
Ethical Boundaries for the Use of Artificial Intelligence in Academic Assessment

Emerging insights into educators' perspectives

Abstract

Purpose: This study examines educators' perceptions of the ethical boundaries for using Artificial Intelligence (AI) in teaching and academic assessment, distinguishing AI support from AI delegation using flipped learning as a pedagogical lens.

Design/methodology/approach: A questionnaire was administered to 148 faculty members from Universidade Autónoma de Lisboa, Portugal. The paper adopts an exploratory quantitative case-study design. Analysis used descriptive statistics, normalised Shannon entropy, Pearson correlation, and a paired-samples t-test.

Findings: Respondents were open to AI support for routine, preparatory, and lower-order learning activities, but more cautious when AI was associated with evaluation, fairness, interpretation, and assessment decisions. Teacher judgment was strongly prioritised, while trust in AI-generated fair and unbiased evaluation remained limited. Entropy analysis showed relatively clear boundaries around AI bias and human judgment, but more contested boundaries around AI-supported evaluation, learner independence, and critical reflection.

Research limitations/implications: The single-institution convenience sample should be interpreted as exploratory and analytically transferable rather than statistically generalisable.

Practical implications: Institutions should clarify where AI may support assessment and where human evaluative authority must remain explicit.

Social implications: The findings highlight the need for shared institutional norms.

Originality/value: The paper proposes the Dual Approach as an interpretive framework connecting ethical responsibility, cognitive complexity, and flipped learning to distinguish AI support from delegated evaluative authority. Its contribution lies in locating the ethical boundary at the transfer of assessment responsibility and identifying where educators' perceptions indicate crisp or fuzzy boundaries.

Keywords: Academic Ethics; AI in Education; Assessment Responsibility; Trust in AI; Bloom's Taxonomy; Flipped Learning.

1 Introduction

Artificial Intelligence (AI) has become increasingly present in higher education, including teaching, learning support, feedback, assessment design, and academic decision-making. Its integration into education offers opportunities for personalisation, accessibility, efficiency, and new forms of pedagogical support (Holmes et al., 2023; Luckin et al., 2016; UNESCO, 2021). However, when AI is used in academic assessment, the issue is not only technical performance or pedagogical efficiency. A central ethical question emerges: *when AI participates in assessment, who ultimately bears moral responsibility for evaluative judgment?*

Academic assessment is not merely a technical procedure. It affects students' academic progression, recognition, opportunities, and future trajectories. For this reason, AI-supported assessment cannot be treated as ethically neutral. The challenge is not simply to decide whether AI should be accepted or rejected, but to identify where AI is perceived as legitimate support and where it becomes problematic delegation of evaluative responsibility.

Current debates on AI in education address trust, privacy, fairness, bias, academic integrity, and the changing role of teachers (Holmes et al., 2023; UNESCO, 2021; Zawacki-Richter et al., 2019). These concerns are especially visible in assessment contexts, where AI may assist with feedback, grading support, question generation, or the design of computer-based tests (Matos and Dediu, 2026). Yet assessment credibility cannot be addressed only through detection, proctoring, or technical safeguards. It also requires careful assessment design and a clear allocation of responsibility between AI tools and human educators (Evangelista, 2025; UNESCO, 2021).

This paper approaches the problem through the idea of ethical boundaries. Some boundaries may be relatively crisp, where educators broadly agree that AI use is acceptable or unacceptable. Others may remain fuzzy, where views are dispersed, contested, or institutionally unsettled. These grey zones are important because they indicate areas where professional norms and institutional guidance may still be under development.

Bloom's Taxonomy is relevant to this problem because it distinguishes levels of cognitive complexity, from remembering and understanding to analysis, evaluation, and creation (Anderson and Krathwohl, 2001). In this study, Bloom's Taxonomy is used as a pedagogical lens, not as an ethical theory. It helps examine whether educators perceive different boundaries for AI use in lower-order and higher-order cognitive tasks. Flipped learning is understood here as a pedagogical model in which students' initial engagement with content occurs before class, while classroom time is reserved for interaction, discussion, application, and evaluation. In this setting, AI-supported preparation can be distinguished from human-led evaluative activity.

Existing research has examined AI ethics in education, human oversight in assessment, cognitive complexity, and flipped learning from several complementary perspectives. However, these strands are often discussed separately. Less attention has been given to how cognitive complexity, pedagogical context, and evaluative responsibility can be combined to distinguish AI assistance from the delegation of consequential academic judgment. This study addresses that gap through the Dual Approach.

This paper therefore examines AI in academic assessment as a problem of ethical boundary-setting: *when is AI perceived as legitimate support, and when does it become problematic delegation of evaluative responsibility?* Two research questions (RQ) guide the study:

RQ1. How do educators distinguish between acceptable AI support and problematic delegation of evaluative responsibility in teaching and academic assessment?

RQ2. Which ethical boundaries appear relatively clear, and which remain fuzzy or contested, when educators evaluate AI use across learning, assessment, and higher-order cognitive activities?

The paper contributes by interpreting educators' survey responses through the Dual Approach, which connects three dimensions: the ethical allocation of responsibility, cognitive complexity as interpreted through Bloom's Taxonomy, and the pedagogical organisation of learning through flipped learning. The distinctive contribution lies in using these dimensions together to identify where AI remains supportive and where it risks substituting for human evaluative authority.

The article is organised as follows. Section 2 reviews the literature, Section 3 presents the conceptual framework, Section 4 describes the methodology, Section 5 reports the results, Section 6 discusses the findings, and Section 7 presents the conclusions, limitations, and future work.

2 Literature Review

AI in higher education assessment. AI in education has been associated with personalised learning, adaptive support, automated feedback, question generation, and increased efficiency in teaching and assessment processes (Holmes et al., 2023; Luckin et al., 2016; UNESCO, 2021). In assessment, AI may support the design of computer-based tests, the evaluation of question clarity, the alignment between questions and the topics

taught, and the enrichment of question databases with diverse question types (Matos and Dediu, 2026). These uses suggest that AI can contribute to the preparation and improvement of assessment instruments.

However, AI support in assessment must be distinguished from the delegation of assessment authority. Support refers to assistance, feedback, structuring, or preliminary analysis. Delegation occurs when AI-generated outputs serve as the basis for evaluative judgment without sufficient human interpretation, validation, or accountability. This distinction is central to the ethical problem addressed in this paper.

Ethical judgment, trust, and responsibility. The ethical dimension of AI-supported assessment is multidimensional. Ethical evaluation cannot be reduced to a single judgment of approval or rejection, but may involve dimensions such as fairness, responsibility, acceptability, harm, and institutional expectations (Reidenbach and Robin, 1990). Ethical sensitivity and ethical evaluation should also be distinguished: educators may recognise that a situation has ethical relevance while still differing in how they judge the acceptability of a given practice (Shawver and Sennetti, 2009).

This distinction is especially important in AI-supported assessment, where awareness of possible risks may coexist with willingness to use AI tools. Technology acceptance research similarly indicates that ethical judgment can shape acceptance of intelligent technologies, particularly when such technologies intervene in sensitive human capacities or decision-making contexts (Olar-te-Pascual et al., 2021). In assessment, trust is therefore not only a technical matter. It is linked to responsibility, legitimacy, and confidence that AI-generated outputs will not undermine fairness or human accountability.

Bias, fairness, integrity, and accountability. AI-supported assessment raises concerns about fairness, privacy, bias, transparency, and academic integrity (Holmes et al., 2023; UNESCO, 2021). AI may help reduce certain forms of human bias, for example by promoting consistency or providing additional information. At the same time, AI systems may reproduce or amplify biases embedded in training or learning data, model design, or algorithmic decision-making processes. For this reason, human oversight remains necessary, especially in contexts where evaluation affects student recognition and progression.

Academic integrity also requires more than detection or proctoring. AI-aware assessment design should reduce opportunities for misuse by making tasks more authentic, reflective, and connected to higher-order learning (Evangelista, 2025). This shifts the problem from simply preventing misconduct to designing assessments where AI use can be guided, interpreted, and bounded.

Cognitive complexity and AI-supported assessment. Bloom's Taxonomy provides a widely used framework for distinguishing levels of cognitive complexity: remembering, understanding, applying, analysing, evaluating, and creating (Anderson and Krathwohl, 2001). In AI-supported education, this framework has been used to examine how AI tools may support different types of learning activity (Akdeniz et al., 2025; Faraon et al., 2023; Hachoumi et al., 2025; Zhao and Fu, 2025).

A common assumption is that AI is particularly effective at supporting Lower-Order Thinking Skills (LOTS), such as remembering and understanding, because many of these tasks can be partially automated. The role of AI in Higher-Order Thinking Skills (HOTS), such as analysis, evaluation, creativity, and critical thinking, is more complex. Some authors caution that AI should not be relied upon independently for HOTS-related tasks (Hachoumi et al., 2025; Li, 2024). However, recent work suggests a more nuanced position: generative AI may support aspects of advanced reasoning by helping users compare alternatives, generate explanations, or identify possible lines of argument. At the same time, it may shift critical thinking toward verification, response integration, and task stewardship (Lee et al., 2025). Overreliance on AI may also reduce opportunities for independent problem-solving and critical engagement when users accept AI outputs without verification (Naqvi et al., 2025).

For this reason, HOTS should not be described as entirely inaccessible to AI. Rather, they should be understood as human-led processes in which AI may provide support, while interpretation, evaluation, ethical judgment, and responsibility remain dependent on learners and teachers.

Flipped learning and active learning as assessment contexts. Flipped learning provides a useful context for distinguishing AI-supported preparation from human-led evaluative activity. In flipped learning, initial engagement with content takes place before class, while classroom time is used for interaction, discussion, application, and higher-order learning. Recent work links flipped learning with AI-supported preparation, generative AI, Bloom's Taxonomy, and student–teacher–GenAI interactions (Kwan et al., 2025; López-Villanueva et al., 2024; Tang et al., 2025). Lecturer-perception research also suggests cautious optimism: AI may support preparation, resource development, and engagement, but raises concerns about over-reliance, superficial learning, and the need for pedagogical guidance (West and Van der Westhuizen, 2026). In this context, AI may support explanations, summaries, quizzes, adaptive resources, or feedback, while classroom time can remain focused on interpretation, discussion, assessment, and reflective judgment.

This structure can also support active learning methodologies such as Problem-Based Learning (PBL), Challenge-Based Learning (CBL), Simulation-Based Learning (SBL), Research-Based Learning (RBL), and Project-Based Learning (PjBL), where classroom activity shifts toward analysis, evaluation, creation, collaboration, and reflective judgment (Hmoud and Ali, 2024; Murphy and O'Neill, 2022).

3 Conceptual Framework

The ethical framework adopted in this study draws on widely recognised AI ethics principles. Across AI ethics guidelines, five recurring domains are particularly relevant to AI-supported assessment: transparency, justice and fairness, non-maleficence, responsibility, and privacy (Jobin et al., 2019). These domains help clarify why assessment involving AI cannot be treated only as a technical or pedagogical matter. Opaque outputs, biased data or models, privacy risks, and unclear responsibility may affect students' academic recognition and progression. In this paper, these principles are operationalised through the distinction between AI support and AI delegation: AI may assist teaching and assessment processes, but final evaluative authority remains ethically sensitive when fairness, accountability, and student outcomes are at stake.

The conceptual framework of this study is the Dual Approach, which distinguishes between AI support and AI delegation in teaching and assessment. AI support refers to the use of AI to assist with preparation, explanation, structuring, or providing preliminary feedback, while leaving interpretation and final responsibility with human educators. AI delegation refers to situations in which AI-generated outputs replace or determine evaluative judgment, especially in contexts where fairness, interpretation, accountability, and academic recognition are at stake.

The Dual Approach therefore extends rather than replaces the perspectives reviewed above. AI ethics frameworks identify the values at stake, including fairness, transparency, privacy, responsibility, and accountability. Bloom's Taxonomy provides a pedagogical lens on cognitive complexity, while flipped learning distinguishes preparatory activity from classroom interaction and higher-order engagement. The Dual Approach adds an explicit responsibility-allocation perspective: it asks not only what cognitive activity AI supports or where that activity occurs, but whether human educators retain interpretation, accountability, and final evaluative authority.

This distinction is central because the ethical issue is not simply AI use itself, but the point at which AI support shifts the locus of evaluative responsibility. AI may be acceptable in routine, preparatory, or lower-order activities, but becomes more sensitive when it is used to interpret complex student work or inform final assessment decisions.

Bloom's Taxonomy is used in this framework as a pedagogical lens for cognitive complexity. Lower-order activities, such as remembering, understanding, and routine application, may be more compatible with AI-supported preparation, practice, or feedback. Higher-order activities, such as analysis, evaluation, and creation, involve interpretation, contextual judgment, and responsibility-sensitive decisions. Table I summarises this relationship by linking Bloom's cognitive levels to examples of AI support and the corresponding need for human oversight.

Table I. Alignment of AI Capabilities with Bloom's Taxonomy and Required Human Oversight

Cognitive Level	Examples of AI Support	Ethical Need for Human Oversight
Remember	Flashcards, definitions	Low to Medium
Understand	Summarizing texts, concept maps	Medium
Apply	Solving typical exercises, grammar correction	Medium
Analyze	Categorization and comparison tools	High
Evaluate	Essay scoring support, peer-review simulation	High
Create	Brainstorming, content generation	High

Note: Cognitive levels follow Bloom's revised taxonomy.

Source: Oversight levels are the authors' interpretive synthesis, informed by the AI ethics and assessment literature reviewed in this article.

Table I should not be read as a deterministic classification. Rather, it functions as a heuristic for identifying where AI support becomes more ethically sensitive. As cognitive complexity increases, the legitimacy of delegating evaluative judgment to AI decreases. Bloom's Taxonomy, therefore, helps identify where ethical boundaries may become more contested, but it is not treated as an ethical theory in itself.

Flipped learning provides the contextual frame for this distinction. AI-supported preparation may be appropriate before class, where learners engage with explanations, summaries, quizzes, examples, or feedback. Classroom time can then be reserved for discussion, problem-solving, critical reflection, and assessment of higher-order learning. In this structure, AI supports learning preparation, while teachers remain responsible for guiding interpretation and making evaluative judgments.

The Dual Approach combines a pedagogical and an ethical dimension. Pedagogically, it aligns AI use with the type of learning activity and the level of cognitive complexity. Ethically, it asks whether the use of AI preserves or weakens human responsibility, fairness, accountability, and trust. The framework does not assume that all AI use in assessment is problematic. Instead, it proposes that the acceptability of AI depends on the task, the degree of cognitive complexity, and the extent to which evaluative authority remains human.

This framework also allows the study to distinguish between crisp and fuzzy ethical boundaries. Crisp boundaries are indicated by strong convergence among respondents, suggesting shared expectations about acceptable or unacceptable AI use. Fuzzy boundaries are indicated by dispersed responses, neutrality, or disagreement, suggesting that professional and institutional norms remain unsettled. The empirical analysis therefore examines both what educators think about AI-supported assessment and how consensual or contested those perceptions are.

4 Methodology

A survey of 148 faculty members at the Universidade Autónoma de Lisboa in Lisbon, Portugal, further explores perceptions of Artificial Intelligence in Education (AIED). The questionnaire was intentionally designed to foreground ethical decision-making in assessment contexts. Items were grouped to capture perceptions of trust, responsibility, fairness, and integrity, and to examine how these ethical concerns vary across cognitive levels defined by Bloom's Taxonomy.

4.1 Study Design and Sample

This study employed a quantitative, cross-sectional, non-experimental survey design. A convenience sampling approach was used, based on respondents' voluntary participation. Data collection was concentrated at the Universidade Autónoma de Lisboa (UAL), located in Lisbon, Portugal. The final sample comprised N =

148 complete responses. The respondents are primarily university faculty members, most holding a PhD, typically aged between 30 and 65 years, and residing in the Lisbon region. Although data collection was concentrated at UAL, the sample included broader academic and institutional experience: around 45% of participants also collaborated with other higher education institutions or research centres, involving at least 15 different institutions. In addition, around 8% of participants held foreign nationality, mainly from Ibero-American and European contexts, representing at least six countries. This indicates some diversity in age, academic trajectories, institutional experience, and international academic and cultural backgrounds.

Given the exploratory nature of this study and its focus on ethical perceptions rather than behavioural measurement, the convenience sample is appropriate for identifying emerging normative patterns. However, this sampling strategy also has limitations. Since participation was voluntary, the sample may include self-selection bias: faculty members with stronger interest in AI, assessment, pedagogical innovation, or ethical issues may have been more likely to respond. The sample should therefore not be interpreted as statistically representative of higher education faculty in general, nor as evidence of a national pattern. Rather, the findings should be read as analytically transferable insights from a contextualised case study, since the issues examined are relevant across higher education, while the specific balance of perceptions may be shaped by national traditions, institutional culture, disciplinary background, and local policy environments.

4.2 *Data Collection Procedure and Ethical Considerations*

The data collection period spanned one month, from September 15, 2025, to October 15, 2025. The questionnaire was administered through dual modes: online (via institutional email) and in-person sessions (during the internal seminar titled "Artificial Intelligence in Teaching: AI, critical thinking and learning active methodologies: a perspective using the Bloom Taxonomy" – Jornadas Pedagógicas do Ensino Superior - JOPES).

In both the remote and in-person settings, participants accessed the survey online using a Google Forms link to ensure data standardisation, anonymity, and confidentiality. Therefore, all responses, including those collected during JOPES sessions, were registered anonymously and digitally.

Participation was strictly voluntary. The study adhered to the ethical principles applicable to human-subject research in opinion surveys, ensuring the anonymity and confidentiality of all responses throughout the process.

At the time of data collection, neither the authors' department nor their research centre had an ethics committee with jurisdiction over this type of anonymous educational survey. Participation was voluntary, and responses were collected anonymously through an online questionnaire. No direct identifiers or sensitive personal information were collected.

An exact response rate could not be reliably calculated because the questionnaire was distributed via both institutional email and in-person seminars, and because responses were collected anonymously.

4.3 *Instrument and Measurement*

The data collection instrument was a structured questionnaire comprising 26 closed-ended items. The questionnaire was designed to examine educators' perceptions of the ethical boundaries for AI use in teaching and assessment, with particular attention to the distinction between AI as support for learning and AI as support for evaluation. The items were organised into five thematic dimensions: awareness and understanding, perceived usefulness, ethics and integrity, intention and adoption, and pedagogical alignment with Bloom's Taxonomy and flipped learning.

The five dimensions were linked to the Dual Approach as follows. Awareness and understanding provided the contextual baseline for respondents' familiarity with AI and awareness of bias. Perceived usefulness captured AI as pedagogical support. Ethics and integrity operationalised the responsibility dimension through privacy, academic integrity, trust in AI-generated evaluation, and the precedence of human judgment. Intention and adoption captured willingness to use AI within these ethical constraints. Pedagogical alignment examined the cognitive and contextual dimension through Bloom's LOTS/HOTS distinction and flipped learning. Together, these dimensions allowed the instrument to examine not only whether educators

accept AI, but under which cognitive, pedagogical, and responsibility conditions AI is perceived as support rather than delegation.

Most items were measured using a 5-point Likert scale, ranging from 1 = Strongly Disagree to 5 = Strongly Agree. Some awareness and pedagogical items used categorical response formats where appropriate. Examples of items include: “I am aware that AI tools may contain inherent biases that can affect educational outcomes”; “I trust AI tools to provide fair and unbiased evaluation feedback”; “I believe human teacher judgment should always take precedence over AI-generated assessments”; “I believe AI tools are more effective in supporting lower-order cognitive skills”; and “I am cautious about using AI tools to assess higher-order cognitive tasks.”

The questionnaire was developed as an exploratory instrument for this case study and was not intended as a previously validated psychometric scale. Its purpose was to capture patterns of perception across the ethical and pedagogical dimensions relevant to the Dual Approach. To assess internal consistency across the closed-ended items, Cronbach’s Alpha was calculated, yielding 0.88, indicating strong reliability for exploratory analysis.

4.4 Data Analysis

Data processing and analysis were performed using SPSS v28 (Statistical Package for the Social Sciences). The analysis included:

- Descriptive Statistics: Calculation of the Mean ($\bar{X}$), Standard Deviation (SD), and frequency distributions for all items.
- Inferential Statistics: To empirically examine key relationships and contrasts implied by the Dual Approach framework, inferential tests were applied.
 - Pearson’s Correlation Coefficient (r): Used to examine the relationship between the Intention to Use (D15) and the Trust in Unbiased Evaluation (C14), which helps quantify the observed adoption paradox.
 - Paired Sample t -Test: Conducted to compare the means of perceived AI effectiveness in supporting LOTS (E23) versus caution in assessing HOTS (E24), providing empirical support for the proposed cognitive division of labour.

The level of statistical significance was set at $p < 0.05$.

5 Survey Questions and Results

This section presents the results of a survey conducted with a total sample of $N=148$ faculty members at Universidade Autónoma de Lisboa in Lisbon, Portugal. The findings are organised into five core thematic areas, providing a detailed breakdown of the educational community’s perceptions of AIEd across awareness, ethics, pedagogical alignment, usefulness, and intention.

To complement the mean values, normalised Shannon entropy was computed for each response distribution. Entropy measures the degree of dispersion or uncertainty in a probability distribution (Shannon, 1948). In this study, the response percentages for each item were treated as a distribution across the available response categories. The entropy value was normalised by the maximum possible entropy for the number of categories in each item, allowing comparison across 5-point Likert items, binary Yes/No items, and the four-category Bloom’s Taxonomy item.

$$H_N = \frac{-\sum_{i=1}^k p_i \ln(p_i)}{\ln(k)}$$

Lower normalised entropy indicates that responses are concentrated in fewer categories, suggesting stronger consensus, whereas higher entropy indicates more dispersed responses, suggesting ambiguity, disagreement, or a less clearly defined ethical boundary. Thus, while the mean indicates the general direction of responses, entropy indicates how consensual or contested that direction is.

5.1 Awareness and Understanding

This subsection assesses the respondents' familiarity with AI tools in education and their comprehension of basic functions and implications, using a Likert scale.

Table II. Distribution and Descriptive Statistics for AI Awareness and Understanding (Likert Items, N=148)

Item	Survey Question	Responses [%]					$\bar{X}$	SD	H_N
		1	2	3	4	5			
A1	I am aware that AI tools are being used in educational settings.	0.0%	1.4%	6.8%	35.1%	56.8%	4.47	0.69	0.579
A2	I understand the basic functions and capabilities of AI tools in education.	1.4%	0.0%	8.1%	54.1%	36.5%	4.24	0.72	0.599
A5	I am aware that flipped learning methodologies can be enhanced by AI tools.	0.0%	4.1%	27.0%	36.5%	32.4%	3.97	0.88	0.756

Note: Response categories: 1-Strongly Disagree; 2-Disagree; 3-Neutral; 4-Agree; 5-Strongly Agree. Percentages may not total 100 due to rounding.

Source: Authors' own survey data.

The results in Table II indicate a generally high level of AI awareness and understanding among participants. Item A1 recorded the highest mean ($\bar{X} = 4.47$), with 91.9% of respondents agreeing or strongly agreeing, demonstrating that the use of AI in education is a widely recognised reality. Understanding of basic functions (A2, $\bar{X} = 4.24$) and awareness of AI's potential to enhance flipped learning (A5, $\bar{X} = 3.97$) are also notably high, suggesting a knowledgeable respondent base.

Table III. Yes/No Responses for Awareness of Bias and Mitigation (N=148)

Item	Survey Question	Yes	No	H_N
A3	I am aware of the potential for teacher bias in assessment and consider how AI tools might help mitigate it.	68.9%	31.1%	0.894
A4	I am aware that AI tools may contain inherent biases that can affect educational outcomes.	91.9%	8.1%	0.406

Source: Authors' own survey data.

Table III shows that 91.9% of participants answered "Yes" to item A4, indicating awareness that AI tools may contain biases affecting educational outcomes. In item A3, 68.9% of participants reported awareness of potential teacher bias in assessment and considered that AI tools might help mitigate it.

5.2 Ethics and Integrity

This subsection explores ethical concerns, trust, and the importance of human judgment, using a Likert scale.

Table IV. Distribution and Descriptive Statistics for Ethics and Integrity (N=148)

Item	Survey Question	Responses [%]					$\bar{X}$	SD	H_N
		1	2	3	4	5			
C11	I am concerned about data privacy when using AI tools.	0.0%	4.1%	16.2%	50.0%	29.7%	4.05	0.79	0.704
C12	I believe the use of AI tools for academic assignments raises concerns about academic integrity.	1.4%	8.1%	1.4%	51.4%	37.8%	4.16	0.91	0.642
C13	I believe human teacher judgment should always take precedence over AI-generated assessments.	1.4%	0.0%	6.8%	27.0%	64.9%	4.54	0.74	0.545
C14	I trust AI tools to provide fair and unbiased evaluation feedback.	9.5%	18.9%	48.6%	21.6%	1.4%	2.86	0.91	0.795

Note: Response categories: 1-Strongly Disagree; 2-Disagree; 3-Neutral; 4-Agree; 5-Strongly Agree.

Source: Authors' own survey data.

Ethical concerns (Table IV) are highly pronounced. Item C13 received the highest mean in this subsection ($\bar{X} = 4.54$), with 91.9% of respondents agreeing or strongly agreeing that human teacher judgment should take precedence over AI-generated assessments. Concerns over academic integrity (C12, $\bar{X} = 4.16$) and data privacy (C11, $\bar{X} = 4.05$) also received high levels of agreement. Item C14 recorded the lowest mean in the survey ($\bar{X} = 2.86$), with 48.6% of respondents selecting Neutral and 28.4% selecting Disagree or Strongly Disagree.

5.3 Pedagogical Alignment with Bloom's Taxonomy

This subsection assesses the perceived alignment and use of AI tools within pedagogical frameworks, using a Likert scale.

Table V. Distribution and Descriptive Statistics for Pedagogical Alignment (Likert Items) (N=148)

Item	Survey Question	Responses [%]					$\bar{X}$	SD	H_N
		1	2	3	4	5			
E19	I actively encourage students to use AI tools as part of their learning process.	2.7%	14.9%	20.3%	39.2%	23.0%	3.65	1.08	0.876
E20	I perceive AI tools as a potential challenge to the traditional role of teachers in my teaching context.	0.0%	6.8%	8.1%	45.9%	39.2%	4.18	0.85	0.690
E21	I use AI tools to support and complement my own student evaluations.	12.2%	20.3%	24.3%	20.3%	23.0%	3.22	1.34	0.985
E23	I believe AI tools are more effective in supporting lower-order cognitive skills.	4.1%	23.0%	36.5%	23.0%	13.5%	3.19	1.07	0.898
E24	I am cautious about using AI tools to assess higher-order cognitive tasks.	0.0%	4.1%	14.9%	52.7%	28.4%	4.05	0.77	0.689
E25	I consider Bloom's Taxonomy when deciding whether to use AI tools in teaching or assessment.	4.1%	6.8%	50.0%	23.0%	16.2%	3.41	0.98	0.804
E26	I believe flipped learning provides the right conditions for methodologies such as PBL, CBL, SBL, RBL, and PjBL.	0.0%	1.4%	52.7%	29.7%	16.2%	3.61	0.77	0.654

Note: Response categories: 1-Strongly Disagree; 2-Disagree; 3-Neutral; 4-Agree; 5-Strongly Agree.

Source: Authors' own survey data.

Table V shows strong agreement that AI poses a challenge to the traditional teacher role (E20, $\bar{X} = 4.18$) and caution regarding the use of AI for assessing HOTS (E24, $\bar{X} = 4.05$). Agreement was lower for AI effectiveness in supporting LOTS (E23, $\bar{X} = 3.19$), encouragement of student AI use (E19, $\bar{X} = 3.65$), and flipped learning as a context for AI-supported methodologies (E26, $\bar{X} = 3.61$).

A paired-samples t-test comparing E23 and E24 showed a statistically significant difference ($t(147) = 9.87$, $p < 0.001$), with greater caution in assessing HOTS than in assessing AI effectiveness in supporting LOTS. The interpretation of this difference in relation to the Dual Approach is developed in the Discussion.

Table VI. Categorical Responses for Bloom's Taxonomy Use (N=148)

Item	Survey Question	Responses [%]				$\bar{X}$	SD	H_N
		1	2	3	4			
E22	I structure my teaching and assessment practices according to Bloom's Taxonomy (e.g., remember, understand, apply, analyze, evaluate, create).	18.9%	12.2%	28.4%	40.5%	2.91	1.13	0.934

Note: Response categories: 1-No, not familiar; 2-No, but familiar with it; 3-Yes, occasionally; 4-Yes, systematically.

Source: Authors' own survey data.

Table VI indicates that Bloom's Taxonomy is a prevalent framework among participants. A combined 68.9% of respondents structure their practices either systematically (40.5%) or occasionally (28.4%) according to the Taxonomy. Only 18.9% reported being unfamiliar with it, reinforcing the idea that pedagogical alignment is being considered within a known, established framework.

5.4 Perceived Usefulness

This subsection measures perceptions of AI's utility in teaching and learning using a Likert scale.

Table VII. Distribution and Descriptive Statistics for Perceived Usefulness (N=148)

Item	Survey Question	Responses [%]					$\bar{X}$	SD	H_N
		1	2	3	4	5			
B6	AI tools help me teach or learn more effectively.	2.7%	5.4%	18.9%	37.8%	35.1%	3.97	1.01	0.811
B7	I believe AI tools contribute positively to academic performance and learning outcomes.	2.7%	10.8%	18.9%	39.2%	28.4%	3.80	1.06	0.856
B8	I feel more independent in my learning when using AI tools.	4.1%	16.2%	36.5%	25.7%	17.6%	3.36	1.08	0.900
B9	AI tools help me reflect more critically on the subjects I study.	2.7%	17.6%	20.3%	37.8%	21.6%	3.58	1.10	0.886
B10	I believe that combining AI with flipped learning fosters deeper critical thinking.	4.1%	5.4%	29.7%	37.8%	23.0%	3.70	1.02	0.842

Note: Response categories: 1-Strongly Disagree; 2-Disagree; 3-Neutral; 4-Agree; 5-Strongly Agree.

Source: Authors' own survey data.

Overall perceived usefulness (Table VII) is strong, particularly concerning the general effectiveness (B6, $\bar{X} = 3.97$) and positive contribution to outcomes (B7, $\bar{X} = 3.80$). However, respondents show lower agreement regarding the AI's ability to foster independence (B8, $\bar{X} = 3.36$) and support critical reflection (B9, $\bar{X} = 3.58$). This suggests that while AI is valued as an efficient tool, its perceived role in complex pedagogical processes remains moderate, with the neutral category garnering a large share of responses (36.5% for B8, 20.3% for B9).

5.5 Intention and Adoption

This subsection focuses on the intention to use and recommend AI tools in the future, measured on a Likert scale.

Table VIII. Distribution and Descriptive Statistics for Intention and Adoption (N=148)

Item	Survey Question	Responses [%]					$\bar{X}$	SD	H_N
		1	2	3	4	5			
D15	I intend to continue using AI tools in my teaching or learning activities.	1.4%	5.4%	10.8%	41.9%	40.5%	4.15	0.92	0.738
D16	I would recommend the use of AI tools to colleagues or peers.	2.7%	4.1%	14.9%	39.2%	39.2%	4.08	0.98	0.774
D17	I believe AI will significantly transform the teacher's role within the next five years.	0.0%	8.1%	8.1%	28.4%	55.4%	4.31	0.94	0.678
D18	I think AI tools will become essential components of future educational environments.	1.4%	4.1%	8.1%	33.8%	52.7%	4.32	0.89	0.683

Note: Response categories: 1-Strongly Disagree; 2-Disagree; 3-Neutral; 4-Agree; 5-Strongly Agree.

Source: Authors' own survey data.

Table VIII shows high levels of intention and adoption. Respondents reported a strong intention to continue using AI (D15, $\bar{X} = 4.15$) and to recommend it (D16, $\bar{X} = 4.08$). They also agreed that AI will transform the teacher's role (D17, $\bar{X} = 4.31$) and become an important component of future educational environments (D18,

$\bar{X} = 4.32$). Over 80% of participants agreed or strongly agreed with the predictive statements (D17 and D18), indicating that respondents expect AI to become increasingly integrated into education.

To examine the relationship between adoption intention and trust in AI-generated evaluation, a Pearson correlation was calculated between D15 and C14. The analysis showed a moderate positive association ($r = 0.421$, $p < 0.001$), indicating that respondents reporting greater trust in AI-generated fair and unbiased evaluation also tended to report stronger intentions to continue using AI tools.

5.6 Entropy-Based Reading of Ethical Boundaries

To complement the mean-based analysis, we computed normalised Shannon entropy for the response distributions. Entropy distinguishes the direction of opinion from the degree of consensus: the mean indicates whether respondents tend to agree or disagree, while entropy indicates whether responses are concentrated or dispersed across categories. Lower entropy therefore suggests clearer consensus, whereas higher entropy suggests ambiguity, disagreement, or a less settled boundary.

Table IX. Items with the Lowest and the Highest Normalised Entropy

Entropy	Item	Theme	$\bar{X}$	H_N	Response Pattern
Lower	A4	Awareness and Understanding	91.9% Yes	0.406	Strong consensus on AI bias
Lower	C13	Ethics and Integrity	4.54	0.545	Strong consensus on human judgment
Lower	A1	Awareness and Understanding	4.47	0.579	Strong consensus on AI presence
Lower	A2	Awareness and Understanding	4.24	0.599	Strong consensus on basic understanding of AI tools
Higher	E21	Pedagogical Alignment	3.22	0.985	Highly dispersed views on AI supporting evaluation
Higher	E22	Pedagogical Alignment	2.91	0.934	Dispersed use of Bloom's Taxonomy
Higher	B8	Perceived Usefulness	3.36	0.900	Dispersed views on AI and learner independence
Higher	E23	Pedagogical Alignment	3.19	0.898	Dispersed views on AI and LOTS support
Higher	B9	Perceived Usefulness	3.58	0.886	Dispersed views on AI tools supporting critical reflection

Source: Authors' own calculations based on survey data.

Table IX presents the items with the lowest and highest normalised entropy values. Lower entropy values were observed for A4, C13, A1, and A2. Higher entropy values were observed for E21, E22, B8, E23, and B9. These patterns are interpreted in the Discussion.

6 Discussion

This section interprets educators' perceptions of AIEd. The empirical results of this study reveal a dual perspective on AIEd: a strong drive for technological adoption coexisting with profound ethical scepticism. This caution toward AI in assessment does not appear to reflect a general rejection of the technology, but rather an underlying concern about preserving moral accountability in evaluative practices. The high intention to adopt and the recognition of AI's transformative role (D17, D18) align with the literature on the future of education (Holmes et al., 2023). However, this optimism is tempered by limited trust in AI-generated evaluation: the overwhelming support for the primacy of human judgment (C13) and low trust in AI evaluation (C14) underscore that, regardless of technological capacity, the teacher should remain the ultimate authority in

assessment (Roll and Wylie, 2016). This relationship is further quantified by the statistically significant positive correlation between Intention to Use and Trust in Evaluation ($r = 0.421, p < 0.001$), indicating that greater trust in AI-generated evaluation is associated with stronger intention to use AI tools. This distrust is further reinforced by concerns about academic integrity (C12) and AI bias (A4), which validate the need for ethical and governance safeguards (Kwan et al., 2025).

This coexistence of adoption intention and ethical reservation is also consistent with technology acceptance research, which shows that ethical judgment can shape acceptance of intelligent technologies, particularly when such technologies intervene in sensitive human capacities or decision-making contexts (Olarte-Pascual et al., 2021).

The entropy results clarify where respondents' perceptions are more settled and where they remain contested. Low entropy values for A1 and A4 indicate broad recognition of AI's presence in education and of the risks associated with AI bias, while C13 shows strong convergence around the precedence of human teacher judgment. These results suggest relatively crisp ethical boundaries around bias awareness and final human responsibility. By contrast, higher entropy values for E21, E22, B8, E23, and B9 show more dispersed views when AI is linked to evaluation support, Bloom's Taxonomy use, learner independence, LOTS support, and critical reflection. These fuzzier boundaries suggest areas where institutional norms and professional expectations remain under development.

Pedagogically, the study provides empirical support for the central thesis of this paper: the necessity of a strategic framework guided by Bloom's Taxonomy and the Dual Approach. The high caution shown by respondents in assessing HOTS (E24, $\bar{X} = 4.05$) versus the perception that AI is effective in supporting LOTS (E23, $\bar{X} = 3.19$) indicates an intuitive division of cognitive labour. This distinction is empirically supported by a highly significant difference confirmed by the paired-samples *t*-test ($t(147) = 9.87, p < 0.001$). This pattern suggests greater caution about using AI in complex cognitive and evaluative tasks. Flipped learning may provide a useful context in which AI supports preparatory activities while contact time emphasises teacher-led higher-order learning and critical reflection.

With respect to RQ1, the findings indicate that educators distinguish acceptable AI support from problematic delegation mainly according to the location of evaluative responsibility. AI appears more acceptable when it assists preparation, routine activity, feedback, or lower-order work while human educators retain interpretation and final judgment. The boundary becomes more restrictive when AI is used in fairness-sensitive or consequential assessment and risks substituting for human evaluative authority.

With respect to RQ2, comparatively crisp boundaries appear around awareness of AI bias and the precedence of human teacher judgment. By contrast, fuzzier boundaries emerge around AI-supported evaluation, learner independence, LOTS support, and critical reflection, where response distributions are more dispersed. These findings suggest that consensus is strongest where accountability and final assessment responsibility are directly implicated, whereas the appropriate scope of AI as a pedagogical support tool remains less settled.

7 Concluding Remarks

This study examined educators' perceptions of the ethical boundaries for using AI in teaching and academic assessment. The central problem was not whether AI should be adopted in higher education, but where educators perceive AI as legitimate support and where they view it as problematic delegation of evaluative responsibility. The findings show that respondents did not uniformly accept or reject AI. Instead, they differentiated between AI as a support tool for routine, preparatory, and lower-order learning activities and AI as a potential substitute for human evaluative judgment.

In direct answer to RQ1, respondents distinguished support from delegation according to whether AI assists learning and assessment while human educators retain interpretation and final evaluative authority. In answer to RQ2, ethical boundaries were comparatively crisp around AI bias and the precedence of human judgment, but remained fuzzy around AI-supported evaluation, learner independence, and critical reflection.

The Dual Approach therefore provides a structured way to relate cognitive complexity and pedagogical context to the locus of evaluative responsibility.

The main contribution of the Dual Approach is to connect a pedagogical distinction with an ethical one. Bloom's Taxonomy helps interpret differences between lower- and higher-order cognitive activities, while the ethical framework clarifies when AI remains supportive and when it risks shifting evaluative authority away from human educators. In this sense, the study does not validate a pedagogical intervention, but offers an interpretive framework for understanding responsibility, trust, fairness, and acceptable AI use in assessment.

These findings have implications for educational practice and institutional governance, but they should be interpreted cautiously given the exploratory single-institution design. Institutions should clarify where AI may support preparation, feedback, and learning activities, and where human evaluative authority must remain explicit. Teacher training should therefore go beyond technical use of AI and include critical evaluation, pedagogical alignment, academic integrity, privacy, fairness, and accountability.

AI-supported assessment therefore depends on maintaining an explicit boundary between assistance and substitution of human judgment.

7.1 *Limitations and Future Work*

A first limitation concerns the single-institution design. Since participants were recruited through UAL, the results may reflect not only individual academic perceptions but also aspects of Portuguese higher-education traditions, UAL's institutional culture, and the local stage of debate on AI in teaching and assessment. Consequently, the findings should not be treated as statistically representative of universities in general or as evidence of a national pattern.

A related limitation concerns potential sampling bias. As the study relied on voluntary participation, respondents may have been more engaged with AI in education, assessment design, or pedagogical innovation than non-respondents. This may have influenced the relatively high levels of AI awareness and intention to adopt AI tools observed in the results. In addition, because the questionnaire was distributed through multiple channels and completed anonymously, the response rate could not be precisely established. These factors limit claims of representativeness and reinforce the interpretation of the study as exploratory rather than generalisable.

Another limitation concerns the respondents' age range. Although the sample included faculty members typically aged between 30 and 65 years, the present study did not analyse whether age influenced perceptions of AI-supported assessment. This may be relevant because trust in automated systems, openness to AI use, and perceptions of responsibility in assessment may vary across career stages and age groups. Future research with larger and more balanced samples should examine whether age or career stage moderates educators' ethical perceptions of AI in teaching and assessment.

The sample was not entirely homogeneous: some participants collaborated with other higher education institutions or research centres, and some international academic backgrounds were represented. This diversity supports analytical transferability, but does not remove the limitations of the single-institution sampling strategy.

Future research should test whether the crisp and fuzzy boundary patterns identified here are reproduced across institutions, disciplines, national contexts, assessment types, and pedagogical models. Multi-institutional studies could also examine whether flipped-learning contexts systematically increase the perceived legitimacy of AI for preparatory activity while preserving stronger human authority in consequential assessment.

We also acknowledge the limitations of using a 5-point Likert scale. Although previous research suggests that increasing the number of response categories has little effect on mean scores or the underlying factor structure, larger scales appear to reduce skewness and provide greater response discrimination (Leung, 2011). Given that our exploratory analysis combines mean scores with response entropy, future research could explore the use of 7- or 10-point Likert scales to determine whether greater response granularity yields additional insights.

Finally, although the study addresses ethical judgment, it does not use a validated multidimensional ethical judgment scale. Future research could combine the present AI-assessment questionnaire with

established instruments for measuring ethical sensitivity and ethical evaluation, allowing a more direct psychometric analysis of ethical judgment in AI-supported assessment contexts.

Data Availability Statement. Data were collected through an anonymous survey. Individual-level data are not publicly available for ethical and confidentiality reasons. Aggregated data may be made available upon reasonable request.

Acknowledgements. This work was funded by national funds through FCT - Fundação para a Ciência e a Tecnologia - project CI2 - Ref^a UID/05567/2023.

Declaration of AI-assisted Tools. We used conversations with ChatGPT, model 5.5 (OpenAI, 2026), as a collaborative support tool at different stages of preparing the manuscript. Regarding the faculty members' questionnaire, ChatGPT was used to support the formulation and refinement of preliminary draft items. The interaction was initiated with the following prompt:

"Please help formulate neutral and academically appropriate questionnaire items for a survey addressed to higher education faculty members about perceptions of AI use in teaching and academic assessment. The items should cover awareness of AI tools, perceived usefulness, academic integrity and ethics, privacy, trust in AI-generated evaluation, human teacher judgment, AI bias, Bloom's Taxonomy, lower-order and higher-order cognitive skills, and flipped learning. The questions should be suitable for a 5-point Likert scale where appropriate, avoid leading or normative wording, and be designed only as draft suggestions to be reviewed and validated by us."

Subsequent prompts were limited to clarification, rewording, and alignment of draft items with the questionnaire sections. All AI-assisted outputs were reviewed, adapted, and approved by the authors, who retained full responsibility for the questionnaire, methodology, analysis, and interpretation. Grammarly (Grammarly, 2026) was later used for readability, grammar, and style; all changes were reviewed before submission.

References

- Akdeniz, H., Clark, T., Roberts, J.L., 2025. Can AI Generate Questions Aligned with Bloom's Taxonomy? A Framework for Gifted Education to Support Teachers. *Journal of Advanced Academics* 36, 100–115. <https://doi.org/10.1177/1932202x251349917>.
- Anderson, L.W., Krathwohl, D.R., 2001. *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. Addison Wesley Longman, New York.
- Evangelista, E.D.L., 2025. Ensuring Academic Integrity in the Age of ChatGPT: Rethinking Exam Design, Assessment Strategies, and Ethical AI Policies in Higher Education. *Contemporary Educational Technology* 17, ep559. <https://doi.org/10.30935/cedtech/15775>.
- Faraon, M., Granlund, V., Rönkkö, K., 2023. Artificial Intelligence Practices in Higher Education Using Bloom's Digital Taxonomy, in: *2023 5th International Workshop on Artificial Intelligence and Education (WAIE)*. Presented at the 2023 5th International Workshop on Artificial Intelligence and Education (WAIE), IEEE, Tokyo, Japan, pp. 53–59. <https://doi.org/10.1109/WAIE60568.2023.00017>.
- Grammarly, 2026. Grammarly [WWW Document]. URL <https://www.grammarly.com/> (accessed 8.5.26).
- Hachoumi, N., Eddabbah, M., Adib, A.E.E., 2025. Enhancing Teaching and Learning in Health Sciences Education through the Integration of Bloom's Taxonomy and Artificial Intelligence. *Informatics and Health* 2, 130–136. <https://doi.org/10.1016/j.infoh.2025.05.002>.
- Hmoud, M., Ali, S., 2024. AIED Bloom's Taxonomy: A Proposed Model for Enhancing Educational Efficiency and Effectiveness in the Artificial Intelligence Era. *The International Journal of Technologies in Learning* 31, 111–128. <https://doi.org/10.18848/2327-0144/cgp/v31i02/111-128>.

- Holmes, W., Bialik, M., Fadel, C., 2023. Artificial Intelligence in Education, in: Stükelberger, C., Duggal, P. (Eds.), *Data Ethics: Building Trust: How Digital Technologies Can Serve Humanity*. Globethics Publications, pp. 621–653. <https://doi.org/10.58863/20.500.12424/4276068>.
- Jobin, A., Ienca, M., Vayena, E., 2019. The Global Landscape of AI Ethics Guidelines. *Nat Mach Intell* 1, 389–399. <https://doi.org/10.1038/s42256-019-0088-2>.
- Kwan, P., Kadel, R., Memon, T., Hashmi, S.S., 2025. Reimagining Flipped Learning via Bloom’s Taxonomy and Student–Teacher–GenAI Interactions. *Education Sciences* 15, 465. <https://doi.org/10.3390/educsci15040465>.
- Lee, H.-P. (Hank), Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., Wilson, N., 2025. The Impact of Generative AI on Critical Thinking: Self-Reported Reductions in Cognitive Effort and Confidence Effects From a Survey of Knowledge Workers, in: *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems, CHI ’25*. Association for Computing Machinery, New York, NY, USA, pp. 1–22. <https://doi.org/10.1145/3706598.3713778>.
- Leung, S.-O., 2011. A Comparison of Psychometric Properties and Normality in 4-, 5-, 6-, and 11-Point Likert Scales. *Journal of Social Service Research* 37, 412–421. <https://doi.org/10.1080/01488376.2011.580697>.
- Li, N., 2024. AI as a Threat to Education? Contrasting GPT-3 and Google in Answering Questions Along Bloom’s Taxonomy of Educational Objectives. *Medicon Eng. Themes* 6, 198–205. <https://doi.org/10.55162/mcet.06.198>.
- López-Villanueva, D., Santiago, R., Palau, R., 2024. Flipped Learning and Artificial Intelligence. *Electronics* 13, 3424. <https://doi.org/10.3390/electronics13173424>.
- Luckin, R., Holmes, W., Griffiths, M., Corcier, L.B., 2016. *Intelligence Unleashed: An Argument for AI in Education*. Pearson, London.
- Matos, J.M., Dediu, A.-H., 2026. Using AI for Research and Educational Support: Enhancing the Design of Computer-Based Evaluations, in: Alvarez, I., Arias-Oliva, M., Dediu, A.-H., Silva, N. (Eds.), *Ethical and Social Impacts of Information and Communication Technology*, *Lecture Notes in Computer Science*. Springer Nature Switzerland, Cham, pp. 105–115. https://doi.org/10.1007/978-3-032-01429-0_10.
- Murphy, S., O’Neill, P., 2022. *Assessing Writing to Support Learning: Turning Accountability Inside Out*. Routledge, New York. <https://doi.org/10.4324/9781003296140>.
- Naqvi, W.M., Ganjoo, R., Rowe, M., Pashine, A.A., Mishra, G.V., 2025. Critical Thinking in the Age of Generative AI: Implications for Health Sciences Education. *Frontiers in Artificial Intelligence* 8. <https://doi.org/10.3389/frai.2025.1571527>.
- Olarte-Pascual, C., Pelegrín-Borondo, J., Reinares-Lara, E., Arias-Oliva, M., 2021. From Wearable to Insideable: Is Ethical Judgment Key to the Acceptance of Human Capacity-Enhancing Intelligent Technologies? *Computers in Human Behavior* 114, 106559. <https://doi.org/10.1016/j.chb.2020.106559>.
- OpenAI, 2026. ChatGPT (GPT-5.5) [WWW Document]. URL <https://chat.openai.com/> (accessed 8.5.26).
- Reidenbach, R.E., Robin, D.P., 1990. Toward the Development of a Multidimensional Scale for Improving Evaluations of Business Ethics. *Journal of Business Ethics* 9, 639–653. <https://doi.org/10.1007/BF00383391>.
- Roll, I., Wylie, R., 2016. Evolution and Revolution: Combining Intelligent Tutoring Systems and Collaboration. *International Journal of Artificial Intelligence in Education* 26, 883–902. <https://doi.org/10.1007/s40593-016-0103-9>.

-
- Shannon, C.E., 1948. A Mathematical Theory of Communication. *Bell System Technical Journal* 27, 379–423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>.
- Shawver, T.J., Sennetti, J.T., 2009. Measuring Ethical Sensitivity and Evaluation. *Journal of Business Ethics* 88, 663–678. <https://doi.org/10.1007/s10551-008-9973-z>.
- Tang, V., Painho, M., Vorobeva, D., 2025. Integrating Tailored Generative AI into the Flipped Classroom: A Pilot Implementation in Higher Education. *Innovations in Education and Teaching International* 1–19. <https://doi.org/10.1080/14703297.2025.2523898>.
- UNESCO, 2021. AI and Education: Guidance for Policy-Makers. UNESCO. <https://doi.org/10.54675/PCSP7350>.
- West, F.M., Van der Westhuizen, C.P., 2026. Artificial Intelligence and the Flipped Classroom: Lecturer Perceptions in a South African Private Higher Education Institution. *International Journal of Learning, Teaching and Educational Research* 25. <https://doi.org/10.26803/ijlter.25.5.46>.
- Zawacki-Richter, O., Marín, V.I., Bond, M., Gouverneur, F., 2019. Systematic Review of Research on Artificial Intelligence Applications in Higher Education – Where are the Educators? *International Journal of Educational Technology in Higher Education* 16, 39. <https://doi.org/10.1186/s41239-019-0171-0>.
- Zhao, N., Fu, Q., 2025. Under the Framework of Deep Learning Cognitive Theory and Bloom's Taxonomy: Investigating the Role of Artificial Intelligence in Fostering Higher-order Thinking in Engineering Education, in: *Proceedings of the 7th International Conference on Computer Science and Technologies in Education*. IEEE, Wuhan, China, pp. 892–896. <https://doi.org/10.1109/cste64638.2025.11092073>.