*Corresponding Author's Email: mopather@emeris.ac.za
Proceedings of the International Academic Conference on Education, Teaching and Learning
Vol.3, Issue. 1, 2026, pp. 67-75
DOI: https://doi.org/10.33422/iacetl.v3i1.1844

ISSN: 3030-1270 online

The Polemics of Ethical AI Usage in Higher Education: A Case Study Investigating the Axiological Expectations Mismatch

Mohini Meghan Grobler

Independent Institute of Education, Emeris, Faculty of Commerce, South Africa

Abstract

The rapid integration of generative artificial intelligence (GenAI) into higher education has created significant tensions around authorship, assessment authenticity, and academic integrity. This paper introduces **axiological expectations mismatch** as a conceptual framework for understanding a core governance problem: the divergence between institutional value frameworks and the ethical reasoning students apply when using AI in their academic work. Drawing on an interpretive qualitative case study at a single private higher education institution in South Africa, the study analyses twelve institutional documents produced between 2021 and 2025, supplemented by descriptive trend data from 3,854 plagiarism incidents over the same period. Schwartz's (2012) theory of basic values provides the theoretical lens; reflexive thematic analysis is the analytic method. Three findings emerge: first, a shift from prohibition to conditional permission for AI use, contingent on disclosure and authorship accountability; second, a reframing of academic integrity as a developmental process rather than a purely disciplinary matter; and third, evidence that policy adaptation improved institutional capacity to recognise and classify AI-related misconduct before it reduced its incidence. The paper argues that AI-related integrity disputes are better understood as conflicts between competing values (fairness, accountability, efficiency, and innovation) than as individual moral failings. Implications for policy design, assessment reform, and faculty development are discussed.

Keywords: Axiology; Academic Integrity; Artificial Intelligence; Value Mismatch; Ethics in Education

1. Introduction

Generative AI tools such as ChatGPT have moved rapidly from novelty to routine in student academic practice. Students now use them for brainstorming, outlining, paraphrasing, editing, and code generation (Cotton et al., 2023; Moya et al., 2024). A 2025 survey by the Higher Education Policy Institute found that 88% of students used generative AI for assessed work, with 92% using AI tools of some kind (Freeman, 2025). Jisc (2025) reports similar patterns and documents strong student demand for clearer, course-specific guidance.

Institutional policy has not kept pace. UNESCO (2023a) has documented that most higher education institutions lack formal AI policies, with significant regional variation in readiness. The Quality Assurance Agency (QAA, 2023) and the Tertiary Education Quality and Standards Agency (TEQSA, n.d.) have both moved away from prohibition-focused frameworks toward governance models centred on disclosure, human oversight, and assessment redesign.

This paper argues that the challenge posed by GenAI is fundamentally **axiological** in nature. Traditional academic integrity frameworks assume a settled hierarchy of values: honesty, fairness, accountability, and independent effort. Students may approach the same tools through a different set of priorities, such as efficiency, adaptability, and practical problem-solving. When these orientations collide, the result is an **axiological expectations mismatch**, a conflict between competing interpretations of what legitimate academic work means in an AI-rich environment rather than simple rule-breaking.

This paper presents an interpretive qualitative case study of one private South African institution that developed increasingly explicit AI guidance between 2021 and 2025. Three research questions structure the inquiry:

- **RQ1:** How do institutional documents construct the boundaries of ethical and unethical AI use in student assessment?
- **RQ2:** What value tensions emerge within these documents when academic integrity is considered alongside innovation, authorship, and student support?
- **RQ3:** How does the institution seek to resolve these tensions through policy language, guidance, and integrity procedures?

2. Theoretical Framework and Literature Review

Early research on AI in higher education focused primarily on technical applications and performance outcomes, with limited attention to ethics or institutional values (Zawacki-Richter et al., 2019). The emergence of generative AI has reoriented this field considerably. Cotton et al. (2023) argue that tools like ChatGPT pose genuine risks to conventional understandings of plagiarism and originality while also opening pedagogical opportunities. Moya et al. (2024) call for AI to be understood not simply as a cheating tool but as a challenge involving literacy, equity, and policy.

Comparative policy work reveals that institutional responses remain uneven. Aristombayeva et al. (2025) found that university AI policies are often reactive, fragmented, and unevenly implemented, with institutions frequently articulating strong values without operational clarity. The gap between policy rhetoric and practical implementation is a recurring feature of this literature.

Global South contexts deserve specific attention. Rather than framing African institutions as lagging behind, recent scholarship demonstrates that some are developing context-sensitive, enabling governance models (Thaldar et al., 2025). The University of KwaZulu-Natal, for example, built a responsible AI framework aligned with national priorities, disclosure norms, and capacity-building objectives, demonstrating that governance innovation is not confined to the Global North.

The study's theoretical anchor is Schwartz's (2012) theory of basic values, which organises values into motivational patterns of compatibility and conflict. Relevant tensions for this study include *openness to change* versus *conservation*, and *self-enhancement* versus *self-transcendence*. In an AI context, values such as self-direction and achievement may motivate

tool experimentation, while conformity, security, and benevolence may underpin institutional concern for fairness and trust. Koscielniak and Bojanowska (2019) link these value orientations to academic dishonesty, finding that socially oriented values negatively predict dishonest conduct while personally focused values (power, hedonism, stimulation) predict it positively.

Howard et al. (2010) complicate simple accounts of plagiarism by showing that problematic source use is often bound up with developing writing proficiency rather than deliberate deception. This distinction matters for AI governance: not all inappropriate AI use reflects intentional misconduct. Faculty research supports this reading. Alsharefeen and Al Sayari (2025) found that staff frequently prefer educative over punitive responses, particularly where policy ambiguity or limited student understanding appears to be the root cause.

Assessment redesign has become central to the literature. QAA (2023) advocates for authentic, synoptic assessments rather than detection-only approaches. Khlaif et al. (2025) propose an "Against, Avoid, Adopt, Explore" framework to help faculty redesign tasks for an AI-enabled environment. Policy clarity, AI literacy, and assessment reform are increasingly treated as inseparable concerns.

Reflexive thematic analysis, as described by Braun and Clarke (2023), provides the methodological foundation for this study. Their approach emphasises theoretical transparency, iterative theme development, and active researcher engagement with data. It is well suited here because institutional policy texts are value-laden documents that construct meaning rather than neutrally record it.

3. Methodology

Research Design

This study uses an interpretive qualitative case study design, augmented by embedded descriptive analysis of institutional trend data. The case is bounded by a single private higher education institution and the period 2021 to 2025. Qualitative document analysis takes priority; the quantitative data provide contextual texture rather than causal evidence.

Institutional Context

The institution is a private higher education provider in South Africa, offering programmes across business, law, education, information technology, and humanities. It operates primarily through distance and blended delivery modes and serves a predominantly undergraduate student population. It is anonymised to preserve confidentiality. Resource constraints, digital access inequalities, and policy implementation capacity all shape how AI governance problems materialise in this context.

Data Sources

The primary dataset comprises twelve institutional documents produced between 2021 and 2025: academic integrity policies, responsible AI guidance, assessment guidelines, student-facing communications, and integrity procedure manuals. Documents were selected if they were formally issued during the study period, engaged substantively with academic integrity, authorship, AI use, or assessment, and provided normative or procedural guidance. Administrative notices and duplicate versions were excluded.

The secondary dataset consists of 3,854 plagiarism incident records spanning the same period. Each record includes the year and month of the incident, the academic school, assessment type, student level, and the institution’s classification of AI involvement. These data are used descriptively to contextualise the document analysis and do not support causal claims.

Defining AI-Related Misconduct

"AI-related" misconduct refers to cases formally classified by the institution as involving generative AI assistance, AI-generated text submission, or undisclosed AI-supported authorship substitution, interpreted under the classification framework operative at the time. Because detection methods and classification criteria evolved substantially over the study period, an observed increase in AI-related cases may reflect actual behavioural change, improved institutional identification capacity, revised classification protocols, or some combination of all three.

Analytic Approach

Document analysis followed the reflexive thematic analysis process described by Braun and Clarke (2023): familiarisation, open coding, theme development, review, and interpretive writing. Initial coding was semantic, focusing on explicit meanings. The quantitative data were examined alongside the policy timeline to assess whether policy shifts corresponded with changes in case recognition or distribution, not to test causal hypotheses.

Limitations and Reflexivity

The single-institution design limits generalisability. The findings describe institutional discourse and governance, not student experience or intent. Trustworthiness was supported through repeated corpus readings, cross-document comparison, memo-writing, and attention to disconfirming evidence.

A more fundamental constraint bears directly on the core theoretical claim. The axiological expectations mismatch concept posits a divergence between institutional and student value frameworks, yet this study analyses only one side of that divergence. Student values are inferred entirely from what institutional documents say about students, not from what students say about themselves. This is not a gap that can be filled by reading the documents more carefully; it is a structural limitation of the data. The mismatch framing therefore functions as an interpretive hypothesis rather than a demonstrated finding. It is consistent with the institutional evidence and with external survey data on student AI use (Freeman, 2025; Jisc, 2025), but it cannot be confirmed without direct access to student reasoning. Claims in this paper about student value orientations should be read with that constraint in mind. Establishing or refining the mismatch concept empirically requires future work that gathers student perspectives through interviews, focus groups, or survey instruments designed to capture value priorities in AI-use contexts.

4. Findings

Embedded Trend Overview

Of the 3,854 plagiarism incidents recorded between 2021 and 2025, 680 (17.64%) were classified as AI-related. The most striking pattern is temporal: AI-related cases rose from 0% in 2021 to 92.52% of all reported incidents by 2025. That figure should not be read as a straightforward measure of rising misconduct. It reflects two processes that the data cannot separate: actual changes in student behaviour as generative AI tools became more accessible, and improvements in institutional capacity to identify and classify AI-related cases as distinct from conventional plagiarism. Both were almost certainly operating across this period.

Case distribution varied by school. The School of Education recorded the highest volume (1,187 cases, 30.8%), followed by Management (981, 25.5%), Information Technology (673, 17.5%), Law (623, 16.2%), and Humanities and Social Sciences (390, 10.1%). The School of Law recorded the highest proportion of AI-related cases within its total, suggesting that disciplinary context shaped both how AI misuse manifested and how it was detected.

Assessment format was also a significant variable. The largest share of cases involved Assignment 1 (1,147, 29.8%), followed by regular assignments (455, 11.8%) and portfolio-of-evidence submissions (395, 10.2%). Take-home written assessments appear most vulnerable because they provide more opportunity for external text generation and make authorship harder to verify.

The institution also reported statistically significant associations across several variables, summarised in Table 1. All figures are drawn from institutional reporting and have not been independently recalculated; they are included as contextual descriptors of patterning rather than independently verified inferential statistics.

Table 1: *Reported Statistical Associations Between Case Variables and AI Classification (Institutional Data, 2021–2025)*

Variable	**Test**	χ^2 / U	**Significance (*p*)**
Academic school vs. AI classification	Chi-square	148.31	< .001
Assessment type vs. AI classification	Chi-square	626.08	< .001
Academic level vs. AI classification	Chi-square	118.41	< .001
Offender status (first-time vs. repeat)	Mann–Whitney U	not disclosed	< .001

Note. "Not disclosed" indicates the test statistic was not reported in institutional records. All figures are from institutional reporting and have not been independently verified.

64.89% of undergraduate cases involved first-time offenders, a pattern consistent with norm confusion or limited policy internalisation rather than deliberate cheating.

Policy Adaptation Timeline

The institutional documents show a phased evolution. Between 2021 and 2022, integrity discourse centred on conventional plagiarism and originality. In 2023, GenAI was explicitly identified as a distinct integrity challenge. By 2024 to 2025, policy language had become substantially more differentiated, distinguishing among acceptable AI use (disclosed, limited support), concerning but disclosed AI use (warranting developmental intervention), and undisclosed AI misuse (warranting formal investigation). This tripartite categorisation reflects a deliberate effort to move beyond binary prohibition and develop a more proportionate governance vocabulary.

Theme 1: From Prohibition to Conditional Permission

Earlier documents positioned AI primarily as a threat to integrity, emphasising control and prevention. Later texts adopted a more differentiated stance: AI use was permissible when disclosed, limited in scope, and compatible with the assessment's learning objectives. This shift reflects a reordering of institutional values, moving from conformity and security toward

conditional endorsement of self-direction and innovation. Ethical legitimacy became contingent rather than presumed. AI was not treated as a neutral study aid but as a permitted resource only when consistent with authorship accountability.

Theme 2: Developmental Integrity Rather Than Pure Punishment

Across the corpus, the institution consistently paired disciplinary language with remediation, reflective learning, and student support, particularly for disclosed but problematic AI use. Students were positioned simultaneously as potential offenders and as learners in an unfamiliar technological environment. This orientation aligns with sector guidance (QAA, 2023; UNESCO, 2023b) and with faculty research showing that staff often prefer educative responses where policy ambiguity or limited understanding appears to be the root cause (Alsharefeen & Al Sayari, 2025).

This developmental model had limits. Its effectiveness depended on sufficiently precise guidance. Where policy language remained broad, students continued to struggle to identify what meaningful disclosure required or where supported writing became authorship substitution. Under those conditions, remediation functioned as a reactive measure rather than a genuinely educative one.

Theme 3: Persistent Ambiguity at the Assistance and Authorship Boundary

The third theme is the persistent instability of the line between legitimate assistance and inappropriate authorship transfer. Institutional documents endorsed AI as a support tool while warning against overreliance and non-disclosure, yet the precise boundary remained unclear. Activities such as idea generation, structural planning, paraphrasing, and language editing may be educationally beneficial while still raising questions about who is performing the assessed intellectual work. This ambiguity is where the **axiological expectations mismatch** is most visible. Institutions defend fairness, accountability, and authentic scholarship. Students may interpret the same tools through values of efficiency and practical competence. The dispute is not simply about whether AI was used but about what that use means within a specific learning context. HEPI and Jisc both report that students frequently perceive mixed institutional signals and want clearer, module-specific guidance (Freeman, 2025; Jisc, 2025).

Theme 4: Policy Adaptation Changed Governance Before It Changed Incidence

The most defensible reading of the trend data requires holding two explanations apart. The first is actual incidence: the number of students genuinely engaging in AI-related misconduct may have risen as tools became more accessible. The second is detectability and classification: as policies became more explicit, the institution developed clearer criteria for identifying and categorising AI-assisted cases, which means cases that previously went unrecorded or were folded into generic plagiarism counts began appearing as a distinct category. Both processes were almost certainly operating simultaneously over this period, and the data cannot separate them. What the trend line captures, therefore, is the combined product of behavioural change and improved institutional recognition, not a clean measure of either alone.

This matters for how the findings are interpreted. An institution that sees AI-related cases rise sharply after introducing a new policy has not necessarily failed; it may have succeeded in making previously invisible conduct visible and classifiable. The high proportion of first-time undergraduate offenders further supports this reading: repeat, deliberate offending would be expected to produce a different profile. Students operating under evolving expectations they had not yet internalised are more consistent with what the data show. The tripartite categorisation of AI use appears to have improved governance by enabling more proportionate responses, directing developmental cases toward remediation and reserving formal investigation for serious or deliberately concealed misconduct.

5. Discussion

The trend data described above do not establish that more students were cheating over time; they establish that more cases were being identified and classified as AI-related. That distinction is the starting point for interpreting everything that follows. Policy adaptation at this institution did not immediately reduce misconduct; it initially made misconduct more legible and classifiable. Legibility and incidence are not the same thing. A rising case count after a policy change may mean more misconduct is occurring, or it may mean the same misconduct is now being identified and recorded that was previously invisible. The data in this study cannot disentangle these two processes, and claims about actual behavioural trends should be read with that in mind. What can be said is that governance improved: the institution developed a more proportionate response architecture and a clearer vocabulary for distinguishing types of AI-related conduct. The concept of **axiological expectations mismatch** helps explain why incidence is likely to remain high even as governance improves: where students and institutions hold different value hierarchies, disputes over legitimate AI use persist regardless of how clearly the rules are written.

This interpretation is consistent with international evidence. UNESCO (2023a) documents global movement toward more explicit AI guidance alongside continued unevenness in implementation. HEPI and Jisc find that student AI adoption consistently outpaces institutional policy consolidation (Freeman, 2025; Jisc, 2025). The South African case examined here illustrates these dynamics in a context shaped by resource constraints and digital inequality, conditions that make governance both more urgent and more difficult.

This study extends values-based approaches to academic integrity. Where Koscielniak and Bojanowska (2019) show how individual values shape vulnerability to dishonesty, the present case shows how institutional documents encode specific value priorities and seek to shape behaviour through moral framing. The mismatch concept bridges institutional discourse and probable student interpretation, offering an analytical framework for understanding integrity disputes without requiring direct access to student reasoning, though that access remains a priority for future research.

The findings reinforce the consensus that detection alone is insufficient (QAA, 2023; TEQSA, n.d.; Khlaif et al., 2025). The case also contributes to African scholarship on AI governance by demonstrating that a South African institution can develop a contextually responsive, enabling model rather than simply importing restrictive frameworks from elsewhere (Thaldar et al., 2025).

6. Implications for Policy and Practice

Temper short-term expectations. Institutions should not interpret initial increases in identified AI-related cases as evidence of policy failure. In periods of rapid change, policy clarification may improve reporting and classification before it reduces incidence. Improved governance may look, temporarily, like a worsening problem.

Move beyond generic guidance. Blanket statements about when AI is "permitted" are insufficient. Policies need task-specific boundaries, clear articulation of what counts as a meaningful intellectual contribution, and a pedagogical rationale for disclosure that goes beyond procedural compliance.

Provide discipline-specific examples. Different disciplines require different guidance. In essay-based humanities modules, students might use AI for brainstorming or language feedback with prompt disclosure. In law, AI may assist with issue-spotting but not with uncited legal reasoning. In IT, AI can support debugging if students annotate AI-assisted sections and

can defend their logic. In management, AI-generated strategic scenarios are acceptable when critically evaluated and independently justified. These distinctions align with assessment redesign frameworks that champion process evidence and transparent human-AI collaboration (QAA, 2023; Khlaif et al., 2025).

Invest in staff development. Inconsistent faculty interpretation of AI policies directly amplifies the axiological mismatch that institutions seek to reduce. Sustained professional development is not optional.

Treat academic integrity as an educational project. Where institutions articulate values clearly, contextualise expectations at the task level, and distinguish developmental misuse from deliberate concealment, they are more likely to build genuine integrity than to produce defensive compliance.

7. Conclusion

This paper has argued that the challenge of generative AI in higher education is fundamentally axiological. Through an interpretive qualitative case study integrated with embedded institutional trend data, it shows that policy adaptation played a pivotal role at the institution under study, but its initial impact was to make misconduct more visible and manageable rather than to reduce its frequency. The concept of **axiological expectations mismatch** explains why this is likely to be a common pattern: when students and institutions hold different value frameworks, governance problems persist even as policies become clearer. Resolution requires better alignment of assessment design, pedagogical rationale, and institutional communication, not just better rules.

This study contributes a values-based explanation for why policy clarification does not rapidly reduce AI-related misconduct, demonstrates how qualitative discourse analysis can illuminate the ethical architecture of AI governance, and underscores the need to pair policy with assessment redesign, staff development, and discipline-specific guidance. Future research should extend this work through comparative institutional analysis, cross-disciplinary case comparison, and direct qualitative engagement with students and faculty to examine how the axiological expectations mismatch is resolved in practice.

References

Alsharefeen, R., & Al Sayari, N. (2025). Examining academic integrity policy and practice in the era of AI: A case study of faculty perspectives. Frontiers in Education, 10, Article 1621743. https://doi.org/10.3389/feduc.2025.1621743

Aristombayeva, M., Satybaldiyeva, R., Maung, B. M., & Lee, D. (2025). Guiding the uncharted: The emerging (and missing) policies on generative AI in higher education. Frontiers in Education, 10, Article 1644081. https://doi.org/10.3389/feduc.2025.1644081

Braun, V., & Clarke, V. (2023). One size fits all? What counts as quality practice in (reflexive) thematic analysis? International Journal of Transgender Health, 24(1), 1–8. https://doi.org/10.1080/26895269.2023.2163990

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. Journal of Further and Higher Education, 48(2), 228–239. https://doi.org/10.1080/14703297.2023.2190148

Freeman, J. (2025, February 26). Student generative AI survey 2025. Higher Education Policy Institute. https://www.hepi.ac.uk/reports/student-generative-ai-survey-2025/

Howard, R. M., Serviss, T., & Rodrigue, T. K. (2010). Writing from sources, writing from sentences. Writing & Pedagogy, 2(2), 177–192. https://doi.org/10.1558/wap.v2i2.177

Jisc. (2025). Student perceptions of AI 2025. https://www.jisc.ac.uk/reports/student-perceptions-of-ai-2025

Khlaif, Z. N., Alkouk, W. A., Salama, N., & Abu Eideh, B. (2025). Redesigning assessment in the era of generative artificial intelligence: A qualitative study of faculty motivations and challenges. Education Sciences, 15(2), Article 174. https://www.mdpi.com/2227-7102/15/2/174

Koscielniak, M., & Bojanowska, A. (2019). The role of personal values and student achievement in academic dishonesty. Frontiers in Psychology, 10, Article 1887. https://doi.org/10.3389/fpsyg.2019.01887

Moya, B. A., Eaton, S. E., Pethrick, H., Hayden, K. A., Brennan, R., Wiens, J., & McDermott, B. (2024). Academic integrity and artificial intelligence in higher education contexts: A rapid scoping review. Canadian Perspectives on Academic Integrity, 7(3). https://doi.org/10.55016/ojs/cpai.v7i3.78123

Quality Assurance Agency for Higher Education. (2023). Reconsidering assessment for the ChatGPT era: QAA advice on developing sustainable assessment strategies. https://www.qaa.ac.uk/docs/qaa/members/reconsidering-assessment-for-the-chat-gpt-era.pdf

Schwartz, S. H. (2012). An overview of the Schwartz theory of basic values. Online Readings in Psychology and Culture, 2(1), Article 11. https://scholarworks.gvsu.edu/orpc/vol2/iss1/11/

Tertiary Education Quality and Standards Agency. (n.d.). Gen AI – academic integrity and assessment reform. https://www.teqsa.gov.au/guides-resources/higher-education-good-practice-hub/gen-ai-knowledge-hub/gen-ai-academic-integrity-and-assessment-reform

Thaldar, D., Botes, M., Badru, A., Chenia, H., Duma, S., Dlamini, S. B., Amin, N., Hugo, W., Govender, R., Bruce-Brand, J., Vosloo, A., Koorbanally, N. A., & Chuturgoon, A. (2025). Generative AI governance in higher education: A case study from Africa. Frontiers in Political Science, 7, Article 1666661. https://doi.org/10.3389/fpos.2025.1666661

UNESCO. (2023a). Guidance for generative AI in education and research. https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research

UNESCO. (2023b). ChatGPT and artificial intelligence in higher education: Quick start guide. https://unesdoc.unesco.org/ark:/48223/pf0000385146

UNESCO. (n.d.). UNESCO survey: Two-thirds of higher education institutions have or are developing guidance on AI use. https://www.unesco.org/en/articles/unesco-survey-two-thirds-higher-education-institutions-have-or-are-developing-guidance-ai-use

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education: Where are the educators? International Journal of Educational Technology in Higher Education, 16, Article 39. https://doi.org/10.1186/s41239-019-0171-0