Using AI Effectively and Ethically in Academic Work: Student and Staff Attitudes, Policy Clarity and Academic Integrity at Myanmar Imperial University

Dr Selvakumar Ramachandran, FBCS, FRSA
University of Oxford
Oxford, United Kingdom
selvakumar.ramachandran@conted.ox.ac.uk

Dr Poht Poht Kyi
Myanmar Imperial University
Yangon, Myanmar
pohtpohtkyi@miu.org.mm

***Abstract*—Generative artificial intelligence (AI) has moved from novelty to expectation in academic life faster than university policy has adapted. This paper reports a pre-workshop survey of 85 students and staff at Myanmar Imperial University (MIU), Yangon, collected on 8 October 2026 before a session on using AI effectively and ethically. ChatGPT (90.6%) and Google Gemini (69.4%) dominated use, mainly for brainstorming, summarising and explaining concepts. Judgements diverged most for AI-drafted prose: 20.0% rated a full AI-written paragraph "always fine" and 38.8% "never acceptable." Although 69.4% rated MIU's AI policy as clear, perceived clarity was unrelated to having felt unsure whether an AI use would count as misconduct (Fisher's exact p = 1.00); 82.4% reported being, or possibly having been, unsure. Confidence in knowing the acceptable-use line correlated with comfort explaining one's AI use aloud (Spearman's ρ = 0.28, p = .0099) but not with worry about breaking an unknown rule (ρ = 0.05, p = .65), suggesting that policy ambiguity carries an anxiety cost independent of confidence. We place these results in a review of 49 verified open-access sources on academic integrity, detector reliability, tiered "traffic-light" policies and equity of access, and close with tiered policy recommendations for MIU and similar institutions in under-resourced, non-English-first-language settings.**

***Index Terms*—academic integrity, AI literacy, assessment design, generative AI, higher education policy, Myanmar, traffic-light framework**

I. INTRODUCTION

Generative AI tools have reached everyday student use faster than any earlier academic technology. Word processors, the internet and email were each, in their own time, optional novelties before they became simply how work gets done. Large language models such as ChatGPT have compressed that same adoption curve into a matter of months rather than years [1], [2]. Universities everywhere are still writing, and in many cases rewriting, the rules for how students and staff may use these tools. The evidence on how students actually behave, and how they feel about the uncertainty they face, remains limited for most institutions. It is especially limited outside North America, the UK, and a small number of well-resourced university systems in Asia and Europe.

This paper addresses that gap for one institution: Myanmar Imperial University (MIU), a private university in Yangon that offers undergraduate and MBA/Master's programmes mainly in Business, Management and Marketing. On 8 October 2026, the first author, an academic at the University of Oxford, Director of Kerckhoffs Ltd and Professor at MIU, delivered a workshop to MIU students and staff titled "Using AI Effectively and Ethically in Academic Work." The workshop combined several elements. It used historical framing based on the medieval lectio and disputatio tradition, in which a student proved their understanding by defending it aloud, not simply by writing it down. It introduced a four-tier "AI Use Scale" (Fig. 1) running from "Always Fine" through "Say You Used It" and "Ask First" to "Never." It discussed real cases of AI failure, including the fabricated legal citations sanctioned in Mata v. Avianca [3] and the chatbot-invented refund policy upheld against Air Canada in Moffatt v. Air Canada [4]. It also gave practical guidance on using AI as, in the workshop's words, "a helper for learning, not a writer who does your work."

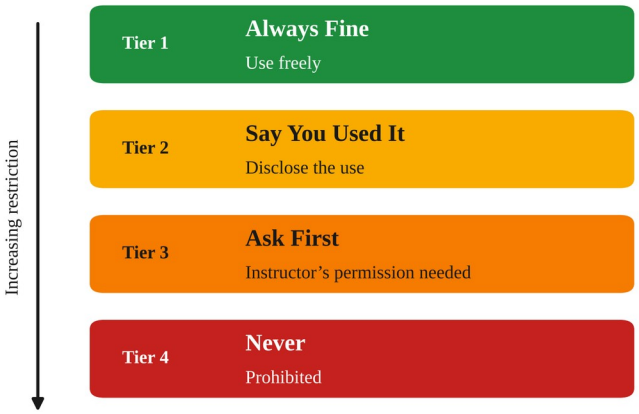


Fig. 1. The four-tier AI Use Scale taught in the workshop, from least to most restrictive.

Immediately before the workshop, attendees completed a short, voluntary survey distributed through Jotform. The survey covered their AI tool use, their sense of institutional policy clarity, their confidence in telling acceptable from unacceptable use apart, and their comfort explaining their own AI use aloud. This last question was a direct echo of the disputatio principle the workshop would go on to teach. Eighty-six submissions were received. After removing one exact duplicate, 85 individual responses make up the sample analysed in this paper.

This paper has four aims. The first is to summarise the open-access academic and institutional-policy literature on generative AI in higher education. This covers academic integrity and misconduct, the reliability of AI-detection tools, tiered or "traffic light" institutional frameworks for permitted AI use, and the pedagogical and equity dimensions of AI

adoption. This literature provides context for interpreting the MIU data. The second aim is to describe the survey methodology, sample and analytic approach transparently. The third aim is to report descriptive and inferential findings on MIU respondents' AI use patterns, acceptability judgements, policy-clarity perceptions, and the statistical relationships between confidence, comfort and worry. The fourth aim is to turn these findings, read alongside the literature, into concrete policy recommendations for MIU and for similar institutions operating in non-English-first-language, resource-constrained settings.

A note on scope. This is a practitioner-oriented study rather than a definitive research article. It presents a single-institution, cross-sectional, self-selected sample collected around a single workshop, so its statistical findings should be read as suggestive rather than confirmatory. Where the sample size and design allow a meaningful inferential test, that test is reported, together with its limitations. Where they do not, the paper says so rather than overstating what 85 responses from one event can show. Section VII (Limitations) returns to this point in full.

II. SYSTEMATIC LITERATURE REVIEW

A. Review Question and Search Strategy

This review asked four questions. First, how does the open-access academic and institutional-policy literature describe the academic-integrity risks created by generative AI, and how reliable are the tools built to detect it? Second, how do tiered or "traffic light" frameworks classify permitted levels of AI use? Third, how is institutional AI policy designed, and how does it relate to student uncertainty? Fourth, what are the pedagogical, equity and assessment-design implications of AI adoption in higher education?

This review did not have access to subscription databases such as Scopus or Web of Science. The search instead drew on Google Scholar results, arXiv, SSRN, ERIC, DOAJ, institutional repositories, EDUCAUSE, and publicly available university policy pages. Search terms included combinations such as "generative AI academic integrity higher education," "AI detection tool accuracy study," "AI hallucination citation fabrication," "traffic light AI use policy university," "RAG framework generative AI assessment," "tiered AI use framework higher education," "student perceptions generative AI survey," "AI literacy framework higher education," "digital divide generative AI access," and "oral assessment viva generative AI." Fifty-five separate search queries were run across three parallel search threads, producing roughly 500 screened result snippets. Each candidate source was checked individually before it was included: the author names, publication year, title and venue were confirmed against the source's own page, DOI record, or repository entry. Sources that could not be verified this way were excluded. This included anonymous pages, sources with attributions that could not be confirmed, and marketing pages with no named author. Forty-nine sources passed this verification step. They are summarised below and listed in full in the References section.

This review is structured and transparent, but it is not exhaustive in the way a database search would be. It does not claim the completeness of a registered systematic review run against subscription indices. Institutions wishing to cite it in a formal academic submission should treat it as a strong starting bibliography rather than a final one.

B. Academic Integrity, Misconduct and the Authorship Question

The arrival of fast, fluent and low-cost text generation reopened long-settled questions about what counts as a student's own work. Cotton et al. [5] describe the core tension directly. Generative AI offers real benefits for engagement and access, but it also increases the risk of plagiarism and fabricated analysis. The authors argue that an effective institutional response needs combined policy, training and multiple detection methods, not any single fix. That institutional response has in practice been slow. Perkins and Roe [6] analysed 142 institutional academic-integrity policies using corpus linguistics and found that none had been meaningfully revised to address generative AI, despite how quickly it had spread. They propose adding explicit "technological explicitness" to the standard academic-integrity framework, so that policies name AI directly rather than relying on older, vaguer plagiarism language. The Committee on Publication Ethics [7] has taken the clearest institutional position available: AI tools cannot be listed as authors on scholarly work, because they cannot take responsibility for a publication's content or manage conflicts of interest, and any AI use must be disclosed in a paper's methods. The International Center for Academic Integrity's [8] long-standing six-value framework of honesty, trust, fairness, respect, responsibility and courage remains the conceptual baseline that most AI-era integrity scholarship, including Perkins and Roe's work, builds on and extends.

Large-scale student-opinion data support this picture of policy lagging behind practice. Parsons and Johnston [9] surveyed student perspectives on generative AI at the University of Liverpool, and the results directly informed that institution's integrity-code revisions. This offers a close methodological parallel to the present study: a single-institution survey feeding directly into policy. The consistent finding across this literature is not that students are indifferent to integrity. It is that the rules they are meant to follow were often written for a world before AI existed.

C. Can AI Use Be Reliably Detected, and Should Policy Assume It Can?

A second body of evidence complicates any institutional strategy that relies on detection as the main integrity mechanism. Weber-Wulff et al. [10] tested 14 AI-text detectors, including Turnitin's, and found them "neither accurate nor reliable." The detectors were systematically biased toward labelling text as human-written, and they were easily defeated by light paraphrasing. Ibrahim et al. [11] tested detection across 32 real university courses and found that OpenAI's own classifier missed 49% of actual ChatGPT submissions, while GPTZero produced an 18% false-positive rate on genuinely human-written work. A simple paraphrase attack pushed false-negative rates to between 95% and 98%. OpenAI [12] withdrew its own public AI-text classifier after acknowledging a "low rate of accuracy": it correctly flagged only 26% of AI-written text, while incorrectly flagging 9% of human-written text as AI-generated. Liang et al. [13] add an equity concern that is directly relevant to MIU. GPT detectors are systematically

biased against non-native English writers, flagging their human-written text as AI-generated at higher rates. This is a serious concern for an institution whose students write academically in a second or third language. Sadasivan et al. [14] offer a theoretical explanation for why this problem will not simply be solved by better engineering. Detector accuracy is bounded by the statistical distance between human and AI text, and this distance shrinks as language models improve. This means reliable detection becomes structurally harder, not easier, over time. Watermarking [15] is the leading technical alternative. It embeds a statistically detectable bias in a model's own word choices at the point of generation, but it depends on cooperation from the model provider and remains vulnerable to simple editing.

The practical implication, which is echoed directly in the MIU workshop material, is that policy built mainly around catching AI use through detection software carries two risks. It risks falsely accusing students who did genuine work, and it risks giving false reassurance to students who evade weak detectors. This point is taken up again in Section II-D.

D. AI Hallucination and Confidently Wrong Output

A closely related risk is that generative AI produces fluent, confident and sometimes entirely fabricated content. The field calls this "hallucination" [16]. The consequences are not hypothetical. In Mata v. Avianca [3], two New York attorneys submitted a ChatGPT-drafted legal brief that cited court cases which did not exist, with fabricated quotations attributed to real-sounding but nonexistent precedent. The presiding judge sanctioned the attorneys for failing to verify their own filing. Dahl et al. [17] measured how common this risk is in a professional setting. Across several large language models, hallucination rates on verifiable legal questions ran from 69% to 88%, and were worse on lower-court material. The models were also poorly calibrated: confident even when wrong. The same pattern reached the corporate world in Moffatt v. Air Canada [4], where a customer-service chatbot invented a bereavement-fare refund policy that did not exist. The Civil Resolution Tribunal of British Columbia held the airline bound by its chatbot's fabricated promise regardless. Gaur [18] sets out a similar taxonomy of failure in a scholarly context, including fake citations and misrepresented findings, and recommends retrieval-grounding and citation verification as safeguards. These safeguards matter because confident fabrication, not obvious error, is the dangerous case. A student or professional who does not independently check AI output has no reliable way of knowing when it has gone wrong.

E. Tiered and "Traffic Light" Frameworks for Permitted AI Use

The workshop's own four-tier "AI Use Scale" (Always Fine, then Say You Used It, then Ask First, then Never) belongs to a fast-growing institutional practice of tiered, colour-coded AI-use frameworks. In effect, this is a red-amber-green (RAG) risk model applied to coursework. It is not a one-off teaching device; it reflects converging institutional practice worldwide. The University of Leeds's published staff guidance [19] defines Red (no AI, because it would undermine the assessed learning outcome), Amber (specified, limited use) and Green (AI integral and expected) categories for summative assessment, with provision for reasonable adjustments under the Equality Act. The University of Birmingham's School of English, Drama and Creative Studies documents [20] a similar default three-tier Red/Amber/Green system, applied across its modules from the 2024/25 academic year, with course leaders able to override the default for individual assessments. The College of Staten Island (CUNY) [21] has published an institutional policy built explicitly around a red/yellow/green "Traffic Light Model," distinguishing prohibited use, restricted use with citation, and permitted use with citation. Leiden Law School's 2026 staff guidance [22] uses a four-level model (Prohibited, Text editing, Brainstorming-structure, Full task-based use) that is structurally identical to MIU's own scale. The University of Manchester adopted a Senate-approved four-tier scheme (Prohibited, Minimal, Permitted, Required) that classifies AI use at the level of each individual assessment [1]. Barnard and Cupido [23] report on a South African institution's implementation of a Traffic Light Model and find the framework useful in practice, but dependent on lecturers applying it consistently from one assessment to the next. This is a direct caution for how MIU would need to put its own scale into practice across modules and instructors. Beale [2] places this convergence within a broader argument: layered, clearly defined-use policies, combined with assessment redesign, are currently the most defensible institutional response available, given the detection and hallucination evidence reviewed above.

This pattern is itself worth noting. Institutions on at least four continents have independently arrived at a three- or four-tier, colour-coded AI-use framework. This consistency suggests the format answers a genuine, widely shared institutional need. It also gives MIU a documented precedent to build on, rather than a new experiment to design from scratch.

F. Policy Clarity, Student Anxiety and Self-Reported AI Use

A second strand of this literature looks not just at what policies say, but at how clearly students perceive them, and at what that perceived clarity does to behaviour and anxiety. Escobar et al. [24] interviewed stakeholders about Stanford's AI governance and found that relying on course-level instructor discretion, without an institution-wide code, produced inconsistent enforcement and student confusion. The authors recommend a formal institutional AI code of practice. Curnock Cook and Brennan [25] summarise a UK multi-institution roundtable involving LSE, Maynooth, Exeter and King's College London. They report that unclear AI rules measurably raise student anxiety about accidental misconduct, and that students worry less about getting it wrong when rules are stated clearly and consistently. Tsao [26] interviewed 58 students and 12 teachers at the University of Hong Kong and found that vague institutional AI policy produced inconsistent enforcement and, in several reported cases, wrongful misconduct accusations. The study calls explicitly for clearer policy communication, framed as a protection for honest students rather than only a disciplinary tool against dishonest ones. Hingle and Johri [27] surveyed 61 students on how they understand AI and algorithmic systems, and found persistent confusion even among technically engaged students. They argue for earlier, more explicit AI-literacy instruction. This finding complicates the assumption that a clear policy on its own, without accompanying AI-literacy teaching, is enough.

Several recent multi-institution student surveys offer direct comparisons to the present study, both in method and in substance. Fawns et al. [28] ran 20 focus groups across four Australian universities, involving 79 students, and identified five recurring tensions that students themselves name: efficiency against integrity, use against unclear policy, empowerment against dependency, access against equity, and engagement against unreliable output. This framework maps closely onto the MIU open-text responses reported in Section IV-G. Pramjeeth [29] surveyed 1,866 students across five South African private higher-education institutions and found effects of gender and study level on AI-reliability perceptions, along with inconsistent disclosure behaviour. Morari et al. [30] surveyed 608 students at an Irish university and found persistent gaps in policy awareness, continuing uncertainty about what counts as acceptable use even where policy existed, and a stated student preference for reporting concerns to the institution rather than to peers. Two studies offer closer comparisons to MIU's own Southeast Asian, non-English-first-language context. Mota et al. [31] conducted a qualitative study of 35 final-year students at a university in Lesotho and found genuine learning benefit alongside strong calls for a university-wide AI policy, given currently unclear boundaries. Xie et al. [32] interviewed Chinese undergraduates and found that AI use was shaped as much by access, language and domestic-platform regulation as by institutional policy, with students adopting workarounds such as VPN use that carry their own ethical and privacy complications.

G. Pedagogy, Critical Thinking and AI as a Learning Tool Rather Than a Writing Tool

A fourth strand of literature addresses the workshop's central teaching claim: that AI should function as "a helper for learning, not a writer who does your work." EDUCAUSE's AI-literacy frameworks [33], [34] and Zhou and Schofield's [35] conceptual framework for AI literacy in learning development converge on a similar structure. Each identifies technical understanding, evaluative judgement, practical application and ethics as the key components, and each treats evaluative judgement, meaning the ability to assess AI output critically rather than simply accept it, as the core skill, rather than fluency in prompting. Gerlich [36] found empirically that heavier reliance on AI tools correlates with increased "cognitive offloading" and reduced critical-thinking engagement, particularly among younger and more AI-dependent users. A 12-study systematic review by Martahani et al. [37] found that the effect of generative AI on metacognitive learning is conditional. Structured scaffolding and deliberate prompting support self-regulated learning, while unstructured and unrestricted use produces what the authors call "metacognitive laziness." Degen [38] puts the alternative into practice directly, designing a Socratic-questioning AI tutor meant to help students develop their own research questions rather than supplying answers. This is a modern technological echo of the disputatio tradition the MIU workshop invokes historically. However, Barshay's [39] account of a multi-school study of the Khanmigo AI tutor found that most students actively avoided the tool's Socratic hints and asked for direct answers instead, and that measured learning gains traced back to the underlying practice material rather than to the AI tutor itself. This is a useful caution: "AI as Socratic helper" is a design intention, not an automatic outcome. Students will often take the path of least cognitive resistance unless assessment design actively works against that pull.

Equity of access is a persistent concern throughout this strand of literature. Warschauer's [40] foundational digital-divide scholarship argued, well before generative AI existed, that access to technology alone does not produce equitable outcomes without accompanying social and institutional support. Newer literature repeatedly confirms this finding in the AI context. Rahman and Saunders [41] report focus-group evidence that paid AI subscriptions give some students a material academic advantage. One participant in their study described the risk that "academic help may be determined by financial resources rather than academic competence." Doyle and Wilson [42] argue that the digital divide has shifted from device and internet access toward AI-tool affordability and capability specifically, and they recommend institution-negotiated AI access as an equity measure. Cheng's [43] study across two Chinese universities similarly found that AI-related inequality arises from the combination of tools, infrastructure, institutional norms and peer support, not from access alone. The study concludes that such inequalities are "produced through everyday participation conditions." This is a useful caution for any MIU response that assumes providing institutional AI licences alone would resolve the access concern.

H. Assessment Redesign and the Return of Oral and Verbal Evaluation

The final strand of literature speaks directly to the workshop's lectio/disputatio theme: the idea that a thesis defence today and a job interview tomorrow serve a similar function. Storey [44] argues explicitly for reclaiming the oral thesis defence as the authoritative check on genuine understanding, precisely because AI can now generate competent written text. Brumby et al. [45], writing in Times Higher Education, propose "sampled viva moderation," meaning an oral re-examination of roughly 10% of students on their written submissions. They present this as a scalable, more inclusive alternative to either a full oral examination for everyone or unreliable detection software, and frame it explicitly as assessing understanding rather than policing AI use. Church et al. [46] describe an AI-assisted viva-voce simulator, built partly at Oxford, which questions students on their own submitted work while a human examiner keeps the final judgement. This is a direct technological bridge between the oral tradition and AI-era tools, and the authors flag equity and access questions in how it would be deployed. Su et al. [47] and Ransome [48] both set out broader frameworks for redesigning assessment around the presence of generative AI, rather than against it, using process-based, iterative and oral components in place of purely written output. Lehigh University's employer-partner survey [49] closes the loop the workshop draws between academic and career expectations. Employers increasingly expect graduates to demonstrate AI literacy and to be able to explain and justify their own use of AI. This is the same "defend your understanding aloud" standard that is driving academic assessment redesign, now appearing as an employer-side expectation too.

I. Synthesis

Taken together, this literature points to four practical conclusions that are directly relevant to MIU. First, institutional AI policy at most universities has lagged behind the technology's spread, and vague policy measurably increases student anxiety as well as the risk of both wrongful and missed misconduct findings. Clarity is therefore protective, not merely administrative. Second, AI-detection software is not a reliable enforcement mechanism on its own, and it is specifically biased against non-native English writers, a group that makes up a large share of MIU's student body. Policy should not rely on detection as a primary integrity safeguard. Third, tiered, colour-coded "traffic light" AI-use frameworks, structurally identical to the one the MIU workshop already teaches, represent converging international practice across several continents and institution types. This gives MIU a validated template to build on, rather than a new design problem to solve from scratch. Fourth, assessment redesign that reintroduces oral, process-based or "explain it yourself" components is both a sound teaching response to AI's writing ability and an increasingly explicit employer expectation. This directly supports the workshop's own central idea.

III. METHODOLOGY

A. Design and Procedure

This study used a cross-sectional, self-administered online survey, built and distributed on the Jotform platform. Attendees completed it immediately before the first author's workshop at Myanmar Imperial University on 8 October 2026. The workshop slides describe the survey to participants as short, taking about five minutes, voluntary, and intended to inform both the live session and future AI-policy development at the university. A second, brief follow-up survey was planned for the end of the session, but it was not available for this analysis (see Section VII).

B. Instrument

The instrument began with two demographic questions: year of study or role, and field of study. It then asked 17 substantive questions. These covered: how often the respondent used AI tools for coursework; which named AI tools they had used, selected from a list including ChatGPT, Google Gemini, Microsoft Copilot, Claude and Grammarly's AI features, with an open "other" option; what they had used AI for, selected from a list including brainstorming, summarising, explaining concepts, checking grammar, analysing data, translating text, exam preparation, writing paragraphs or essays, and writing or checking code; how clear they found their institutional AI policy; their self-rated confidence, on a 1 to 5 scale, in knowing the line between acceptable and unacceptable use; whether they had ever been unsure if a particular AI use would count as misconduct; their self-rated comfort, on a 1 to 5 scale, explaining their own recent AI use aloud; their acceptability judgement, on a four-point scale of "Always fine," "Fine if I disclose it," "Only with the instructor's permission" and "Never acceptable," for four specific use-scenarios that correspond to the workshop's AI Use Scale tiers; their worry, on a 1 to 5 scale, about being wrongly accused of AI use; their worry, on a 1 to 5 scale, about accidentally breaking an unknown rule; their view on whether all students have equal access to good AI tools; and a single open-text question asking what they most wanted to understand better about using AI in their studies.

C. Participants and Data Preparation

Eighty-six Jotform submissions were received. One row was an exact duplicate: it had the identical respondent identifier and identical responses across every substantive item, so it was removed, leaving an analytic sample of 85. Respondents were classified by role as Student, Staff/Faculty, or Unknown/Not specified. This classification was based on their stated year-of-study or role field, and, where that field was blank but the respondent's title indicated a role, on corroborating information. This produced 75 students (88.2%), 3 staff or faculty (3.5%), and 7 respondents whose role could not be determined from the data provided (8.2%). Respondents' reported field of study was Business (57, 67.1%), Management (24, 28.2%), or Marketing (4, 4.7%). Among students, year-of-study groups ranged from second-year undergraduates to final-year undergraduates, the largest single group at 33 respondents, to Master's or MBA students, 17 respondents. Each respondent is referred to in this paper only by an arbitrary sequential identifier (P001 to P085); no respondent is named or otherwise identifiable. The de-identified dataset underlying this paper is available from the authors on request.

D. Analytic Approach

Descriptive statistics, including frequencies, percentages, means, medians and standard deviations, summarise all survey items. For the four acceptability-scenario items, responses were mapped onto the workshop's own four-tier labels of Always Fine, Say You Used It, Ask First and Never. This allows direct comparison with the AI Use Scale framework taught in the session. For items where respondents could select more than one option, such as tools used and uses of AI, each selected option was counted on its own, so percentages do not add up to 100%. Open-text responses were coded inductively by a single coder using thematic analysis [50]; a response could carry more than one theme.

Inferential tests examined six relationships chosen in advance for their practical relevance to MIU policy. These were: whether self-rated policy clarity relates to ever having been unsure if an AI use counted as misconduct; whether role, comparing students with staff or faculty, relates to confidence; whether field of study relates to frequency of AI use; whether role relates to perceived equality of access; the correlation between confidence and comfort explaining AI use aloud; and the correlations among the worry measures and between confidence and worry about an unknown rule. Two-by-two and other small contingency tables were tested with Fisher's exact test. Larger tables used Pearson's chi-square test, with Monte Carlo permutation p-values where expected cell counts were small. Associations between ordinal Likert items were tested with Spearman's rank correlation. All tests used a conventional significance threshold of alpha = .05. Given the sample size and the number of comparisons, the results should be read as exploratory signals rather than confirmatory hypothesis tests. No correction for multiple comparisons was applied, which is a limitation discussed in Section VII. All analysis was carried out in Python, using pandas and SciPy.

IV. FINDINGS

A. Sample Composition

Of the 85 respondents, 75 (88.2%) identified as students, 3 (3.5%) as staff or faculty, and 7 (8.2%) did not give enough information to classify. Among students, final-year undergraduates were the largest single group, with 33 respondents, followed by Master's or MBA students (17), and second- and third-year undergraduates (13 and 9 respectively), with the remaining three describing themselves as, for example, "Bachelor" or "Level 7". Respondents' field of study was weighted heavily toward Business (57, 67.1%) and Management (24, 28.2%), with a small Marketing group (4, 4.7%). This is consistent with MIU's stated academic focus (Fig. 2).

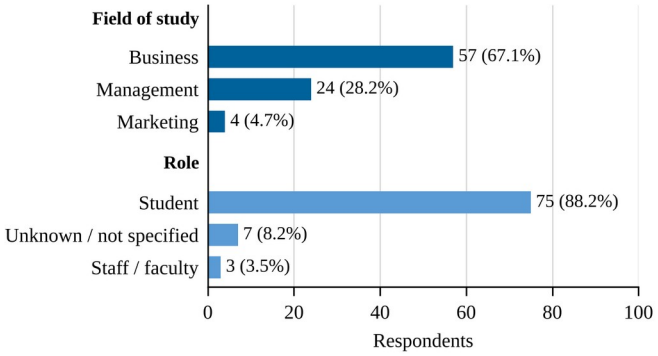


Fig. 2. Sample composition by field of study and role (n = 85).

Self-reported AI-use frequency was high across the sample. Thirty-two respondents (37.6%) reported using AI tools every day for coursework, and 40 (47.1%) reported using them a few times a week. Together, 84.7% of respondents use AI at least weekly (Fig. 3). Only 12 (14.1%) reported monthly use, and a single respondent (1.2%) reported rarely using AI. No respondent reported never using AI.

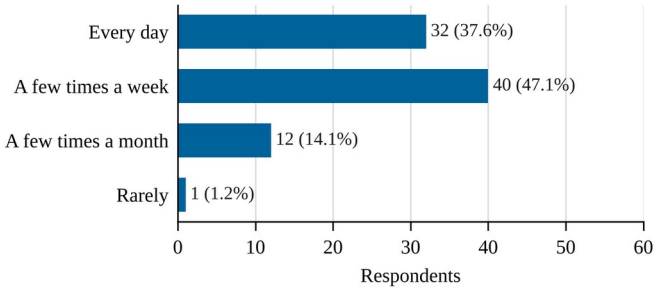


Fig. 3. Frequency of AI tool use for coursework (n = 85).

B. Which Tools, and for What

ChatGPT was the dominant tool, used by 77 respondents (90.6%), followed by Google Gemini, used by 59 respondents (69.4%). Grammarly's AI features (16, 18.8%), Claude (13, 15.3%) and Microsoft Copilot (12, 14.1%) were well behind these two leading tools. DeepSeek, NotebookLM and Quillbot were each used by only one or two respondents (Fig. 4).

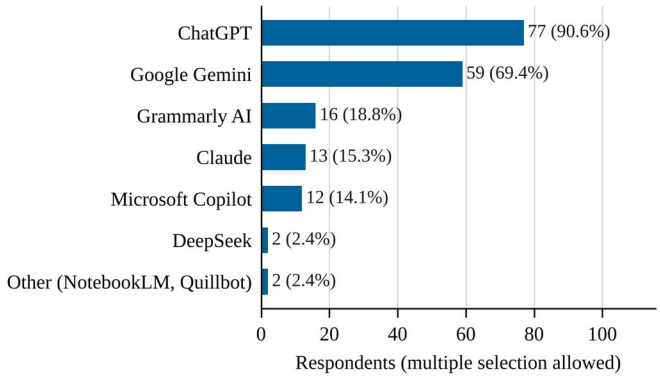


Fig. 4. AI tools used for academic work (multiple selection allowed; percentages of n = 85 do not sum to 100%).

The most common uses (Fig. 5) were brainstorming ideas (61, 71.8%), summarising readings (55, 64.7%) and having a concept explained (53, 62.4%). These all sit within the "helper for learning" category the workshop recommends. Using AI to analyse data (41, 48.2%), check grammar (30, 35.3%) and translate text (29, 34.1%) were also common. Notably, 21 respondents (24.7%) reported using AI to write paragraphs or essays for them. This is the use-case the workshop's framework treats as highest-risk, so while it is a minority practice, it is far from a negligible one.

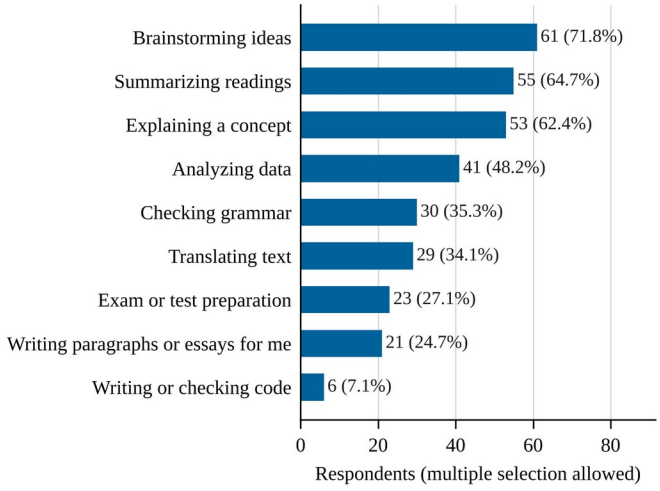


Fig. 5. Reported uses of AI in academic work (multiple selection allowed; percentages of n = 85 do not sum to 100%).

C. The AI Use Scale in Practice: Where Agreement Breaks Down

Mapping the four acceptability-scenario items onto the workshop's own tier labels shows where MIU respondents agree and where they do not (Fig. 6 and Table I). For low-stakes uses, namely checking work for errors and brainstorming, the largest share of respondents judged the behaviour "Always Fine" (53.6% and 41.7% respectively), and very few respondents (2.4% in both cases) called it never acceptable. Summarising articles and then writing the analysis independently split more evenly between "Always Fine" (41.2%) and "Say You Used It" (35.3%). This suggests that most respondents see disclosure, rather than prohibition, as the right norm for this use.

The picture changes sharply for AI-drafted prose submitted with only minor edits. Just 20.0% judged this "Always Fine," a further 21.2% chose "Say You Used It," 20.0% chose "Ask First," and the largest single group, 38.8%, judged it "Never" acceptable. This is the only one of the four scenarios where "Never" is the most common response. It is also the only scenario where opinion is spread close to evenly across all four tiers, rather than concentrated at one end. This is empirical confirmation that this scenario is exactly the contested boundary case the workshop's framework is designed to clarify.

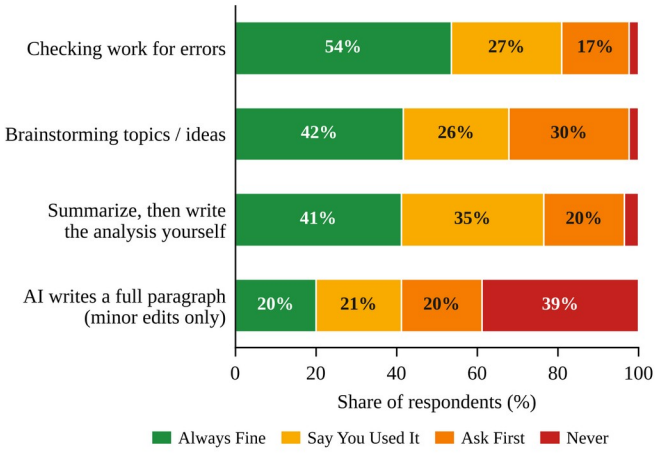


Fig. 6. Acceptability judgements across four AI-use scenarios, mapped to the AI Use Scale tiers. Percentages are of respondents answering each item (n = 84 for the first two scenarios, 85 for the others).

TABLE I
AI USE SCALE RESPONSES BY SCENARIO (% OF RESPONDENTS)

AI-use scenario	n	Always Fine	Say You Used It	Ask First	Never
Checking work for errors	84	53.6	27.4	16.7	2.4
Brainstorming topics/ideas	84	41.7	26.2	29.8	2.4
Summarise, then write analysis yourself	85	41.2	35.3	20.0	3.5
AI writes a full paragraph (minor edits only)	85	20.0	21.2	20.0	38.8

D. Policy Clarity, Uncertainty and Confidence

Most respondents rated their institution's AI policy as clear. Fifty-nine (69.4%) said "Yes, it's clear," 21 (24.7%) said "Somewhat clear," and only 5 (5.9%) reported no clear policy or were unsure whether one existed. Despite this, 38 respondents (44.7%) said they had been unsure, at some point, whether a specific use of AI would count as academic misconduct, and a further 32 (37.6%) answered "Not sure" to the same question. This means that 82.4% of respondents reported either having actually felt uncertain, or that it was plausible they had, against only 17.6% who said they had never felt unsure.

These two findings do not line up statistically, which is a notable result. Comparing respondents who rated policy as "clear" against those who rated it as not fully clear, the share who had ever been unsure about misconduct was nearly identical: 81.4% of the "clear" group, against 84.6% of the "not fully clear" group, had ever felt unsure (Fisher's exact test, p = 1.00, showing no detectable association). *In other words, self-rated policy clarity did not predict freedom from the felt uncertainty the policy is meant to resolve.* This matches the findings of Tsao [26] and Perkins and Roe [6] that institutional policy often lags behind the ambiguous cases it needs to cover. It suggests that a policy can read as clear in general while still leaving genuine uncertainty in edge cases, which is exactly the kind of boundary that the AI Use Scale's middle tiers, "Say You Used It" and "Ask First," are designed to handle.

Confidence in knowing the acceptable-use line averaged 3.65 out of 5 (standard deviation = 0.67), and comfort explaining one's AI use aloud, right now, averaged 3.38 (standard deviation = 0.77). Both are moderately positive, but well short of universal assurance. Worry about being wrongly accused of AI use averaged 3.47 (standard deviation = 1.08), and worry about accidentally breaking an unknown rule averaged 3.68 (standard deviation = 0.97), the highest mean of the four Likert measures collected (Fig. 7).

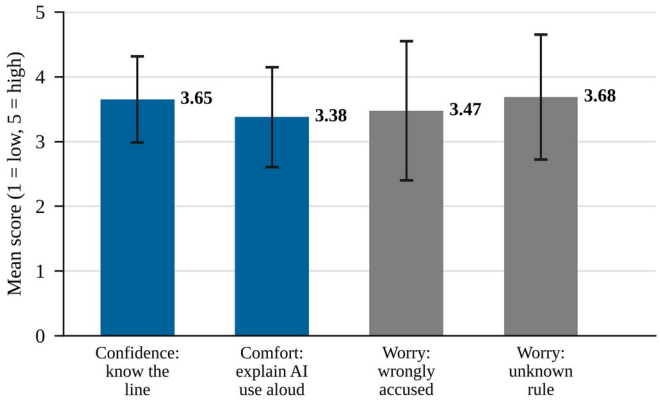


Fig. 7. Mean self-reported confidence and worry scores, 1 (low) to 5 (high), with standard-deviation whiskers (n = 85).

E. How Confidence, Comfort and Worry Relate to One Another

The full pattern of correlations among the four confidence and worry measures is shown in Fig. 8.

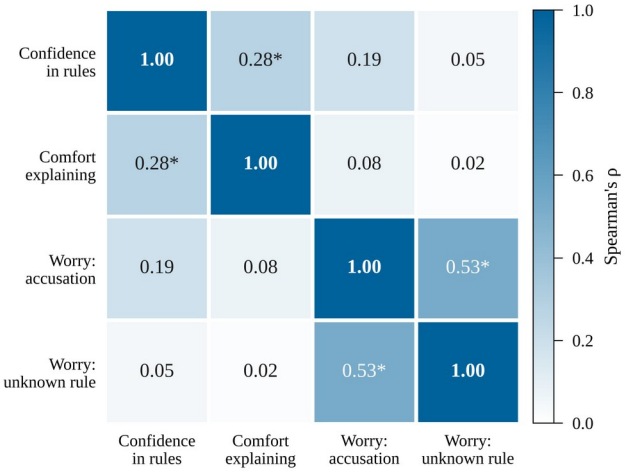


Fig. 8. Spearman correlations among confidence and worry measures (n = 85). An asterisk marks a correlation significant at p < .05; the diagonal is 1.00 by definition.

Confidence in knowing the acceptable-use line correlated significantly, though moderately, with comfort explaining one's AI use aloud (Spearman's rho = 0.28, p = .0099, n = 85). Respondents who felt surer of the boundary were also more comfortable defending their own recent AI use in conversation. *This is the clearest quantitative support in this dataset for the workshop's central idea: that the ability to explain one's AI use out loud is a meaningful, not an arbitrary, sign of genuine understanding of where the line sits.*

The two worry measures correlated strongly with each other (rho = 0.53, p \< .0001). Respondents who worried about being wrongly accused also tended to worry about unknowingly breaking an unstated rule. This is consistent with a single underlying "ambiguity anxiety," rather than two separate concerns. Confidence also showed a weaker, borderline correlation with worry about false accusation (rho = 0.19, p = .085), while comfort explaining aloud showed no meaningful relationship with either worry measure (rho = 0.08 and rho = 0.02 respectively, neither significant).

By contrast, confidence in knowing the acceptable-use line showed no significant relationship with worry about breaking an unknown rule (rho = 0.05, p = .65). *In plain terms, feeling sure about the rules you do know does not reduce fear of a rule you do not know exists.* This null result is informative in its own right. It shows that confidence and anxiety are not simply two ends of the same scale among MIU respondents, and that reassurance-focused messaging alone, such as stating that "the policy is clear," is unlikely to resolve the specific fear of an unstated rule. Only direct, explicit and regularly updated guidance is likely to close that particular gap.

Two further comparisons returned no significant association, and are reported here for completeness (Table II), with their limitations noted. Field of study showed a non-significant trend toward higher frequency of AI use among Management and Marketing respondents compared with Business respondents (permutation chi-square test, p ≈ .11). Given the small Marketing subgroup (n = 4), this should be read as a descriptive pattern rather than a confirmed effect. Role, comparing students with staff or faculty, showed no significant association with either confidence or perceived equality of access. However, the staff/faculty subgroup (n = 3) is far too small to support any reliable conclusion either way. This comparison is reported only so that a future, better-powered survey can test it properly.

TABLE II
INFERENTIAL TESTS NOT SHOWN IN FIG. 8

Comparison	Test	Result	Interpretation
Policy clarity × ever unsure re: misconduct (n = 85)	Fisher's exact	p = 1.00	No association; clarity did not reduce felt uncertainty
Field of study × frequency of use (n = 85)	Permutation χ^2	p ≈ .11	Non-significant; small Marketing subgroup
Role (student/staff) × confidence (n = 78)	χ^2	p = .90	Underpowered (staff n = 3)
Role (student/staff) × equal access (n = 78)	χ^2	p = .46	Underpowered (staff n = 3)

F. *Perceived Equity of Access*

Asked whether all students have equal access to good AI tools, 56 respondents (65.9%) said yes, 13 (15.3%) said no, and 16 (18.8%) were not sure. This means just over a third of respondents either perceive inequity or are unsure whether it exists. Read alongside the findings of Rahman and Saunders [41] and Cheng [43], this figure is more likely an undercount than an overcount. Students who hold premium AI subscriptions themselves may be the least likely to notice a gap that disadvantages peers who do not.

G. *What Respondents Want to Understand Better: Qualitative Themes*

Forty-nine respondents (57.6%) answered the optional open-text question asking what they most wanted to understand better about using AI in their studies. Eight of these entries were non-informative (a full stop, a dash or a bare "Yes") and eight more were too brief or off-topic to classify, leaving 33 codable responses. These were coded inductively into six themes (Fig. 9); a response could carry more than one theme. The largest theme, effective and skilful use (17 respondents), covered requests to learn better prompting, study workflows and productivity techniques. Eleven respondents asked directly for clearer rules and boundaries: what exactly counts as acceptable and, in one case, how AI detection works and "how to avoid being wrongly accused of using AI." This is a direct, real-world echo of the false-positive risk documented by Liang et al. [13] and Weber-Wulff et al. [10]. Six respondents wanted to use AI without becoming over-dependent on it, wishing to keep "producing my own original work" and "not relying on it too much," which mirrors the empowerment-versus-dependency tension that Fawns et al. [28] report. A further six wanted help with research, reliable sources and verifying data, four asked how to apply AI in business or career contexts, and three raised fairness, detection or trust concerns.

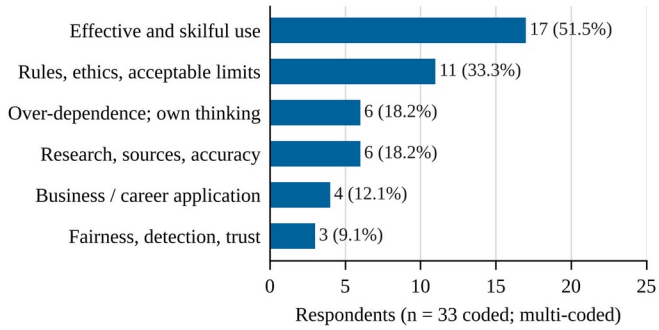


Fig. 9. Themes in respondents' open-text answers (33 codable responses; responses may carry more than one theme).

Two individual responses deserve direct attention. One respondent (P058) added an unprompted methodological caution to their answer: "you cannot expect students to answer this survey honestly. Obviously, if a student used AI to write their assignment, they won't say they did it in this survey... maybe you should consider it in your analysis." This is an important observation, and it is addressed directly in the Limitations section below. A second respondent (P082) raised a fairness concern that the instrument did not otherwise capture: that growing reliance on AI risks devaluing "genuine understanding" relative to "something that looks convincing," making the effort put in by students who read carefully and think independently feel comparatively "invisible." Taken together, these two responses show that MIU's AI-literate students are not simply passive subjects of policy. Several are already thinking, without being asked, about the same integrity and fairness questions this paper investigates.

V. DISCUSSION

Three findings stand out when the MIU data is read against the literature reviewed in Section II.

AI adoption at MIU is already near-universal, but it is concentrated in learning-helper uses rather than writer uses. A total of 84.7% of respondents use AI tools at least weekly, overwhelmingly ChatGPT and Google Gemini. The dominant activities, namely brainstorming, summarising and having concepts explained, sit comfortably within the "helper for learning" approach the workshop recommends, and which the AI-literacy literature [33], [35] treats as sound practice. The AI Use Scale data in Section IV-C shows that respondents draw this same line themselves, independently of the workshop: low-stakes helper uses cluster heavily at "Always Fine," while AI-drafted prose submitted with minor edits is the one scenario where "Never" is the most common judgement. This is an encouraging finding, since it suggests MIU students are not naive about where the serious risk lies. However, the 24.7% who report having used AI to write paragraphs or essays for them, and the 61.2% of respondents who placed that same scenario somewhere other than "Never," show that the contested middle ground is a live practical issue, not a hypothetical one.

Perceived policy clarity is not doing the reassurance work it is meant to do. The near-identical rate of felt uncertainty between respondents who rated MIU's AI policy as clear and those who did not (Section IV-D) is the single most policy-relevant finding in this paper. It is consistent with Perkins and Roe's [6] finding that institutional policies rarely cover AI's genuinely ambiguous cases, even where such policies exist, and with the findings of Tsao [26] and Escobar et al. [24] that vague or inconsistently applied policy leaves anxiety intact even when students can point to a document. Combined with the non-significant correlation between confidence and worry about an unknown rule (rho = .05), this points to two somewhat separate problems. One is a stated-policy problem: does a document exist, and is it generally understood? The other is a felt-ambiguity problem: am I certain that no edge case I have not anticipated will catch me out? MIU's current messaging appears to address the first problem without resolving the second.

The workshop's "explain it aloud" idea has direct empirical support in this dataset. The significant correlation between confidence and comfort explaining AI use aloud (rho = .28, p = .0099) is exactly the relationship predicted by the workshop's disputatio-to-defence framing: students who understand the acceptable-use boundary are more willing to describe their own practice openly. Read alongside Storey [44], Brumby et al. [45] and Su et al. [47] on oral and process-based assessment, this strengthens the case that a verbal-explanation component, whether a full viva, a sampled oral check, or simply an in-class "explain your AI use" exercise, is not only a defensible teaching idea but one with measurable grounding in how MIU's own students already connect confidence with disclosure comfort.

A fourth, smaller point is worth mentioning. The equity findings (Section IV-F) and the Liang et al. [13] detection-bias finding together suggest that MIU should treat equity not as a secondary concern, but as a core part of its AI-integrity response. A third of respondents perceive unequal access to good AI tools, and any policy response that leans on AI-detection software risks disproportionately flagging the genuine work of MIU's non-native-English-speaking students, the exact population Liang et al. found such tools penalise.

VI. RECOMMENDATIONS

These recommendations follow directly from the findings in Section IV, read against the converging institutional practice documented in Section II. They are offered for MIU's consideration and are, in several cases, transferable to other under-resourced, non-English-first-language institutions facing the same gap between rapid AI adoption and unsettled policy.

A. Adopt a Formal, Published Tiered AI-Use Framework

The workshop's four-tier AI Use Scale already functions, informally, as the shared reference point this survey's respondents used, without being told to, when answering the acceptability questions. MIU should formalise it by publishing a short, plain-language policy document in both Burmese and English, modelled on the published RAG and tiered guidance from the University of Leeds and Leiden Law School [19], [22]. Every course syllabus should then state, for each assessment, which tier applies. This follows the University of Manchester's assessment-level classification model [1] and takes on board the caution of Barnard and Cupido [23] that a framework only works if lecturers apply it consistently from one assessment to the next.

B. Close the Felt-Ambiguity Gap Directly, Not Only Through a Policy Document

Because self-rated policy clarity did not reduce felt uncertainty (Section IV-D), MIU should not assume that publishing the tiered framework on its own will resolve the problem. A maintained, searchable list of frequently asked questions covering concrete edge cases, updated each term as new scenarios arise, a named contact point for "ask first" queries with a committed response time, and brief AI-policy refreshers at the start of each course are all likely to do more to close this gap than simply restating that a policy exists.

C. Do Not Lean on AI-Detection Software as a Primary Integrity Mechanism

Given the detection-reliability evidence in Section II-C, and its specific bias against non-native English writers [13], which is directly relevant to MIU's student body, any misconduct process should treat AI-detector output as, at most, a reason to start a conversation with the student. It should never be used as standalone evidence. A disclosure-first model, built around the "Say You Used It" tier, together with process-based evidence such as drafts, prompt logs or an oral explanation, is more defensible than relying on detector scores.

D. Introduce a Verbal-Explanation or Sampled-Viva Component to High-Weight Assessments

Following Brumby et al. [45] and Storey [44], MIU should pilot a short, low-stakes "explain your work" component, even just a five-minute sampled conversation on a subset of submissions, for assessments that fall in the "Ask First" or "Never" tiers. This paper's own finding that confidence correlates with comfort explaining aloud (Section IV-E) suggests such a component would be a valid measure of genuine understanding, not merely a punitive one, consistent with the workshop's own approach.

E. Treat Equitable Access as Part of the Integrity Strategy

With a third of respondents perceiving unequal access to AI tools, MIU should look into negotiating an institution-wide licence for a baseline AI tool, which would reduce the gap between students who can and cannot afford premium tiers. This is consistent with the recommendation of Doyle and Wilson [42]. At the same time, MIU should recognise, following Cheng [43], that licence access alone will not resolve inequality rooted in infrastructure, language or peer support. These issues need parallel attention in AI-literacy teaching.

VII. LIMITATIONS

This paper's findings should be read within several clear boundaries.

- *Single institution, single event.* All data come from one workshop at one university. The findings may not generalise to MIU's wider student body, most of whom did not attend, to other faculties, or to other institutions.
- *Self-selected, convenience sample.* Attendance at an optional AI-ethics workshop is likely to select for students who are already more engaged with, or more anxious about, AI policy than the general student population. This could inflate both reported confidence and reported worry compared with a random sample.
- *Cross-sectional design, with no post-session comparison available.* The workshop slides indicate that a brief follow-up survey was planned for the end of the session, but that data was not available for this analysis. A before-and-after comparison would have let this paper test whether the workshop itself changed confidence, worry or acceptability judgements. This is arguably the single most useful test this design could offer, and its absence is this paper's most significant methodological gap. MIU is encouraged to capture and analyse that follow-up data as a direct companion to this report.
- *Self-report and social-desirability bias.* As one respondent pointed out without being asked (Section IV-G), a student who has used AI in ways they know are unauthorised has little incentive to disclose this on a survey run by their own university immediately before an ethics workshop. Reported rates of higher-risk uses, such as AI-drafted paragraphs, should therefore be treated as a floor rather than a ceiling on actual prevalence. The true rate of higher-risk use is plausibly higher than what is reported here.
- *Small and uneven subgroups.* The staff/faculty subgroup (n = 3) and the Marketing field-of-study subgroup (n = 4) are too small to support reliable subgroup comparisons. The corresponding tests in Section IV-E are reported for transparency and for future reference, not as confirmed findings.
- *No correction for multiple comparisons.* Several inferential tests were run at a conventional significance threshold of .05 with no correction applied, so there is a meaningful chance that at least one result is a false positive by chance alone. The two strongest findings, namely the policy-clarity and uncertainty null result and the confidence and comfort-explaining correlation, are both consistent with independent prior literature (Section II). This increases confidence in them beyond what the p-values alone would justify, but the results should still be read as hypothesis-generating rather than confirmatory.
- *Literature review scope.* The review in Section II draws on openly accessible sources rather than a full subscription-database search. While 49 sources were individually verified, it should be treated as a strong starting bibliography rather than an exhaustive, registered systematic review (Section II-A).
- *Author positionality.* The first author delivered the workshop, and the second author is the founder and Vice-Chancellor of the institution studied. Respondents may have answered with this in mind, and the authors' closeness to the setting may shape interpretation. The results are reported descriptively, with their limitations stated, and independent replication would be valuable.

VIII. CONCLUSION

Myanmar Imperial University's students have already adopted generative AI tools at a rate consistent with global trends. Reassuringly, they largely draw their own line in much the same place as the workshop's AI Use Scale: comfortable with AI as a brainstorming and explanation partner, but divided

and cautious about AI as a ghostwriter. What this survey makes visible, and what a document simply stating that policy is "clear" cannot fix on its own, is a persistent undercurrent of felt uncertainty that does not track stated policy clarity, together with a measurable and worth-protecting relationship between understanding the rules and being willing to explain one's own work aloud.

The international literature reviewed here offers MIU a path forward that is grounded in evidence rather than speculation. This means formalising the tiered framework the workshop already teaches, closing the ambiguity gap with living guidance rather than a static document, treating detection software as a starting point for conversation rather than a verdict, bringing verbal explanation back into assessment design, and treating equity of access as part of the integrity strategy rather than as a separate issue. The closing message to workshop participants, to "use AI to think further, not to think less," is on this evidence not only good advice but one that MIU's own students are already largely trying to follow. The institution's task now is to build the policy infrastructure that makes this easier, not harder, to do with confidence.

REFERENCES

[1] E. Thompson, "University of Manchester adopts four-tier AI policy for student assessments," *EdTech Innovation Hub*, 2026.

[2] R. Beale, "Adapting university policies for generative AI: Opportunities, challenges, and policy solutions in higher education," 2025, arXiv:2506.22231.

[3] *Mata v. Avianca, Inc.*, 678 F. Supp. 3d 443 (S.D.N.Y. 2023).

[4] *Moffatt v. Air Canada*, 2024 BCCRT 149 (Civil Resolution Tribunal of British Columbia).

[5] D. R. E. Cotton, P. A. Cotton, and J. R. Shipway, "Chatting and cheating: Ensuring academic integrity in the era of ChatGPT," *Innov. Educ. Teach. Int.*, vol. 61, no. 2, pp. 228–239, 2024, doi: 10.1080/14703297.2023.2190148.

[6] M. Perkins and J. Roe, "Decoding academic integrity policies: A corpus linguistics investigation of AI and other technological threats," *Higher Educ. Policy*, vol. 37, pp. 633–653, 2024, doi: 10.1057/s41307-023-00323-2.

[7] Committee on Publication Ethics (COPE), "Authorship and AI tools," position statement, 2023. [Online]. Available: https://publicationethics.org/guidance/cope-position/authorship-and-ai-tools

[8] International Center for Academic Integrity (ICAI), "Fundamental values of academic integrity," n.d. [Online]. Available: https://www.academicintegrity.org/fundamental

[9] B. N. Parsons and H. Johnston, "Student perspectives on the use of generative artificial intelligence technologies in higher education [data collection]," Univ. Liverpool, 2023, doi: 10.17638/datacat.liverpool.ac.uk/2297.

[10] D. Weber-Wulff *et al.*, "Testing of detection tools for AI-generated text," *Int. J. Educ. Integrity*, vol. 19, no. 1, Art. no. 26, 2023, doi: 10.1007/s40979-023-00146-z.

[11] H. Ibrahim, F. Liu, R. Asim *et al.*, "Perception, performance, and detectability of conversational artificial intelligence across 32 university courses," 2023, arXiv:2305.13934.

[12] OpenAI, "New AI classifier for indicating AI-written text," 2023. [Online]. Available: https://openai.com/index/new-ai-classifier-for-indicating-ai-written-text

[13] W. Liang, M. Yuksekgonul, Y. Mao, E. Wu, and J. Zou, "GPT detectors are biased against non-native English writers," *Patterns*, vol. 4, no. 7, Art. no. 100779, Jul. 2023.

[14] V. S. Sadasivan, A. Kumar, S. Balasubramanian, W. Wang, and S. Feizi, "Can AI-generated text be reliably detected?" 2023, arXiv:2303.11156.

[15] J. Kirchenbauer, J. Geiping, Y. Wen, J. Katz, I. Miers, and T. Goldstein, "A watermark for large language models," 2023, arXiv:2301.10226.

[16] Z. Ji *et al.*, "Survey of hallucination in natural language generation," *ACM Comput. Surv.*, 2022, arXiv:2202.03629.

[17] M. Dahl, V. Magesh, M. Suzgun, and D. E. Ho, "Large legal fictions: Profiling legal hallucinations in large language models," 2024, arXiv:2401.01301.

[18] M. Gaur, "Hallucinations in scholarly LLMs: A conceptual overview and practical implications," UMBC Faculty Collection, MD-SOAR repository, 2026. [Online]. Available: http://hdl.handle.net/11603/41556

[19] Univ. Leeds, "Using RAG assessment categories," generative AI guidance for staff and students, n.d. [Online]. Available: https://generative-ai.leeds.ac.uk/staff-guidance/ai-for-student-education/rag-assessment-categories/

[20] Univ. Birmingham, School of English, Drama and Creative Studies, "Case study 12: Implementing a framework for the use of generative AI across a diverse school," QAA member case study, n.d.

[21] College of Staten Island, CUNY, "AI use policy—Traffic Light Model," institutional policy document, n.d.

[22] Leiden Law School, Leiden Univ., "Using AI in written assignments: Four levels for lecturers and students," staff announcement, May 6, 2026.

[23] A. Barnard and N. Cupido, "Leveraging artificial intelligence in assessments: Good practices through the Traffic Light Model," *J. Ethics Higher Educ.*, vol. 6, no. 1, pp. 71–95, 2025, doi: 10.26034/fr.jehe.2025.8334.

[24] S. Escobar, S. Lombardo, E. Kim, and J. Al Shalabi, "Governing generative AI at Stanford: Recommendations for the use of artificial intelligence by undergraduates and faculty," *GRACE: Global Rev. AI Community Ethics*, vol. 4, no. 1, 2026.

[25] M. Curnock Cook and B. Brennan, "AI for everyone—but not everything," *HEPI Blog*, Feb. 27, 2025. [Online]. Available: https://www.hepi.ac.uk/2025/02/27/ai-for-everyone-but-not-everything/

[26] J. Tsao, "False positives and real anxiety: What happens when universities can't define AI cheating," *HKU Common Core*, Nov. 24, 2025. [Online]. Available: https://commoncore.hku.hk/2025/11/24/false-positives-and-real-anxiety-what-happens-when-universities-cant-define-ai-cheating/

[27] A. Hingle and A. Johri, "A research study on student conceptions of artificial intelligence," in *Proc. ASEE Southeast Sect. Conf.*, Arlington, VA, USA, 2023, doi: 10.18260/1-2--44979.

[28] T. Fawns *et al.*, "Gen AI and student perspectives of use and ambiguity: A multi-institutional study," in *Proc. ASCILITE*, 2024.

[29] S. Pramjeeth, "Ethics in the age of generative AI: Student perceptions, use and transparency in South African private higher education," *J. Educ.*, no. 103, pp. 42–61, 2026, doi: 10.17159/2520-9868/i103a03.

[30] V. Morari, D. Grimes, and D. Hawe, "Academic integrity and generative artificial intelligence: Views and perceptions of students in an Irish university," *J. Acad. Ethics*, vol. 24, Art. no. 17, 2025.

[31] L. Mota, M. Mothabeng, and M. M. Seoane, "Perceived impact of generative artificial intelligence on students' learning experiences: A case of Botho University, Lesotho," in *Proc. Botho Univ. Int. Res. Conf. (BUIRC 2025)*, Atlantis Highlights Sustain. Dev., 2025, pp. 106–123.

[32] Q. Xie, M. Li, and F. Cheng, "Between regulation and accessibility: How Chinese university students navigate global and domestic generative AI," 2025, arXiv:2506.14377.

[33] M. Kassorla, M. Georgieva, and A. Papini, "AI literacy in teaching and learning: A durable framework for higher education (executive summary)," EDUCAUSE, 2024.

[34] M. Hibbert, E. Altman, T. Shippen, and M. Wright, "A framework for AI literacy," *EDUCAUSE Rev.*, 2024.

[35] X. Zhou and L. Schofield, "Developing a conceptual framework for artificial intelligence (AI) literacy in higher education," *J. Learn. Develop. Higher Educ.*, no. 31, 2024, doi: 10.47408/jldhe.vi31.1354.

[36] M. Gerlich, "AI tools in society: Impacts on cognitive offloading and the future of critical thinking," *Societies*, vol. 15, no. 1, Art. no. 6, 2025, doi: 10.3390/soc15010006.

[37] Martahani, R. Retnowati, S. Kurniasih, R. Suarti, and I. Rachman, "Conditional effects of generative AI on metacognitive learning," *Jurnal PAJAR (Pendidikan dan Pengajaran)*, vol. 10, no. 4, pp. 707–721, 2026.

[38] B. Degen, “Resurrecting Socrates in the age of AI: A study protocol for evaluating a Socratic tutor to support research question development in higher education,” 2025, arXiv:2504.06294.

[39] J. Barshay, “Students tried Khanmigo’s AI tutor. Most didn’t want its help,” *The Hechinger Report / FutureEd*, 2026. [Online]. Available: https://www.future-ed.org/students-tried-khanmigos-ai-tutor-most-didnt-want-its-help/

[40] M. Warschauer, *Technology and Social Inclusion: Rethinking the Digital Divide*. Cambridge, MA, USA: MIT Press, 2004.

[41] N. Rahman and G. Saunders, “Access to premium AI services is a significant concern for students,” *Wonkhe*, 2024. [Online]. Available: https://wonkhe.com/blogs/access-to-premium-ai-services-is-a-significant-concern-for-students/

[42] A. Doyle and A. Wilson, “Bridging the digital divide in higher education: From access to empowerment,” Liverpool John Moores Univ., Liverpool, U.K., 2025. [Online]. Available: https://researchonline.ljmu.ac.uk/id/eprint/28643/

[43] Y. Cheng, “Students’ experiences of generative AI in higher education: Understanding digital inequalities through a sociomaterial lens,” in *Proc. Int. Conf. Networked Learn.*, vol. 15, 2026, doi: 10.54337/nlc.v15.10884.

[44] V. Storey, “When AI writes the doctoral thesis: Reclaiming the oral defence as a learning development intervention,” *J. Learn. Develop. Higher Educ.*, no. 39, 2026, doi: 10.47408/jldhe.vi39.1809.

[45] D. Brumby, A. Cox, A. Sarkar, and S. Gould, “Sampled vivas are pivotal in combating AI cheating,” *Times Higher Educ.*, Mar. 6, 2025.

[46] I. M. Church, L. Drake, and M. Harris, “Using LLMs to support assessment of student work in higher education: A viva voce simulator,” 2025, arXiv:2511.05530.

[47] F. Su, M. O’Dea, M. Wood, and E. Seidl, “Embracing a new world: Authentic assessment designs in an age of generative artificial intelligence (GenAI),” *Perspectives: Policy Pract. Higher Educ.*, 2026, doi: 10.1080/13603108.2025.2601741.

[48] E. Ransome, “Designing for the AI era: A framework for integrating generative AI into assessment in higher education,” *Pract. Res. Higher Educ.*, 2026.

[49] Lehigh Univ. Center for Career & Professional Development, “Employer expectations for AI fluency: Strategic insights from Lehigh’s employer partners,” 2025.

[50] V. Braun and V. Clarke, “Using thematic analysis in psychology,” *Qual. Res. Psychol.*, vol. 3, no. 2, pp. 77–101, 2006, doi: 10.1191/1478088706qp063oa.