International Journal of Modern Education
EISSN: 2637-0905
IJMOE
Volume: 8 Issue: 29 (March 2026) PP. 945-960

INTERNATIONAL JOURNAL OF MODERN EDUCATION (IJMOE)
www.gaexcellence.com/ijmoe

WHEN HIGH SCORES DO NOT MEAN ETHICAL UNDERSTANDING: REVISITING AI ETHICS MEASUREMENT AMONG MATRICULATION STUDENTS

Amal Harun[1]*, Ahmad Zamri Khairani[2]

[1]School of Educational Studies, Universiti Sains Malaysia, Malaysia
amalharun@student.usm.edu.my https://orcid.org/0009-0003-1293-9779
[2]School of Educational Studies, Universiti Sains Malaysia, Malaysia
ahmadzamri@usm.my https://orcid.org/0000-0002-3832-8103
*Corresponding Author

Article Info:

Article history:
Received date: 27.01.2026
Revised date: 12.02.2026
Accepted date: 09.03.2026
Published date: 26.03.2026

To cite this document:

Harun, A., Khairani, A. Z. (2026). When High Scores Do Not Mean Ethical Understanding: Revisiting Ai Ethics Measurement Among Matriculation Students. *International Journal of Modern Education, 8*(29), 945-960.

Abstract:

The rapid integration of Artificial Intelligence AI into educational contexts has led to the development of numerous AI literacy instruments aimed at assessing students' ethical awareness and responsibility. This focus is particularly salient in matriculation colleges, which function as a critical transitional stage between secondary schooling and university education, where students begin to encounter heightened academic autonomy and ethically consequential uses of generative AI. However, most existing measures rely on self-report formats, raising concerns about the validity of inferences drawn from high ethics scores. Thus, this study examines whether self-reported AI ethics scores meaningfully reflect students' ethical understanding and practices in authentic academic contexts. A convergent mixed-methods design was employed, combining a quantitative survey of 355 matriculation students using the ethics-related construct of the Meta AI Literacy Scale (MAILS) with qualitative semi-structured interviews involving three experienced lecturers. Quantitative results indicated statistically high mean score across all ethics subdimensions. In contrast, qualitative findings revealed systematic discrepencies between perceived and enacted ethical practices, including verbatim use of AI-generated content, limited recognition of ethical boundaries, and recurring academic integrity concerns. These findings suggest that self-report measures may overestimate ethical competence and inadequately capture ethical reasoning as enacted in situationally complex learning environments. From a measurement perspective, the results highlight a risk of construct under-representation when ethical AI use is assessed solely through questionnaires. This study contributes to AI literacy research by examining the construct validity of ethics-related self-report scores and

International Journal of Modern Education
IJMOE
EISSN: 2637-0905
Volume: 8 Issue: 29 (March 2026) PP. 945-960

demonstrating the value of behaviourally grounded mixed-methods approaches for strengthening AI ethics measurement.

DOI: 10.35631/IJMOE.829056

Keywords:

AI Ethics, AI Literacy Measurement, Self-Report Bias, Situational Judgement Test, Matriculation Students

Introduction

Recent developments in AI literacy research have led to the widespread use of self-report instruments to assess students' ethical awareness and responsibility in using Artificial Intelligence (AI) tools. Across educational contexts, these instruments frequently report high levels of perceived ethical competence, suggesting that students consider themselves capable of using AI responsibly. However, an important measurement question remains insufficiently examined: what do high self-reported AI ethics scores actually represent? Without behavioural or contextual validation, such scores risk being interpreted as evidence of ethical understanding when they may instead reflect attitudinal endorsement or socially desirable responding.

From a measurement perspective, this concern is not trivial. Self-report instruments assessing ethical constructs are particularly vulnerable to social desirability bias, whereby respondents provide answers aligned with perceived norms rather than actual behaviour (Miller & Research, 2011; Paulhus, 2018; van de Mortel, 2008). More fundamentally, validity theory emphasises that test scores must accurately represent the intended construct (Messick, 1995).When ethical AI use is operationalised primarily through agreement with general statements about responsibility or awareness, there is a risk of construct underrepresentation, as questionnaire items may capture abstract moral endorsement while failing to assess situational judgement, boundary recognition, and decision-making under pressure. In this sense, high ethics scores may reflect perception rather than enacted competence.

This distinction aligns with broader literature documenting a persistent attitude–behaviour gap, in which individuals' reported ethical beliefs do not consistently translate into ethical action in contextually complex situations. In educational settings, students may express strong support for academic integrity while still engaging in dishonest practices when facing time constraints, performance pressure, or ambiguous rules. The emergence of generative AI (GenAI) amplifies

International Journal of Modern Education
IJMOE
EISSN: 2637-0905
Volume: 8 Issue: 29 (March 2026) PP. 945-960

this challenge. Unlike earlier technologies, GenAI can produce sophisticated academic outputs with minimal user input, blurring the boundaries between assistance, collaboration, and misconduct (Ka & Chan, 2025; Nelson et al., 2025; Ortiz-bonnin & Blahopoulou, 2025; Pudasaini et al., 2025). Ethical AI use therefore requires more than agreement with ethical principles; it requires context-sensitive judgement and deliberate application of those principles in authentic learning situations.

Against this backdrop, the present study adopts a measurement-oriented perspective to examine how AI ethics scores should be interpreted among Malaysian matriculation students. Using a convergent mixed-methods design, the study combines quantitative self-report data with qualitative insights from lecturers to explore whether reported AI ethics scores meaningfully represent students' ethical understanding and practices when using generative AI in academic contexts. By examining both students' reported awareness and lecturers' observations of students' AI-related practices, the study seeks to provide a more grounded interpretation of AI ethics measurement and contribute to ongoing discussions concerning construct validity in AI ethics assessment.

To guide this investigation, the study addresses the following research questions:

RQ1: What levels of AI ethics awareness are reflected in matriculation students' self-reported scores on the ethics-related subscale of the Meta AI Literacy Scale (MAILS)?

RQ2: How do lecturers describe students' ethical understanding and practices when using generative AI in academic tasks?

RQ3: To what extent do students' self-reported AI ethics scores align with lecturers' observations of students' AI-related practices?

Literature Review

Ethical AI in Education: From Principles to Practice

Generative Artificial Intelligence (GenAI) is increasingly integrated into students' academic workflows, including writing, summarising, translation, and problem-solving tasks (Hasib & Islam, 2025). While these tools can enhance efficiency and access to information, they simultaneously introduce ethical concerns related to authorship, transparency, academic integrity, and over-reliance on automated systems (Schiff, 2022). Unlike earlier educational technologies, GenAI systems can independently generate sophisticated and contextually appropriate academic content, making the boundary between assistance and misconduct increasingly ambiguous (Sharma & Panja, 2025).

Contemporary AI literacy frameworks consistently position ethics as a central dimension, emphasising responsible use, transparency, fairness, and accountability (Floridi & Cowls, 2022; Long & Magerko, 2020; Mustapha et al., 2025; UNESCO, 2024). However, these frameworks often conceptualise ethics at the level of principle endorsement. The translation of abstract ethical values into context-sensitive academic decision-making remains less clearly operationalised, particularly within transitional educational settings such as matriculation programmes.

International Journal of Modern Education
IJMOE
EISSN: 2637-0905
Volume: 8 Issue: 29 (March 2026) PP. 945-960

The Measurement Challenge: Self-Report and Social Desirability

Most empirical studies assessing ethical AI use rely on self-report questionnaires using Likert-scale formats. These instruments typically measure students' agreement with statements concerning responsible or ethical AI behaviour. While such measures are efficient and scalable, their suitability for assessing ethical constructs warrants closer scrutiny.

Self-report instruments measuring ethical attitudes are particularly vulnerable to social desirability bias, whereby respondents provide answers that align with perceived social norms rather than actual behaviour (Miller & Research, 2011; Paulhus, 2018; van de Mortel, 2008). In educational contexts, where academic integrity is strongly institutionalised as a normative expectation, students may feel compelled to present themselves as ethically responsible regardless of their actual practices. As a result, high self-reported ethics scores may reflect aspirational identity or moral self-concept rather than enacted ethical competence.

This concern is well documented in academic integrity research, which demonstrates that students may endorse ethical principles while still engaging in dishonest behaviours under situational pressures such as time constraints, performance anxiety, or ambiguous guidelines (Hasib & Islam, 2025; Leaton Gray et al., 2025). Similar dynamics are likely in the context of GenAI use, where ethical boundaries are evolving and institutional policies remain inconsistent.

Ethical Awareness Versus Ethical Judgement: The Attitude–Behaviour Gap

A key limitation of questionnaire-based ethics measurement lies in its tendency to capture ethical awareness rather than ethical judgement. Ethical AI use requires more than recognising general principles; it involves identifying when ethical considerations are relevant, weighing competing demands, and making contextually appropriate decisions (Alnsour et al., 2025; Khan et al., 2025; Laine et al., 2025; Mujtaba, 2025).

The broader literature on the attitude–behaviour gap suggests that individuals' stated ethical beliefs do not consistently predict their actions in complex or pressured situations (Yan et al., 2025). Ethical behaviour is influenced not only by moral endorsement but also by contextual cues, perceived risk, peer norms, and institutional clarity (Shaw, 2025). In the case of GenAI, students may believe their AI use is ethical because they conceptualise it as an advanced search tool, even when such use involves verbatim copying or substitution of personal effort.

This distinction is particularly salient in matriculation contexts, where students are transitioning toward greater academic autonomy while institutional guidelines on AI use remain underdeveloped. Under such conditions, self-reported ethical awareness may coexist with practices that educators consider ethically problematic.

Construct Representation and the Case for Context-Based Assessment

From a measurement theory perspective, these concerns relate to issues of construct validity. According to Messick's (1995) unified framework, validity concerns the degree to which evidence supports the interpretation of test scores as representations of the intended construct. When AI ethics is operationalised primarily through agreement-based questionnaire items, there is a risk of construct underrepresentation, whereby important facets of ethical

International Journal of Modern Education
IJMOE
EISSN: 2637-0905
Volume: 8 Issue: 29 (March 2026) PP. 945-960

competence, such as situational judgement, boundary recognition, and decision-making under pressure are insufficiently captured (Alnsour et al., 2025; Asamoah & Amarteifio, 2025; Culture & Humphreys, 2025).

If high ethics scores primarily reflect students' endorsement of abstract moral principles rather than their ability to apply ethical reasoning in authentic situations, then interpreting such scores as indicators of ethical competence may be overstated. Self-report instruments are particularly susceptible to this limitation because respondents often evaluate their own attitudes or intentions in generalised terms rather than demonstrate how they would act in concrete contexts (Podsakoff et al., 2003). As a result, these measures may capture normative agreement with ethical ideals without adequately representing the decision-making processes required in real academic situations involving generative AI.

Addressing this limitation requires assessment approaches capable of capturing context-dependent ethical reasoning. Situational Judgement Tests (SJTs) represent one such approach. SJTs present respondents with realistic scenarios and require them to evaluate, rank, or select behavioural responses to situational dilemmas (Patterson et al., 2016). By embedding ethical evaluation within contextually grounded situations, SJTs more closely approximate the cognitive and judgement processes involved in real-world decision-making. Consequently, SJTs provide stronger construct representation of ethical judgement as enacted reasoning rather than abstract moral endorsement. Within AI-related academic contexts, SJTs may therefore function as a complementary assessment method alongside self-report measures, enabling more robust and defensible interpretations of ethics-related scores.

Methodology

This study employed a convergent mixed-methods design to examine the interpretive alignment between students' self-reported AI ethics scores and lecturers' observations of students enacted ethical practices. In a convergent design, quantitative and qualitative data are collected during a similar phase of the research process, analysed independently, and subsequently integrated at the interpretation stage to identify patterns of convergence or divergence across data strands (Creswell & Plano Clark, 2018). This design was considered appropriate because the purpose of the study was not to use one dataset to explain the other, but rather to evaluate the interpretive meaning of ethics-related survey scores by comparing them with independently generated behavioural evidence. Such an approach enables researchers to triangulate findings from different data sources and strengthen the validity of score interpretation in educational measurement (Fetters et al., 2013). Evidence of convergence between strands would support the construct representation of the survey measure, whereas systematic divergence may indicate potential construct underrepresentation or limitations in how ethical reasoning is captured through self-report instruments (Messick, 1995; Kane, 2013). The quantitative component involved 355 matriculation students drawn from a population of approximately 2,500 students enrolled in a Malaysian matriculation college in Negeri Sembilan during the 2025/2026 academic session. The sample size exceeded the minimum requirement recommended by Krejcie and Morgan (1970) for a population of this size. Participants were recruited using convenience sampling across multiple academic streams, including physical sciences, life sciences, accounting, and computer science.

International Journal of Modern Education
IJMOE
EISSN: 2637-0905
Volume: 8 Issue: 29 (March 2026) PP. 945-960

The qualitative strand comprised three purposively selected matriculation lecturers. Selection criteria included (a) a minimum of ten years of teaching experience at the matriculation level, (b) active involvement in continuous assessment of student coursework, and (c) direct experience and integrating AI in teaching and learning. The smaller qualitative sample reflects the focused interpretive aim of the qualitative strand, which sought context-rich insights from experienced evaluators of student academic practice rather than population-level representation. Recurring patterns across interviews indicated sufficient thematic saturation for the analytic purposes of the study.

The ethics-related construct was measured using the ethics subscale of the Meta AI Literacy Scale (MAILS) developed by Carolus et al. (2023). To ensure contextual relevance to the Malaysian matriculation setting, minor adaptations were made to the wording of several items. Specifically, references to general AI usage were adjusted to reflect students' academic use of generative AI tools within learning and assessment contexts. These adaptations were limited to linguistic and contextual clarification rather than substantive changes to the construct being measured.

For example, items referring broadly to AI applications were rephrased to emphasise AI use in academic tasks such as assignments, problem-solving, and learning activities, which more accurately reflect the contexts in which matriculation students interact with generative AI technologies. The underlying construct of ethical awareness such as considering the societal consequences of AI use and incorporating ethical considerations into AI-related decisions, remained consistent with the original instrument.

Prior to administration, the adapted items were reviewed to ensure clarity and appropriateness for the target population. The adapted scale was administered using a five-point Likert response format ranging from strongly disagree to strongly agree. In the present sample, the ethics subscale demonstrated high internal consistency reliability (Cronbach's $\alpha = .91$), indicating that the adapted items functioned consistently in measuring the intended construct of ethical awareness related to AI use. (George & Mallery, 2003).

Quantitative data were collected via an anonymous online survey. Data analysis was conducted using descriptive statistics, including means, standard deviations, frequency distributions, and percentage agreement rates. Because the study aimed to examine the interpretive meaning of ethics scores rather than to test predictive or causal relationships, inferential modelling was not performed. Emphasis was placed on score patterns and distributional characteristics to assess the overall level of self-reported ethical endorsement.

Qualitative data were obtained through semi-structured face-to-face interviews focusing on lecturers' observations of students' ethical understanding, common AI-related academic practices, and perceived discrepancies between reported and enacted AI use. Interviews were audio-recorded with consent and transcribed verbatim.

Interview data were analysed using reflexive thematic analysis following Braun and Clarke's (2006) six-phase framework: (1) data familiarisation, (2) systematic coding, (3) generation of initial themes, (4) theme review and refinement, (5) defining and naming themes, and (6) production of the analytical narrative. An inductive coding strategy was employed to allow themes to emerge from participants' accounts rather than being imposed by predetermined

International Journal of Modern Education
EISSN: 2637-0905
IJMOE
Volume: 8 Issue: 29 (March 2026) PP. 945-960
theoretical categories. This approach enabled systematic identification of recurring patterns related to ethical reasoning, boundary recognition, and academic integrity practices.
Integration of the quantitative and qualitative findings was conducted at the interpretation stage, consistent with the principles of a convergent parallel mixed-methods design (Creswell & Plano Clark, 2018). Both data strands were analysed independently to preserve the methodological integrity of each source of evidence prior to comparison. The results were subsequently integrated using a narrative weaving strategy, in which patterns observed in the quantitative scores were systematically juxtaposed with themes emerging from the qualitative data (Fetters et al., 2013). This integrative approach enabled the identification of areas of convergence, complementarity, and divergence across the two strands of evidence, thereby supporting a more nuanced interpretation of students' ethical reasoning in the use of AI.

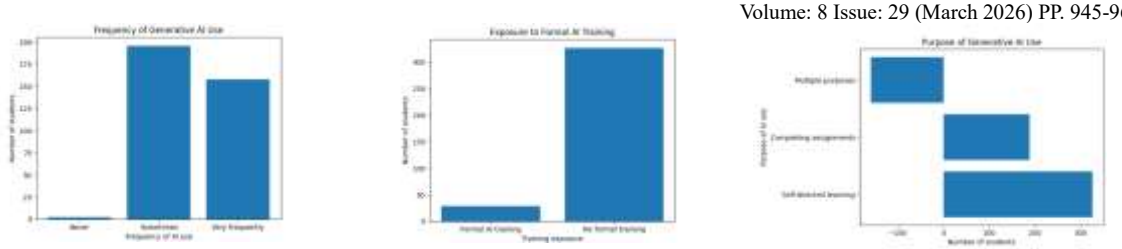
Particular analytic attention was directed toward instances where statistically high mean ethics scores appeared inconsistent with lecturers' observations of students' AI-related practices. Rather than treating these differences as contradictions between datasets, they were interpreted as potentially indicating construct underrepresentation within the self-report measure. In other words, elevated survey scores may reflect students' endorsement of ethical principles at an abstract level rather than their application of ethical reasoning in practice. Examining such discrepancies enabled the integration process to serve a broader measurement purpose, whereby the mixed-methods design functioned as a mechanism for critically evaluating and strengthening the interpretation of the construct captured by the survey scores.

Ethical approval was obtained prior to data collection. Participation was voluntary, and informed consent was obtained from all participants. Students completed the survey anonymously, and no identifying information was collected. Interview participants provided written informed consent and were assured of confidentiality. Pseudonyms were used in reporting qualitative findings, and all data were securely stored in accordance with institutional ethical guidelines.

Results

A total of 355 matriculation students participated in the survey. The sample comprised 117 male and 238 female students from diverse academic streams, including physical sciences (n = 154), life sciences (n = 128), accounting (n = 64), and computer science (n = 9). The use of AI was widespread among respondents. In Figure 1(a), only two students reported never using AI, while the majority indicated using AI either sometimes (n = 196) or very frequently (n = 159). In Figure 1(b), only 29 students reported having received formal instruction related to AI, whereas the vast majority indicated no formal training and relied primarily on self-exploration. In terms of usage context, students reported accessing AI mainly through smartphones, followed by computers and tablets. In Figure 1(c), AI was most commonly used for self-directed learning and for completing academic assignments, with some students indicating broader use across multiple aspects of daily activities.

International Journal of Modern Education
IJMOE
EISSN: 2637-0905
Volume: 8 Issue: 29 (March 2026) PP. 945-960



A B C

Figure 1: Patterns Of Generative AI Use Among Matriculation Student Respondents

Responses to the ethics-related items indicated statistically high mean scores across all ethics subdimension. Across items, approximately two-thirds of respondents selected *agree* or *strongly agree*, with mean scores clustering slightly above four on a five-point scale (**Table 1**). Very few respondents selected the lowest response categories, indicating strong normative endorsement of ethical awareness and consideration in AI use. Based on the mean score classification guidelines proposed by Alkharusi (2022), the overall level of AI ethics reflected in these scores can be categorised as high.

Table 1: Mean Score for AI Ethics Use

Item	Statement	N	Mean	SD	Agree %
E1	Weigh the consequences of AI use for society	352	4.05	0.75	69.4
E2	Incorporate ethical considerations in AI use	354	4.00	0.78	67.2
E3	Analyse ethical implications of AI applications	355	4.02	0.77	67.0

On the surface, the results suggest that students possess a clear understanding of ethical expectations related to generative AI use and perceive themselves as capable of acting responsibly within academic settings.

Qualitative Findings: Lecturers' Observations of Ethical Understanding and Practice

In contrast, thematic analysis of semi-structured interviews with experienced matriculation lecturers revealed a consistent pattern in how students engage with generative AI in academic tasks. Across interviews, lecturers described AI use as widespread and routine, cutting across assignments such as report writing, problem solving, translation, and presentation preparation. Students generally lack clarity about ethical boundaries, frequently perceiving AI as a shortcut for task completion and equating it with traditional search tools.

Concerns related to academic integrity emerged as the most dominant theme. Lecturers reported frequent instances where students submitted AI-generated content verbatim, without adaptation or demonstrated understanding. Such practices were often identifiable *through* mismatches between the sophistication of the language used and students' actual academic level, as well as students' inability to explain or justify the content of their submissions when questioned. AI was commonly perceived by students as a shortcut, particularly under time pressure or when foundational subject knowledge was weak. These observations suggest that AI use was closely tied to issues of honesty, personal effort, and responsibility for one's own learning.

A critical contextual factor highlighted by lecturers was the absence of formal institutional guidelines on AI use at the matriculation level. Without clear policies, lecturers relied largely on personal judgement and teaching experience to decide whether a particular use of AI should

International Journal of Modern Education
IJMOE
EISSN: 2637-0905
Volume: 8 Issue: 29 (March 2026) PP. 945-960

be considered ethical or problematic. This resulted in inconsistent practices across courses, particularly in continuous assessment. Lecturers expressed concern that such ambiguity could compromise fairness, as not all instances of AI use were detected or addressed in the same way. Lecturers also consistently described low levels of ethical awareness among students. Many students were reported to be unaware that using AI-generated content without modification or disclosure constituted an ethical issue. Ethical awareness tended to emerge only after students were confronted or reprimanded, rather than through prior understanding. Responses to such feedback varied, ranging from anxiety to indifference, indicating that ethical awareness was often reactive rather than internalised.

Despite these concerns, lecturers did not reject the use of AI in education. Instead, AI was viewed as a supportive tool with legitimate pedagogical value, particularly for saving time, clarifying concepts, and supporting independent learning. However, lecturers strongly emphasised that AI should not replace students' critical thinking or the educator's role. Students were seen as needing explicit guidance to question, verify, and reflect on AI outputs, while lecturers positioned themselves as ethical gatekeepers who must remain *above AI* in guiding learning and assessment decisions. **Table 2** summarises the key themes and illustrative quotes that highlight the gap between students' self-reported ethical awareness and observed practices.

Table 2: Themes and Illustrative Quotes Highlighting the Ethics–Practice Gap

Theme	Illustrative Lecturer Quotes
Widespread AI Use without Ethical Clarity	*"They see AI as a tool to finish the work quickly. They do not really think about whether it replaces their own thinking."*
	"They treat AI just like Google. Copying AI output does not seem wrong to them."
Academic Integrity Concerns	*"The language suddenly becomes very advanced, but when you ask them to explain, they cannot."*
	"When they are rushing or do not understand the topic, AI becomes the fastest solution."
Absence of Institutional Guidelines	*"We in matrics do not have any clear guidelines, so it is hard to say whether this is ethical or not."*
	"Most cases are handled through advice or warnings. There is no standard procedure."
Low Ethical Awareness	*"Most of them do not realise that using AI like this is wrong."*
	"They only start to understand when they get caught or when I explain it to them."
AI as Support, Not Replacement	*"AI can help, but students cannot trust it one hundred percent."*
	"Teachers still need to be above AI. We guide students on how to use it ethically."
	"ChatGPT can give answers, but it cannot guide students the way a teacher does. As a lecturer, I don't just teach content, I correct their

thinking, their attitude, and sometimes their values. Definitely, AI cannot replace that human role."

These responses indicate that lecturers perceive AI as a supplementary tool rather than a substitute for teachers, particularly in areas requiring ethical judgement, contextual understanding, and moral guidance. While AI is acknowledged for its efficiency in information delivery, participants consistently emphasised the irreplaceable human role of teachers in interpreting student behaviour, fostering critical thinking, and guiding ethical academic practices.

Discussion

This study examined whether high self-reported AI ethics scores meaningfully reflect students' ethical understanding as enacted in academic practice. Although survey results indicated strong perceived ethical capability, lecturers' accounts revealed a consistent misalignment between reported ethics and observed behaviour. Rather than demonstrating robust ethical competence, high scores appeared to reflect general ethical awareness that did not reliably translate into ethical judgement during actual use of generative AI. These findings are consistent with (Auwal, 2025), who observed that students tend to prioritise the usefulness of AI-generated outputs over understanding the underlying processes. This outcome-oriented orientation suggests that ethical considerations, particularly transparency and explainability, are often overlooked when AI is perceived as efficient or accurate, indicating a largely instrumental and non-critical approach to AI use in educational contexts.

From a measurement perspective, this finding reveals an important limitation of self-report instruments when they are used to assess complex and context-dependent constructs such as ethics. Based on Messick's validity framework (Messick, 1995), test scores should accurately represent the construct being measured. In this study, the mismatch between survey responses and observed practices suggests construct under-representation, where questionnaire items capture ethical agreement; nonetheless, they do not adequately assess situational judgement or moral reasoning (Theoharakis et al., 2025). As a result, high ethics scores may reflect socially desirable answers or general moral beliefs rather than students' actual ability to handle ethical dilemmas in everyday academic use of AI.

The findings also indicate that students often think about ethics only after they are reminded or corrected, rather than using ethical judgment naturally when using AI. In many cases, ethical concerns only come up when lecturers point them out. This suggests that while students may know ethical rules, they do not automatically apply them during everyday AI use. Being ethical is not just about knowing what is right; nevertheless, it is also about recognising when ethical issues matter and acting on them in real situations (Țîru et al., 2025). When AI is used regularly and mainly to save time or effort, students may overlook ethical issues unless these are clearly explained and highlighted.

Institutional context further amplifies this misalignment. The absence of clear and standardised guidelines on acceptable AI use creates ambiguity for both students and lecturers. In such conditions, students may reasonably interpret generative AI as an extension of conventional information tools, while lecturers assume the role of informal ethical gatekeepers. As a result,

self-reported ethics scores may be anchored in idealised norms rather than shared behavioural expectations, complicating their interpretation.

Taken together, these findings raise concerns about the adequacy of questionnaire-based measures for assessing AI ethics. While self-report instruments are useful for capturing perceived awareness and attitudes, they appear limited in their ability to assess ethical judgement as enacted in practice. This limitation is particularly salient in AI-related contexts, where ethical decisions are rarely binary and often involve trade-offs between learning, efficiency, and accountability.

To address this gap, the study highlights the value of SJTs as a complementary assessment approach. SJTs assess decision-making within realistic scenarios and offer stronger construct representation by operationalising ethics as context-sensitive choice rather than self-perception. In AI-related contexts, SJT-based items can present plausible academic dilemmas and examine how students prioritise ethical considerations under competing demands, thereby aligning assessment more closely with real-world practice (Ortiz-bonnin & Blahopoulou, 2025).

Several limitations should be acknowledged. The use of convenience sampling limits generalisability, and qualitative insights were derived from lecturers' perspectives rather than direct observation of students' reasoning processes. However, the convergent mixed-methods design strengthens interpretive validity by consistently revealing misalignment between reported ethics and observed practice.

Overall, this study cautions against interpreting high self-reported AI ethics scores as direct evidence of ethical understanding or behaviour. The findings support a shift toward multi-method assessment approaches that combine self-report measures with behaviourally grounded tools to capture ethical reasoning as it is enacted in authentic academic contexts.

Limitation of the Study

Despite its contributions, several limitations should be acknowledged. First, the study employed convenience sampling within a single matriculation college, which may limit the generalisability of the findings to other educational contexts or institutions. Students' experiences with generative AI and institutional expectations may vary across different colleges, universities, or national education systems. Future studies could extend this research by including multiple institutions or employing probability sampling methods to improve representativeness.

Second, the qualitative component relied on interviews with a small number of lecturers, which may not fully capture the diversity of educators' perspectives regarding students' ethical use of AI. Although the participants were selected based on extensive teaching experience and recurring themes emerged across interviews, a larger qualitative sample or additional data sources such as classroom observations or student interviews could provide richer insight into students' ethical reasoning processes.

Third, the quantitative analysis focused primarily on descriptive statistics to examine patterns in ethics scores, as the primary objective of the study was to evaluate the interpretive meaning of these scores rather than to test causal relationships. Future research could employ more

International Journal of Modern Education
IJMOE
EISSN: 2637-0905
Volume: 8 Issue: 29 (March 2026) PP. 945-960

advanced statistical analyses or longitudinal designs to examine how ethical reasoning develops over time and how it relates to actual AI usage behaviour.

Finally, the study examined lecturers' observations of students' AI-related practices rather than direct behavioural data from students themselves. While lecturers provide valuable insight into academic practices within assessment contexts, future research could incorporate scenario-based assessments or situational judgement tests to capture students' ethical reasoning in more behaviourally grounded ways.

Taken together, these limitations highlight the need for multi-method approaches to assessing AI ethics, combining self-report measures with behavioural or scenario-based assessments to provide a more comprehensive understanding of students' ethical reasoning in AI-supported learning environments.

Conclusion

This study highlights a critical limitation in current approaches to measuring AI ethics in education and underscores the need for more robust and context-sensitive measurement tools. The findings reveal that high self-reported AI ethics scores do not necessarily reflect students' ethical understanding as enacted in real academic contexts, as questionnaire-based measures may fail to capture key aspects of ethical judgement, such as boundary recognition, situational reasoning, and decision-making under pressure. From a measurement perspective, this highlights the need to advance assessment design beyond self-report instruments through the development of behaviourally grounded tools, such as situational judgement tests, that operationalise AI ethics as applied reasoning within realistic academic scenarios and support more defensible interpretations of ethics-related scores. Beyond measurement, the results also have notable implications for curriculum development. This suggests that ethical AI use should not be treated as implicit knowledge; rather, it is explicitly embedded through structured instruction, scenario-based learning, and aligned assessment practices to ensure that ethical awareness is translated into consistent and responsible AI use across students' educational trajectories.

Acknowledgements:	The authors would like to express their sincere appreciation to the Distinguished Chemistry lecturers and Subject Matter Experts (SMEs) in Biology who generously participated in this study by sharing their experiences with students regarding the integration of artificial intelligence in teaching and learning. The authors also extend their gratitude to the students of Negeri Sembilan Matriculation College for their commitment and cooperation in completing the research questionnaire.
Funding Statement:	No Funding
Conflict of Interest Statement:	The authors declare that there is no conflict of interest regarding the publication of this paper. All authors have contributed to this work and approved the final version of the manuscript for submission to the International Journal of Modern Education (IJMOE).

International Journal of Modern Education
IJMOE
EISSN: 2637-0905
Volume: 8 Issue: 29 (March 2026) PP. 945-960

Ethics Statement:	This study was conducted in accordance with established ethical guidelines for research involving human participants. Informed consent was obtained from all participants prior to data collection. Students provided consent before completing the questionnaire, and all interview participants gave written informed consent before participating in the interviews. Participation was voluntary, and confidentiality and anonymity were assured throughout the study.
Author Contribution Statement:	Both authors contributed substantially to the development of this manuscript. Author A was responsible for the conceptualization, study design, and overall supervision of the research. Author B conducted the data collection, data analysis, and interpretation of the findings. Both authors contributed to the drafting and critical revision of the manuscript and approved the final version prior to submission.

International Journal of Modern Education
EISSN: 2637-0905
IJMOE
Volume: 8 Issue: 29 (March 2026) PP. 945-960

References

Alkharusi, H. (2022). A descriptive analysis and interpretation of data from Likert scales in educational and psychological research. Indian Journal of Psychology and Education, 12(2), 13–16.

Alnsour, M. M., Qouzah, L., Aljamani, S., Alamoush, R. A., & AL-Omiri, M. K. (2025). AI in education: enhancing learning potential and addressing ethical considerations among academic staff—a cross-sectional study at the University of Jordan. *International Journal for Educational Integrity*, *21*(1). https://doi.org/10.1007/s40979-025-00189-4

Asamoah, M. K., & Amarteifio, J. (2025). Exploring the role and ethical concerns of generative AI in doctoral education at the University of Ghana. *SN Social Sciences*, *5*(9), 1–27. https://doi.org/10.1007/s43545-025-01184-9

Auwal, A. M. (2025). Autonomy versus algorithm: a replication study of student perspectives on AI ethical boundaries. *International Journal of Educational Technology in Higher Education*, *22*(1). https://doi.org/10.1186/s41239-025-00570-w

Culture, A. I., & Humphreys, D. (2025). *AI ' s Epistemic Harm : Reinforcement Learning , Collective*. 1–27.

Floridi, L., & Cowls, J. (2022). A unified framework of five principles for AI in society. *Machine Learning and the City: Applications in Architecture and Urban Design*, *1*, 535–545. https://doi.org/10.1002/9781119815075.ch45

Hasib, M., & Islam, M. S. (2025). How University students in Bangladesh engage with ChatGPT: A qualitative study. *Plos One*, *20*(9 September), 1–18. https://doi.org/10.1371/journal.pone.0333089

Ka, C., & Chan, Y. (2025). In understandings of academic misconduct. *Education and Information Technologies*, *30*, 8087–8108.

Khan, S., Mazhar, T., Shahzad, T., Khan, M. A., Rehman, A. U., Saeed, M. M., & Hamam, H. (2025). Harnessing AI for sustainable higher education: ethical considerations, operational efficiency, and future directions. *Discover Sustainability*, *6*(1). https://doi.org/10.1007/s43621-025-00809-6

Laine, J., Minkkinen, M., & Mäntymäki, M. (2025). Understanding the Ethics of Generative AI: Established and New Ethical Principles. *Communications of the Association for Information Systems*, *56*, 1–25. https://doi.org/10.17705/1cais.05601

Leaton Gray, S., Edsall, D., & Parapadakis, D. (2025). AI-Based Digital Cheating At University, and the Case for New Ethical Pedagogies. *Journal of Academic Ethics*, *23*(4), 2069–2086. https://doi.org/10.1007/s10805-025-09642-y

Long, D., & Magerko, B. (2020). What is AI Literacy? Competencies and Design Considerations. *Conference on Human Factors in Computing Systems - Proceedings*, 1–16. https://doi.org/10.1145/3313831.3376727

Messick, S. (1995). Standards of Validity and the Validity of Standards in Performance Asessment. *Educational Measurement: Issues and Practice*, *14*(4), 5–8. https://doi.org/10.1111/j.1745-3992.1995.tb00881.x

Miller, A. L., & Research, A. for I. (2011). Investigating Social Desirability Bias in Student Self-Report Surveys. *Association for Institutional Research*. http://libdata.lib.ua.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED531729&site=eds-live&scope=site

Mujtaba, B. G. (2025). Human-AI Intersection: Understanding the Ethical Challenges, Opportunities, and Governance Protocols for a Changing Data-Driven Digital World. *Business Ethics and Leadership*, *9*(1), 109–126. https://doi.org/10.61093/bel.9(1).109-126.2025

Mustapha, R., Ibrahim, N., Ayub, M. N., Jaafar, A. B., Jusoh, M. K. A. Bin, & Mahmud, M. (2025). From Detection to Prevention: A Human-Centered Design (HCD) Approach to Mitigating AI Misuse in Learning. *International Journal of Basic and Applied Sciences*, *14*(4), 159–167. https://doi.org/10.14419/3t4ghz09

Nelson, A. S., Santamaría, P. V., Javens, J. S., & Ricaurte, M. (2025). Students' Perceptions of Generative Artificial Intelligence (GenAI) Use in Academic Writing in English as a Foreign Language †. *Education Sciences*, *15*(5). https://doi.org/10.3390/educsci15050611

Ortiz-bonnin, S., & Blahopoulou, J. (2025). ChatGPT usage in higher education students. *Social Psychology of Education*, 1–21.

Patterson, F., Zibarras, L., & Ashworth, V. (2016). Situational judgement tests in medical education and training: Research, theory and practice: AMEE Guide No. 100. *Medical Teacher*, *38*(1), 3–17. https://doi.org/10.3109/0142159X.2015.1072619

Paulhus, D. L. (2018). Encyclopedia of Personality and Individual Differences. *Encyclopedia of Personality and Individual Differences*, 1–5. https://doi.org/10.1007/978-3-319-28099-8

Pudasaini, S., Miralles-Pechuán, L., Lillis, D., & Llorens Salvador, M. (2025). Survey on AI-Generated Plagiarism Detection: The Impact of Large Language Models on Academic Integrity. *Journal of Academic Ethics*, *23*(3), 1137–1170. https://doi.org/10.1007/s10805-024-09576-x

Schiff, D. (2022). Education for AI, not AI for Education: The Role of Education and Ethics in National AI Policy Strategies. *International Journal of Artificial Intelligence in Education*, *32*(3), 527–563. https://doi.org/10.1007/s40593-021-00270-2

Sharma, R. C., & Panja, S. K. (2025). Addressing Academic Dishonesty in Higher Education: A Systematic Review of Generative AI's Impact. *Open Praxis*, *17*(2), 251–269. https://doi.org/10.55982/openpraxis.17.2.820

Shaw, D. (2025). The digital erosion of intellectual integrity: why misuse of generative AI is worse than plagiarism. *AI and Society*, 2–4. https://doi.org/10.1007/s00146-025-02362-2

Theoharakis, V., Mylonopoulos, N., & Papadopoulou, K. (2025). AI's learning paradox: how business student' engagement with AI amplifies moral disengagement-driven misconduct. *Studies in Higher Education*, *5079*, 1–18. https://doi.org/10.1080/03075079.2025.2533365

Țîru, L. G., Gherheș, V., Stoicov, I., & Stanici, M. (2025). Not Ready for AI? Exploring Teachers' Negative Attitudes Toward Artificial Intelligence. *Societies*, *15*(12), 1–19. https://doi.org/10.3390/soc15120337

Unesco. (2024). Readiness Assessment Methodology. *Unesco*. https://www.unesco.org/ethics-ai/en/ram

van de Mortel, T. F. (2008). Faking it: Social desirability response bias in selfreport research. *Australian Journal of Advanced Nursing*, *25*(4), 40–48.

Yan, Y., Liu, H., & Chau, T. (2025). A Systematic Review of AI Ethics in Education: Challenges, Policy Gaps, and Future Directions. *Journal of Global Information Management*, *33*(1), 1–50. https://doi.org/10.4018/JGIM.386381

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE.

Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods designs: Principles and practices. *Health Services Research, 48*(6), 2134–2156.

Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement, 50*(1), 1–73.

International Journal of Modern Education
IJMOE
EISSN: 2637-0905
Volume: 8 Issue: 29 (March 2026) PP. 945-960

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*(9), 741–749.

Patterson, F., Zibarras, L., & Ashworth, V. (2016). Situational judgement tests in medical education and training: Research, theory and practice. *AMEE Guide No. 100*. *Medical Teacher, 38*(1), 3–17.

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*(5), 879–903.