nd ed.). Routledge.

Miao, F., & Cukurova, M. (2024). AI competency framework for teachers. UNESCO. https://doi.org/10.54675/ZJTE2084

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, 108(6), 1017–1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x

Mishra, P., Warr, M., & Islam, R. (2023). TPACK in the age of ChatGPT and generative AI. *Journal of Digital Learning in Teacher Education*, 39(4), 235–251.https://doi.org/10.1080/21532974.2023.2247480

Pizzuti, E. (2026). L'intelligenza artificiale come partner didattico per la progettazione di attività didattiche inclusive: Verso una checklist operativa per docenti [Master's thesis, University of Applied Sciences and Arts of Southern Switzerland].

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, 29(5), 489–497. https://doi.org/10.1002/nur.20147

Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems*, 3, 121–154. https://doi.org/10.1016/j.iotcps.2023.04.003

Tomlinson, C. A. (2014). Differentiated instruction. In J. A. Plucker & C. M. Callahan (Eds.), Critical issues and practices in gifted education: What the research says (2nd ed., pp. 197–210). Prufrock Press.

Tucker, C. R., & Novak, K. (2025). Elevating educational design with AI: Making learning accessible, inclusive, and equitable. IMPress.

Wiggins, G., & McTighe, J. (2005). Understanding by design (Expanded 2nd ed.). ASCD.

Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., Li, X., Jin, Y., & Gašević, D. (2024). Practical and ethical challenges of large language models in education: A systematic scoping review. *British Journal of Educational Technology*, 55(1), 90–112. https://doi.org/10.1111/bjet.13370