AI and academic integrity in the Global South: Equity and policy–practice tensions in South African Higher Education

E. G. Makwakwa
Department of Mathematics Education
UNISA
Pretoria, South Africa
https://orcid.org/0000-0002-6565-3235

M. C. Thobejane
Department of ICT Infrastructure and Systems
UNISA
Pretoria, South Africa
https://orcid.org/ 0009-0001-1347-0736

ABSTRACT

Artificial intelligence is reshaping assessment practices in higher education, yet its ethical governance in Global South open distance e-learning (ODeL) contexts is complicated by persistent inequalities in digital access, skills, and institutional capacity. This article examines how AI-related academic integrity challenges emerged in a large undergraduate data handling and probability module within a South African teacher education programme offered through an ODeL institution. Framed by Sociotechnical Systems Theory, the study analyses assessment submissions from a cohort of 3,088 students and a purposive sample of 229 submissions, triangulated with written reflections from experienced module markers. The findings identify recurring integrity challenges that cannot be reduced to individual misconduct, including AI-shaped graphical artifacts and procedural calculations, handwritten and scanned submissions that bypass similarity checking, and non-acceptance of End-User License Agreements that weaken accountability. Marker evidence shows how these practices are amplified by weak digital literacies, device constraints, and equity-driven accommodations that unintentionally create loopholes in integrity systems. The article argues that ethical AI governance in resource-constrained ODeL environments requires alignment among policy, technology, and pedagogy, including inclusive integrity pathways, strengthened support for digital and AI literacy, and assessment designs that foreground reasoning, conceptual understanding, and pedagogical interpretation.

Keywords: Artificial intelligence, academic integrity, Open Distance e-Learning (ODeL), sociotechnical systems theory, Global South, teacher education, ethical AI governance

INTRODUCTION

The rapid integration of artificial intelligence (AI) into higher education has reshaped assessment practices, student learning behaviours, and institutional governance worldwide. Generative AI tools such as ChatGPT and Claude can now produce fluent prose, structured explanations, and data representations that challenge traditional assumptions about authorship, originality, and evidence of learning (Rudolph, Tan, and Tan 2023; Sullivan, Kelly, and McLaughlan 2023). As a result, universities are grappling with how to maintain academic integrity while adapting assessment practices to an AI-pervasive environment.

Much of the emerging scholarship on AI and academic integrity is situated in well-resourced Global North contexts, where stable digital infrastructure, widespread device access, and relatively uniform digital literacies are often assumed (Oyedemi and Mogano. 2018; Swartz et al. 2019). These assumptions do not hold in many Global South contexts, where higher education systems operate under conditions of structural inequality, uneven access to technology, and diverse student preparedness. In such settings, AI-related integrity challenges manifest differently and cannot be addressed through policy or assessment models developed for materially advantaged institutions.

This article aligns with emerging Southern scholarship that cautions against the uncritical transfer of Global North models of AI governance, assessment, and academic integrity into materially unequal educational environments. In South African higher education, questions of AI ethics cannot be separated from historical inequalities, uneven digital infrastructures, and the social justice mandate of widening participation institutions. From this perspective, ethical AI governance involves not only protecting academic standards but also ensuring epistemic justice, equitable participation, and contextually responsive assessment practices within resource-constrained environments.

In South Africa, these challenges are particularly pronounced in the Open Distance and e-Learning (ODeL) sector. ODeL institutions serve large, diverse cohorts of predominantly working-class and first-generation students, many of whom rely on mobile devices, experience intermittent connectivity, and have limited exposure to advanced digital tools (Walters and Kop 2009; Czerniewicz et al. 2020). While the ODeL model is explicitly oriented toward widening access and promoting equity (Department of Higher Education and Training 2014), this mandate creates tension when institutions attempt to enforce academic integrity policies, especially those that rely on technological compliance, such as typed submissions and plagiarism-detection software.

Standard integrity responses, including algorithmic detection tools, high-stakes surveillance, and punitive disciplinary frameworks, risk disproportionately disadvantaging

students with limited digital access, uneven academic preparation, or constrained institutional support. These tensions are experienced most acutely by academic staff serving as markers, who are tasked with implementing integrity policies while evaluating submissions that often reflect a complex mix of AI-mediated work, access constraints, and foundational conceptual gaps. What emerges is not simply a matter of individual student misconduct, but a systemic misalignment between institutional expectations, technological infrastructures, and the lived realities of students operating within constrained sociotechnical environments. Not all integrity-related practices identified in this study stem directly from generative AI use. Some emerged through submission-format practices, technological constraints, or institutional accommodations associated with unequal access conditions. The study therefore situates AI within a broader sociotechnical ecology of academic integrity in ODeL environments.

Traditional academic integrity frameworks tend to conceptualise violations as deliberate acts of dishonesty, locating responsibility primarily within individual students (Bretag 2016; McCabe, Butterfield, and Treviño 2012). However, such individualised framings obscure how AI-related integrity challenges are co-produced by assessment design, institutional policy, technological affordances, and unequal access to resources (Eaton 2021). In ODeL contexts, where students may turn to AI tools to compensate for limited academic support, language barriers, or time poverty, categorising all unauthorised AI use as "cheating" risks reproducing the very inequities these institutions are mandated to address (Boughey and McKenna 2021).

This article examines AI-related academic integrity challenges in a South African ODeL institution through the lens of Sociotechnical Systems Theory (STS). STS conceptualises organisational outcomes as the product of interdependent social and technical systems rather than isolated individual actions (Trist and Bamforth 1951; Mumford 2006). Applying this lens allows the study to move beyond detection-focused or punitive interpretations of AI misuse and instead examine how integrity breakdowns emerge from interactions between students, assessment design, digital infrastructure, institutional policies, and equity-driven accommodations.

Empirically, the study draws on a large-scale case analysis of an undergraduate data handling and probability module with a cohort of 3,088 students. Artefactual analysis of a purposive sample of assessment submissions is triangulated with reflective observations from experienced module markers, whose professional judgments provide critical insight into how AI-shaped work manifests in practice and how integrity policies operate on the ground. Marker perspectives are treated not as anecdotal commentary, but as expert data that illuminate the practical consequences of sociotechnical misalignment in assessment.

By foregrounding a Global South ODeL context and centering the practitioner lens of markers, this article contributes empirical evidence to international debates on AI and academic integrity. It challenges binary framings of AI as either a threat to be prohibited or a tool to be embraced uncritically. Instead, it argues for integrity governance approaches that balance equity with accountability, recognizing that ethical AI integration depends as much on institutional coherence and assessment design as on student compliance. In doing so, the article contributes directly to the special issue theme of "Artificial Intelligence in Higher Education in the Global South: Practices and Possibilities" by examining how AI-related assessment practices, institutional integrity mechanisms, and equity concerns intersect within a large-scale South African ODeL context.

LITERATURE REVIEW AND THEORETICAL FRAMEWORK

This study is grounded in scholarship on artificial intelligence, academic integrity, and equity in higher education, with particular attention to how these issues manifest in Global South open distance e-learning (ODeL) contexts. Rather than treating AI misuse as an individual ethical failure, the literature increasingly recognises integrity challenges as emerging from the interaction between technological systems, institutional practices, and student access conditions. This section reviews key debates on AI and assessment, equity and ODeL, and positions Sociotechnical Systems Theory (STS) as the analytical framework guiding the study.

AI, academic integrity, and assessment

Generative AI and the disruption of assessment

Generative AI tools have fundamentally disrupted conventional assessment practices in higher education. Traditional formats such as essays and short-answer questions, long relied upon to assess understanding and critical thinking, are increasingly vulnerable to AI-generated outputs that are fluent, coherent, and contextually appropriate (Rudolph et al 2023; Swiecki et al. 2022). This development complicates authorship verification and challenges long-standing assumptions about originality and independent student work (Luo 2024).

Beyond authorship concerns, generative AI introduces epistemic risks. Large language models generate responses based on probabilistic patterning rather than verification, resulting in plausible but inaccurate outputs, commonly referred to as hallucinations (Rudolph et al. 2023). In quantitative disciplines such as statistics, these risks are amplified when visually polished representations obscure underlying conceptual or computational errors.

Limits of detection-only integrity approaches

Institutional responses to academic misconduct have historically prioritised detection and deterrence through surveillance technologies, plagiarism-detection software, and proctoring systems (Lawrence-Slater 2001). However, growing scholarship argues that detection-only approaches are increasingly ineffective in the context of generative AI, which evolves more rapidly than detection tools and is easily adapted to bypass text-based checks (Luo 2024).

Moreover, surveillance-driven integrity models risk shifting pedagogical relationships from trust to control, potentially alienating students and obscuring the structural conditions that shape integrity-related behaviours (Owan et al. 2023a). This tension is particularly visible to frontline assessors, who encounter AI-shaped submissions that appear anomalous but cannot always be substantiated through technical evidence, especially when submissions bypass similarity checking.

Illusions of mastery and surface correctness

A critical concern in AI-mediated assessment is the emergence of "illusion of mastery", where students submit outputs that are procedurally correct and linguistically polished yet reflect limited conceptual engagement (Crompton and Burke 2023). Generative AI can mask misunderstanding by producing complete solutions that students accept uncritically, undermining formative learning processes and metacognitive development.

Markers' observations across multiple studies highlight this pattern, noting correct answers presented without explanation, justification, or evidence of reasoning (Rudolph et al. 2023). In this context, assessment validity is compromised when correctness of output is mistaken for evidence of understanding, particularly in disciplines where interpretation and reasoning are central.

Equity, access, and ODeL in the Global South

Digital inequality and structural constraints

In Global South contexts, AI-related integrity challenges intersect with persistent digital inequalities. In South Africa, students' access to devices, connectivity, and digital skills remains uneven, shaped by historical and socioeconomic disparities (Oyedemi and Mogano. 2018; Czerniewicz et al. 2020). Many ODeL students rely primarily on mobile devices and intermittent data access, limiting their ability to comply with technically demanding assessment requirements.

Markers routinely report receiving handwritten or scanned submissions, incorrect file formats, and evidence of difficulty using basic software tools, despite explicit instructions.

Without systemic intervention, AI adoption risks reinforcing educational exclusion by privileging students with greater technological capital.

ODeL scale, asynchronicity, and integrity governance

ODeL institutions operate at scale and rely heavily on asynchronous assessment models. While this enhances access and flexibility, it complicates monitoring and standardisation of assessment conditions (Howell 2024). Ensuring equivalent evidence of learning across diverse contexts requires careful alignment between pedagogy, technology, and assessment design (Chen and Lai 2024).

Equity-driven accommodations, such as accepting scanned submissions or relaxing similarity checking, are often introduced to avoid excluding students with limited access. However, research increasingly notes that such accommodations can generate unintended integrity consequences, including inconsistent enforcement and weakened accountability (Lawrence-Slater 2001; Luo 2024). These tensions are most acutely experienced by markers, who must navigate the practical consequences of policy flexibility in assessment evaluation.

Sociotechnical Systems Theory (STS) as analytical framework

This study adopts Sociotechnical Systems Theory (STS) to analyse academic integrity challenges as emergent, systemic phenomena rather than individual moral failures. STS conceptualises organisations as comprising interdependent social and technical subsystems, where outcomes depend on the alignment, or misalignment, between people, technologies, and institutional structures (Trist and Bamforth 1951; Mumford 2006).

Applied to assessment, STS highlights how integrity challenges arise when technical systems (such as submission platforms, AI tools, and plagiarism-detection software) are poorly aligned with social practices (student access conditions, marker judgement, equity-driven policies). Key STS concepts guiding the analysis include mutual constitution between humans and technologies, variance control at the source of breakdowns rather than through downstream policing, and the identification of policy–practice gaps.

Markers occupy a critical position within this sociotechnical system. As social actors tasked with operationalising technical policies, their experiences provide diagnostic insight into where integrity mechanisms fail in practice. From an STS perspective, their observations are not anecdotal but constitute empirical evidence of systemic misalignment.

By adopting STS, this study moves beyond binary debates about AI prohibition versus permission. Instead, it frames academic integrity as a co-produced outcome shaped by institutional design choices, technological affordances, and structural inequalities. This

framework underpins the analysis of assessment artefacts and marker observations presented in the Findings and informs the study's recommendations for ethical and equitable AI governance in ODeL contexts.

Ethical AI governance in the Global South

Ethical debates on AI in higher education have increasingly moved beyond questions of plagiarism detection toward broader concerns of fairness, transparency, equity, and epistemic inclusion (Alfaleh 2026; Dabis and Csáki 2024). In Global South contexts, ethical AI governance cannot be reduced to surveillance or technological enforcement alone, because students participate in digitally unequal environments shaped by structural disadvantage (Guadu, Dibekulu and Menberu 2025; Swart and Shuttleworth 2021). Scholars have cautioned that integrity systems relying heavily on automated detection tools may unintentionally reproduce exclusion by privileging students with greater digital access and technological fluency (Guadu et al 2025; Swart and Shuttleworth 2021). Ethical governance in ODeL environments, therefore, requires balancing accountability with inclusion, while recognizing how institutional policies, technological infrastructures, and socioeconomic realities shape students' engagement with AI tools (Alfaleh 2026; Dabis and Csáki 2024).

CONTEXT OF THE STUDY

This study was conducted at a large public open distance e-learning (ODeL) institution in South Africa serving geographically dispersed undergraduate students across urban, peri-urban, and rural contexts. The institution enrols high-volume cohorts and delivers teaching and assessment predominantly online and asynchronously, with flexibility and scalability central to its access mandate.

Module and student cohort

The case examined is an undergraduate data handling and probability module offered within a teacher education programme. The module focuses on foundational statistical concepts, including data classification, graphical representation, measures of central tendency and dispersion, and introductory probability. Emphasis is placed on conceptual understanding and pedagogical interpretation, given students' future roles as mathematics teachers.

The cohort comprised 3,088 registered students. As is typical in ODeL contexts, students varied widely in schooling background, academic preparedness, and access to digital resources. While some students had reliable access to laptops and stable internet connectivity, others

relied primarily on smartphones and intermittent data access, resulting in wide variation in submission formats and digital competence.

Assessment design and integrity mechanisms

The assessment analysed was a high-stakes, open-book online task requiring students to construct tables and graphs using spreadsheet software, perform basic statistical calculations, and provide written explanations of their reasoning. Students were instructed to submit typed, text-selectable responses compatible with plagiarism-detection software, and the use of generative AI tools to produce answers was explicitly prohibited.

Institutional academic integrity mechanisms included plagiarism-detection software integrated into the learning management system. Students were required to accept an End-User Licence Agreement (EULA) for similarity reports to be generated. In principle, this ensured consistent screening of submissions. In practice, however, equity-driven accommodations across the institution, particularly for students with limited device access or digital skills, sometimes resulted in the acceptance of scanned or handwritten submissions and instances of EULA non-acceptance.

For markers, this context created a tension between upholding integrity standards and accommodating diverse access conditions. Submissions that bypassed similarity checking or arrived in non-standard formats limited the consistent application of integrity mechanisms and shaped the integrity challenges examined in this study.

METHODOLOGY

Research paradigm and approach

This study adopted an interpretive qualitative approach, situated within a sociotechnical systems framework. The aim was not to quantify the prevalence of academic misconduct, but to examine how AI-related integrity challenges emerged within a specific assessment context through the interaction of technological systems, institutional practices, and student behaviours. An interpretive paradigm is appropriate where the focus is on understanding meaning, practice, and context rather than measurement or prediction (Oates 2006). This aligns with the study's emphasis on patterns of practice and systemic conditions rather than individual attribution.

Research design

A qualitative case study design was employed, focusing on a single high-enrolment undergraduate module's major assessment. Case study designs are particularly suited to

examining complex, context-bound phenomena where boundaries between the phenomenon and its setting are not clearly defined (Gaya and Smith 2016). The study combined artefact analysis of student assessment submissions with qualitative insights from experienced module markers.

Sociotechnical Systems Theory (STS) guided the research design by directing analytic attention to how integrity challenges were co-produced by social actors (students and markers), technical artefacts (AI tools, submission platforms, plagiarism-detection systems), and organisational rules (assessment requirements and integrity policies) (Trist and Bamforth 1951; Mumford 2006).

Data sources and triangulation

Two complementary data sources were used to enable methodological triangulation, strengthening the credibility and interpretive depth of the findings (Braun and Clarke 2006; Maguire and Delahunt 2017).

First, assessment submissions from a cohort of 3,088 students were examined. A purposive sample of 229 submissions was selected for detailed analysis. The sample was not intended to be statistically representative, but to capture a range of integrity-relevant phenomena, including non-standard submission formats, EULA non-acceptance, and artefactual indicators of AI-shaped work. Submissions were treated as digital artefacts, with analysis focusing on formatting features, software signatures (such as LaTeX syntax), graphical representations, and submission metadata.

Second, qualitative written reflections and feedback from three experienced module markers (**Marker 1, Marker 2, and Marker 3**) were analysed. Each marker had assessed several hundred scripts and provided expert observations on recurring errors, suspected AI use, and pedagogical concerns. In line with qualitative research principles, markers were treated as expert informants whose professional judgement provided contextual insight into artefactual patterns observed in the student submissions (Oates 2006).

Triangulation was achieved by comparing independently identified artefactual features with marker observations. Convergence between these data sources strengthened analytic claims, while divergence prompted further interpretive scrutiny rather than exclusion.

Although the study was interpretive rather than quantitative in orientation, indicative prevalence patterns were documented during coding to distinguish isolated anomalies from recurring integrity-related practices. Handwritten submissions emerged as the most common submission irregularity observed across the analysed sample. AI/similarity index indicators were frequent, while EULA non-acceptance and scanned document submissions recurred

across multiple scripts. Submitted question papers and incorrect script uploads were observed in several submissions but occurred less consistently. These patterns were not used for statistical generalisation, but to contextualise the sociotechnical integrity challenges identified in the qualitative analysis.

Data analysis procedures

Data analysis followed a two-stage, iterative thematic process consistent with qualitative analytic frameworks (Braun and Clarke 2006).

In the first stage, student submissions were examined inductively to identify recurring artefactual patterns. Initial codes included AI-shaped graphical representations, templated procedural calculations, handwritten or scanned submissions, EULA non-acceptance, and duplicated formatting structures. This stage focused on identifying *what* patterns appeared across submissions.

In the second stage, artefact codes were integrated with marker reflections to develop higher-order themes. Marker observations were used to interpret *why* particular artefacts were pedagogically problematic and to distinguish isolated errors from systemic patterns. This analytic integration aligns with sociotechnical approaches that emphasise interpretation across social and technical dimensions rather than isolated variables (Mumford 2006).

Ethical considerations

Ethical clearance for the study was obtained from the institutional research ethics committee (Ref #: 7980_2025_PRC_REC_31). All student and marker data were anonymised prior to analysis. The study reports only aggregated patterns and illustrative examples, and no attempt was made to identify or sanction individual students. Given the sensitivity of academic integrity research, findings are framed analytically rather than punitively, consistent with ethical guidance in qualitative educational research (Bretag 2016).

Methodological limitations

Several limitations must be acknowledged. First, the study does not claim certainty in identifying AI-generated work, as current detection tools and analytic methods cannot reliably distinguish AI use from other forms of external assistance (Luo 2024). Second, the findings are context-specific, reflecting a single module within one ODeL institution, and are not intended to be generalisable beyond similar settings.

Finally, the analysis is based on assessment artefacts and marker observations; student motivations and intentions were not directly examined and therefore cannot be inferred.

Despite these limitations, the methodology provides a robust account of how integrity challenges emerge at the intersection of technology, policy, and equity in large-scale ODeL assessment contexts.

FINDINGS

This section reports the key findings from a qualitative analysis of assessment submissions in a large undergraduate data handling and probability module offered through an open distance e-learning (ODeL) institution in South Africa. The analysis draws on artefactual evidence embedded in student submissions, including graphical representations and computational formats inconsistent with prescribed assessment tools and expected student workflow. In specific cases, these artifacts provide strong empirical indicators of AI-shaped work.

Rather than attempting to quantify the prevalence of AI use or claim absolute certainty of detection across the entire cohort, the analysis focuses on recurring patterns of practice and their interaction with institutional policies, submission systems, and equity-driven accommodations. Guided by a sociotechnical systems perspective, the findings illustrate how academic integrity challenges emerged through the interaction of technical infrastructures, institutional practices, and student agency. Indicative prevalence patterns observed during coding further contextualized these findings.

AI-generated graphical artifacts as indicators of integrity breakdown

A prominent pattern in the assessment submissions was the presence of graphical representations that showed strong indicators of AI-generated production. These artifacts appeared across multiple forms of statistical representation, including bar charts, pie charts, histograms, box-and-whisker plots, pictographs, and stem-and-leaf diagrams. While generative AI use cannot be definitively proven in every individual instance, the consistency, structure, and recurrence of these anomalies across many scripts suggest AI-shaped work rather than isolated student error.

This visual polish frequently coexisted with fundamental statistical inconsistencies when examined against the underlying datasets. Bar charts representing categorical variables often displayed touching bars, a convention associated with continuous data rather than nominal categories, or arranged categories alphabetically rather than by frequency (Figure 1). **Marker 1** explicitly noted “persistent errors in tallying and missing axis labels despite graphs looking professionally produced,” while **Marker 3** observed that “correct labeling of the horizontal and vertical axes was very challenging for many students.”

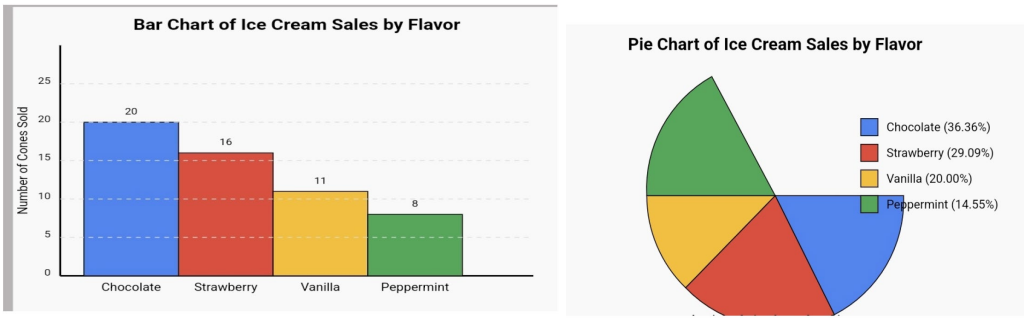

Figure 1: Examples of AI-generated bar and pie charts that show visual sophistication alongside statistical inconsistencies

Pictographs provided particularly strong artefactual indicators. Several submissions included stylized or composite iconography that would be difficult to produce using the prescribed tools. Fractional icons were often represented inconsistently, with only part of a symbol altered to indicate proportionality. Even where a key was provided, the number of icons frequently failed to correspond to the stated frequencies (Figure 2). **Marker 3** reported that "most students did not attempt the pictograph correctly," while **Marker 1** similarly highlighted "incorrect or inconsistent pictogram icons."

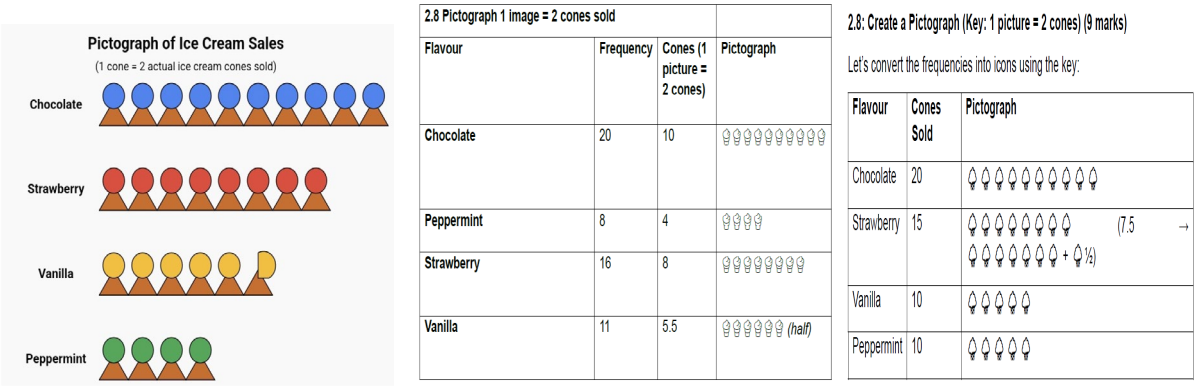

Figure 2: AI-shaped pictographs with inconsistent icon proportionality and mismatches between icons, keys, and frequencies

More compellingly, some stem-and-leaf plots exhibited structural incoherence inconsistent with manual construction. In one submission (Figure 3), repeated stem values appeared within a single line (e.g., "2 2 | 5 6 8"), duplicated rows shared identical formatting, and an empty stem ("7 |") was presented without explanation. Such features violate standard conventions and are unlikely to result from independent student work. Instead, they suggest verbatim reproduction of a malformed template that was not checked against the dataset.

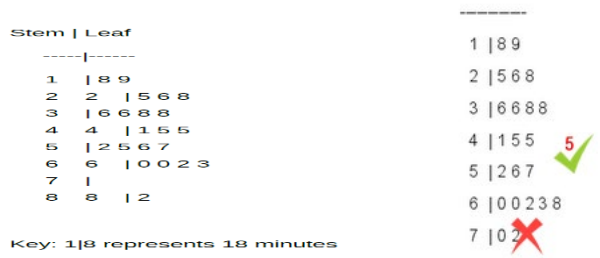

Figure 3: AI-shaped stem-and-leaf plot exhibiting repeated stems, duplicated rows, and unexplained empty stems

Similar inconsistencies appeared in histograms and box-and-whisker plots, where distributions, quartiles, and whiskers did not align with the underlying datasets despite confident interpretations (Figure 4 and Figure 5).

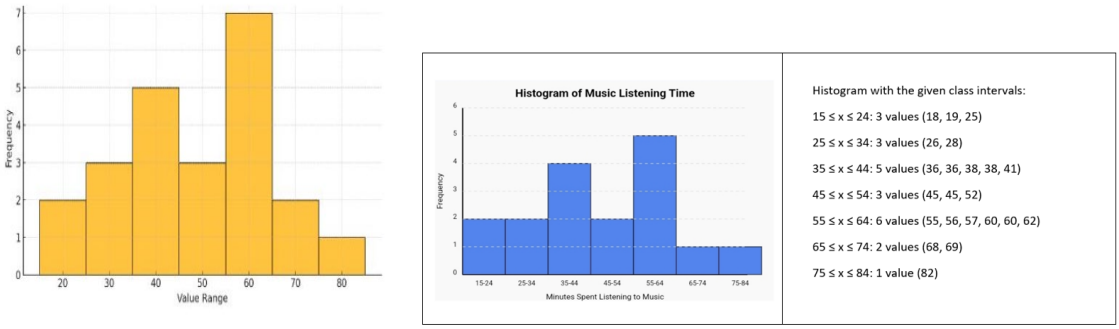

Figure 4: AI-shaped histogram showing smooth visual distribution despite incorrect class intervals and frequencies.

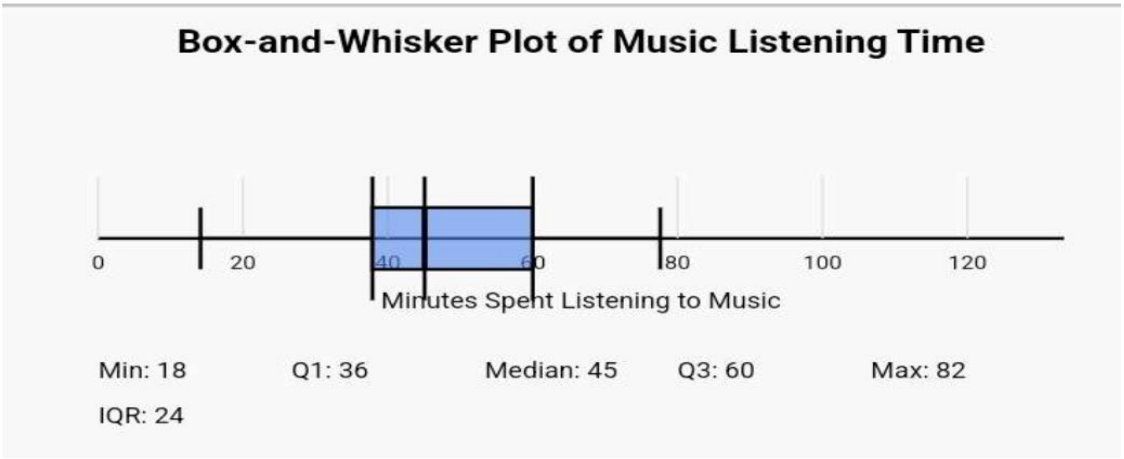

Figure 5: Box-and-whisker plot with quartile and whisker placements inconsistent with the underlying dataset.

From a sociotechnical perspective, these artefacts illustrate reliance on the visual authority of AI-generated outputs, where correctness is inferred from appearance rather than statistical validation.

Circumvention of integrity systems through submission format practices

Beyond AI-shaped content, a second pattern of integrity-related concern related to how assessments were submitted rather than what they contained. Numerous submissions were categorised as handwritten, scanned documents, wrong scripts, or uploaded question papers. Although these practices contravened institutional requirements for typed, text-selectable submissions, they cannot be interpreted solely as deliberate attempts to evade integrity checks.

Handwritten submissions were the most frequently coded category, often reflecting limited device access, reliance on mobile phones, and uneven digital skills rather than advanced conceptual engagement. Closely related were scanned submissions consisting of handwritten pages converted into PDFs or embedded images. These formats disabled text-based plagiarism detection and undermined assessment consistency. A smaller subset of cases involved students uploading the wrong script or the question paper itself, pointing to file management difficulties and gaps in digital literacy rather than intentional misconduct.

A closer examination revealed variation in handwritten submissions. Some displayed procedural variation and visible reasoning characteristic of authentic engagement, while others reproduced AI-shaped or externally generated responses manually through standardized explanations and copied representational structures. Handwriting, therefore, did not necessarily indicate authenticity or absence of AI mediation, but reflected adaptive navigation of integrity systems within constrained digital conditions.

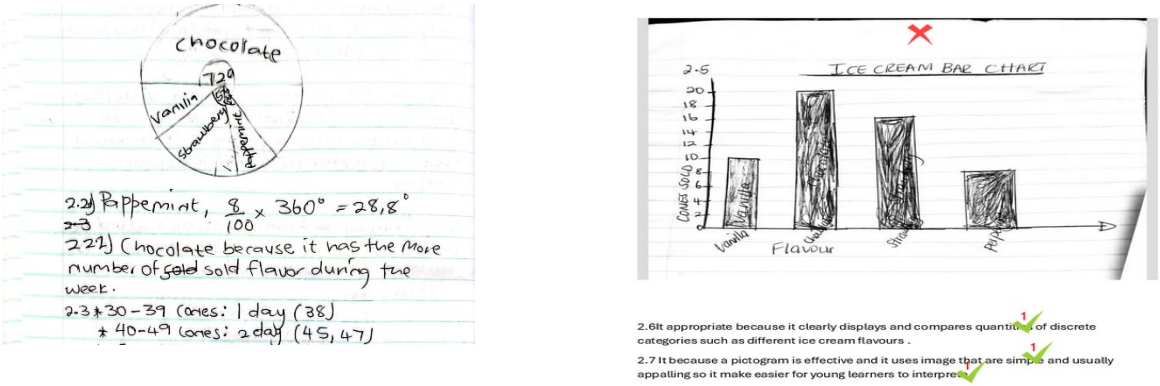

Figure 6: Handwritten and scanned assessment submissions that bypass text-based similarity checking and limit consistent integrity evaluation

Taken together, these practices reveal how technical and administrative constraints can weaken integrity mechanisms, producing outcomes that resemble misconduct without necessarily constituting it. From a sociotechnical perspective, they highlight misalignments between institutional expectations, technological infrastructures, and students' lived realities. From an

STS perspective, the coexistence of authentic handwritten reasoning and manually reproduced AI-shaped responses within the same submission category illustrates how students navigate integrity systems adaptively within materially constrained assessment environments.

Policy–practice gaps in integrity enforcement: EULA non-acceptance

A further integrity-related pattern concerned students' failure to accept the End-User Licence Agreement (EULA) required for similarity checking through Turnitin (Figure 7). When the EULA was not accepted, no similarity report was generated, leaving markers unable to evaluate originality or substantiate concerns about AI-shaped or plagiarised content.

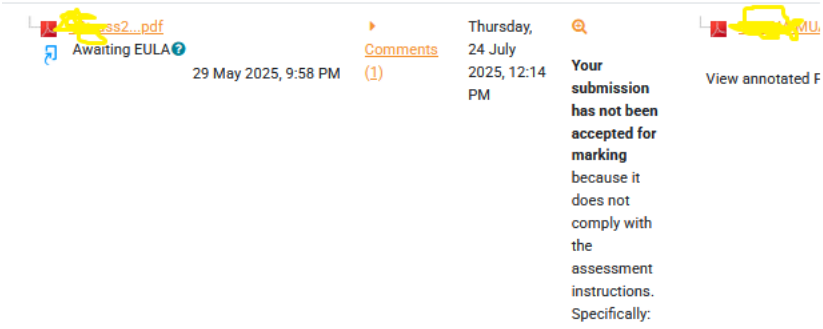

Figure 7: illustrates how EULA non-acceptance at the submission stage prevents automated integrity checks from being applied

Markers frequently expressed uncertainty in these cases. **Marker 2** noted that even when responses appeared atypical, "it was difficult to act without a similarity report." EULA non-acceptance cannot be assumed to be deliberate; students may overlook system prompts due to device limitations, connectivity disruptions, or unfamiliarity with submission interfaces.

Nevertheless, EULA non-acceptance creates exploitable loopholes. Some submissions were exempt from integrity checks while others were not, resulting in uneven enforcement within the same module. From a sociotechnical perspective, this represents a misalignment between policy intent and technical implementation, where equity-driven flexibility dilutes enforcement in the absence of equivalent integrity pathways.

Procedural correctness without conceptual understanding

A further pattern reinforcing the integrity concerns was the submission of procedurally correct but conceptually unintegrated calculations. Several scripts contained relative frequency calculations retaining LaTeX-style syntax (e.g., \frac{20}{55} × 100 = 36.36%), a format neither taught nor required in the module (Figure 8). These calculations were submitted without an accompanying explanation, indicating that while the numerical results were correct, the reasoning resided with the AI tool rather than the student.

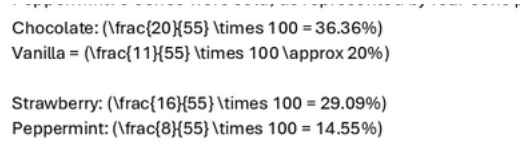

Figure 8: AI-shaped relative frequency calculations containing LaTeX syntax pasted verbatim into student submissions, without accompanying interpretation or contextual explanation

Marker 2 observed that "students have managed to produce correct answers, but it is evident that they generated the answers without modification or effort," while **Marker 1** similarly reported "submitting AI-generated responses that didn't meet the requirements."

Comparable patterns appeared in quartile and dispersion calculations, where identically formatted step-by-step solutions recurred across multiple submissions (Figure 9). This uniformity, combined with earlier conceptual confusion, suggests reliance on AI-generated procedural templates rather than independent reasoning.

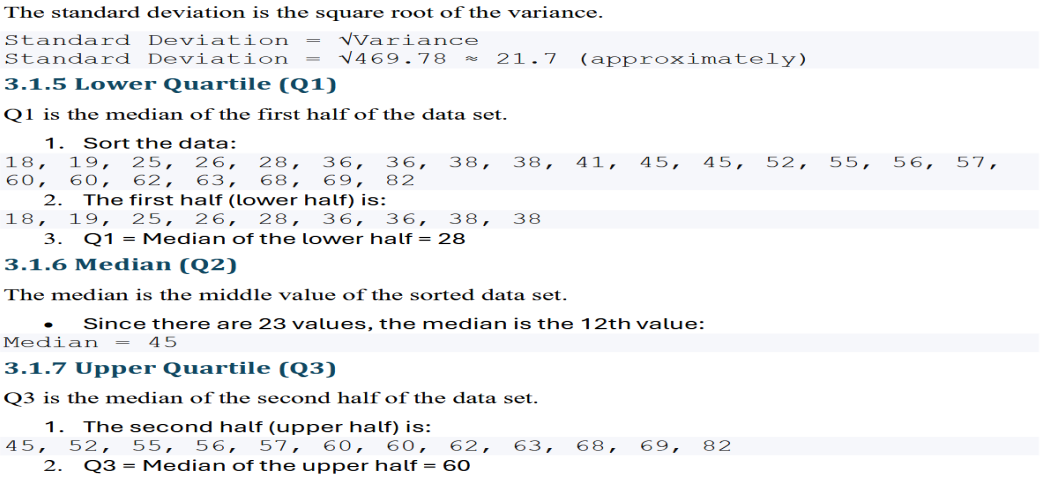

Figure 9: Identically formatted quartile calculations appearing across multiple student scripts, indicating reliance on AI-generated procedural templates rather than independent reasoning

From a learning perspective, these artefacts exemplify an illusion of mastery, where surface correctness masks a lack of conceptual understanding. **Marker 2** explicitly linked this pattern to pedagogical risk, warning that uncritical AI use "moves us back to rote learning rather than the construction of knowledge."

Synthesising integrity challenges in an AI-mediated ODeL environment

Taken together, the findings demonstrate that academic integrity challenges in this ODeL context cannot be attributed to a single factor such as student dishonesty or the availability of

generative AI tools. Instead, they emerge from the interaction of assessment design, technological systems, institutional policies, and students' material conditions of access.

AI-shaped graphical artefacts and procedurally correct but conceptually unintegrated calculations reveal how generative tools can produce polished outputs that substitute for cognitive engagement. Submission-format irregularities and EULA non-acceptance further expose systemic vulnerabilities that weaken integrity enforcement, even in the absence of deliberate evasion.

Collectively, these patterns reinforce a central insight of Sociotechnical Systems Theory: academic integrity breaches are co-produced by social practices and technical arrangements. In resource-constrained ODeL environments, binary responses to AI are unlikely to succeed. Without alignment between policy, technology, and pedagogy, AI risks amplifying existing inequities while undermining assessment validity.

DISCUSSION

Interpreting AI use and academic integrity in an ODEL context

This section interprets the findings through the lens of Sociotechnical Systems Theory (STS) and situates them within existing scholarship on AI, academic integrity, and equity in Global South higher education. Rather than treating AI misuse as an individual ethical failure, the discussion demonstrates how integrity challenges emerge from systemic misalignments between technology, policy, pedagogy, and student access conditions.

AI misuse as a sociotechnical phenomenon, not individual misconduct

The findings indicate that AI-related integrity challenges in the analysed assessment cannot be adequately explained as isolated instances of student dishonesty. Instead, they are best understood as sociotechnical outcomes produced by the interaction of technical systems, institutional practices, and student circumstances. This aligns directly with Sociotechnical Systems Theory, which conceptualises organisational outcomes as the product of interdependent social and technical subsystems rather than individual actions alone (Trist and Bamforth 1951; Mumford 2006).

While some submissions contained strong indicators of AI-generated work, others reflected adaptive practices shaped by structural constraints, including limited device access, reliance on mobile phones, and uneven digital skills. Handwritten submissions ranged from authentic reasoning to manually reproduced AI-shaped responses, illustrating how students navigated integrity systems within constrained sociotechnical conditions. Similar patterns have

been documented in Global South ODeL contexts, where technological access and digital literacy significantly mediate assessment participation (Sullivan et al. 2023; Howell 2024).

From an STS perspective, these practices represent *variance* introduced by misaligned systems: rigid technical requirements interacting with unequal access conditions. The findings therefore reinforce arguments that integrity breaches in digitally mediated assessments are frequently systemic rather than moral in nature (Luo 2024; Owan et al. 2023a). This supports calls to shift integrity discourse away from punitive individualisation toward systemic redesign.

Equity-driven policy adaptations and the integrity paradox

A central tension revealed by the findings concerns the unintended consequences of equity-driven institutional adaptations. Although institutional policy mandates typed submissions and similarity checking through Turnitin, flexibility was exercised in practice to accommodate students with limited digital access. Similar accommodation strategies have been reported across ODeL institutions seeking to balance inclusion with assessment validity (Chen and Lai 2024; Howell 2024).

However, consistent with STS concepts of policy–practice misalignment, these accommodations weakened the consistent application of integrity mechanisms. Submissions exempt from similarity checking created uneven accountability across the cohort, placing markers in ethically and professionally difficult positions. This finding echoes concerns raised by Lawrence-Slater (2001) and Luo (2024) that integrity systems relying on technical enforcement fail when social realities undermine their implementation.

The integrity paradox observed here, where inclusion-driven flexibility simultaneously erodes enforcement, demonstrates the limits of binary policy responses. STS highlights those ethical outcomes require joint optimisation of social and technical systems (Mumford 2006). The findings thus extend existing literature by showing how equity interventions, when not systemically integrated, can unintentionally destabilise integrity governance.

AI-generated artefacts and the illusion of competence

One of the most pedagogically significant findings relates to the emergence of AI-generated artefacts that create an illusion of competence. Visually polished graphs and procedurally correct calculations masked fundamental statistical misunderstandings, allowing surface correctness to substitute for conceptual engagement. This directly reflects concerns raised in the literature regarding AI-enabled "illusion of mastery," where learners mistake AI-produced outputs for evidence of understanding (Crompton and Burke 2023; Rudolph et al. 2023).

Markers' observations that students submitted "correct answers without modification or exertion of effort" resonate strongly with prior studies warning that generative AI can facilitate performance without learning (Luo 2024). In statistics education specifically, where interpretation and validation are central, AI-generated representations can undermine assessment validity if reasoning is not made explicit (Swiecki et al. 2022).

The identification of procedurally correct but conceptually unintegrated calculations further reinforces concerns about surface-level learning. The presence of LaTeX-formatted expressions and identical step-by-step computational templates, neither taught nor required in the module, provides strong empirical evidence of AI-mediated procedural outsourcing. Similar patterns have been noted in studies where students use AI to bypass cognitive effort while preserving formal correctness (Rudolph et al. 2023; Luo 2024). From an STS perspective, this represents a failure of variance control: the assessment system evaluates outputs rather than reasoning processes, allowing deviations in learning to go undetected. This finding supports calls in the literature for assessments that prioritize explanation, critique, and justification over computational accuracy alone (Owan et al. 2023a).

Digital literacy, paraphrasing tools, and the limits of detection

The findings also highlight the limitations of detection-centric integrity strategies. While some AI use was identifiable through stylistic and representational anomalies, other submissions employed paraphrasing tools or hybrid strategies that evaded detection. This reinforces arguments that absence of detectable markers does not equate to absence of AI use (Luo 2024).

Markers' reliance on professional judgement aligns with research recognising educators as critical integrity agents rather than passive users of surveillance technologies (Evangelista, 2025). However, as this study shows, their agency is constrained when technical systems fail, such as in cases of EULA non-acceptance or non-text submissions. This finding strengthens critiques of detection-only approaches and supports pedagogical models that foreground transparency, reflection, and justification.

Situating the findings within the Global South

While AI-related integrity challenges are evident globally, this study demonstrates how they are intensified in Global South ODeL contexts by unequal access, connectivity constraints, and uneven digital literacies. These findings align with scholarship cautioning against importing Global North integrity models into resource-constrained environments without contextual adaptation (Sullivan et al. 2023; Howell 2024).

The contribution of this study lies in its large-scale empirical grounding, informed by marker observations, which reveals how AI, equity, and institutional practice intersect in practice. By foregrounding sociotechnical interactions rather than individual blame, the study advances a context-sensitive framework for understanding AI ethics in higher education.

RECOMMENDATIONS

Toward ethical and equitable AI governance in ODeL

Drawing on the findings, the sociotechnical framing, and the practical insights of markers, several interrelated recommendations are proposed. These recommendations are not positioned as prescriptive solutions, but as principles to guide more coherent and context-sensitive responses to AI use in ODeL environments.

Consistent but Inclusive integrity pathways

Institutions should aim for consistency in the application of academic integrity mechanisms while recognizing structural inequalities among students. Rather than exempting certain assessments or cohorts from similarity checking, a practice that leaves markers without crucial tools, alternative but equivalent integrity pathways should be developed for students with limited technological access. Students unable to access standard digital submission pathways should be supported through equivalent integrity mechanisms such as supervised digital submissions or scaffolded assessments requiring explanation of reasoning processes. Such approaches preserve integrity while reducing exclusion in resource-constrained contexts.

Strengthening digital and AI literacy for students and staff

The findings indicate that many integrity challenges stem not from deliberate misconduct but from limited digital and AI literacy. Institutions should therefore embed explicit training on digital submission practices, responsible AI use, and academic integrity expectations within undergraduate programmes, particularly in the first year. Training should include practical guidance on verifying AI-generated outputs and on the responsible use of generative tools. Equally important is professional development for academic staff. Lecturers and markers require support to interpret AI-shaped student work, apply integrity policies consistently in complex situations, and design assessments that foreground reasoning and interpretation. Without institutional upskilling, policy–practice gaps are likely to persist, leaving markers under-supported in their frontline role as integrity arbiters.

Assessment design as a site of ethical intervention

Assessment design emerged as a critical leverage point within the sociotechnical system. Product-oriented tasks, such as generating graphs, definitions, or performing routine calculations, were particularly vulnerable to AI substitution. Markers' experiences confirm that such tasks often yield superficially correct but conceptually hollow submissions. To mitigate this risk, assessments should prioritise tasks that require explanation, justification, critique, and contextual interpretation.

In statistics education specifically, this includes asking students to evaluate flawed representations, explain methodological choices, or interpret data in relation to real-world contexts. Such designs do not prohibit AI use outright, but reposition AI as an object of critique or a limited support tool rather than a replacement for cognitive engagement, thereby aligning assessment with the pedagogical goal of knowledge construction emphasised by markers.

Aligning equity commitments with integrity governance

The study highlights a tension between equity-driven accommodations and the consistent enforcement of integrity standards. Addressing this tension requires institutional coordination rather than ad hoc compromises at module level. Equity considerations should be integrated into integrity governance frameworks from the outset, ensuring that all students are subject to comparable standards while receiving differentiated forms of support.

In addition, institutional policymakers should recognise that some forms of AI reliance may emerge not only from deliberate attempts to circumvent assessment requirements, but also from students' struggles with conceptual understanding, inadequate academic support, and uneven preparation for digitally mediated learning environments. Ethical AI governance therefore requires more than surveillance and enforcement mechanisms; it also requires investment in responsive teaching practices, academic support structures, and assessment guidance that reduce students' dependence on uncritical AI use as a compensatory strategy.
This alignment may include investment in shared digital infrastructure, clearer institutional guidance on acceptable AI use, and formal channels for integrating marker feedback into policy reviews, as their daily work reveals the real-world consequences of policy gaps. Without such alignment, well-intentioned equity measures risk inadvertently weakening academic standards and creating ambiguity for both students and the staff tasked with evaluating their work.

CONCLUSION

This study examined AI use and misuse in assessment within a South African ODeL context, drawing on a large-scale analysis of undergraduate submissions in a data handling and probability module, triangulated with the professional observations of module markers. Framed by Sociotechnical Systems Theory, the findings demonstrate that academic integrity challenges associated with AI cannot be attributed solely to individual student behaviour. Instead, they emerge from the interaction between technological systems, institutional policies, and the lived realities of students navigating structural inequalities.

The findings revealed integrity-related practices including AI-generated work, handwritten or scanned submissions that bypassed similarity detection, and non-acceptance of End User Licence Agreements. These practices were often intertwined with uneven digital access, limited technological skills, and inconsistent implementation of integrity mechanisms, illustrating how AI misuse functioned both as a shortcut and as a compensatory strategy within constrained learning environments.

By foregrounding the Global South ODeL context, and by valuing the practitioner lens of markers, this article contributes empirical evidence to international debates on AI ethics in higher education. It challenges binary framings of AI as either a threat to be prohibited or a tool to be embraced uncritically. Instead, it argues for governance approaches that balance integrity with equity, recognising that ethical AI integration depends as much on institutional coherence and support structures as on student compliance.

Ultimately, the study suggests that without inclusive and consistently applied integrity frameworks, AI risks amplifying existing inequalities rather than enhancing learning. For AI to contribute meaningfully to higher education in resource-constrained contexts, institutions must invest in aligned policies, assessment redesign, and digital capacity-building. This capacity-building must extend to markers and lecturers, empowering them with the literacy and institutional support to navigate AI-mediated assessment with confidence and pedagogical purpose. Such efforts are essential for safeguarding ethical and equitable educational practice in AI-mediated environments.

More broadly, the study contributes to ongoing Global South debates about the possibilities and limitations of AI in higher education by demonstrating that ethical AI integration depends not only on technological innovation, but also on socially responsive governance, assessment redesign, and institutional sensitivity to structural inequality.

REFERENCES

Alfaleh, M. 2026. “Sustainable AI-Driven Assessment in Higher Education: A Systematic Review of Fairness, Transparency, Pedagogical Innovation, and Governance.” *Sustainability* 18 (2): 785. https://doi.org/10.3390/su18020785.

Bitzer, E., and H. Menkveld. 2023. “Academic Integrity in South African Higher Education: Challenges and Opportunities.” *South African Journal of Higher Education* 37 (2): 1–18.

Boughey, C., and S. McKenna. 2021. *Understanding Higher Education: Alternative Perspectives*. African Minds.

Braun, V., and V. Clarke. 2006. “Using Thematic Analysis in Psychology.” *Qualitative Research in Psychology* 3 (2): 77–101. https://doi.org/10.1191/1478088706qp063oa

Bretag, T., ed. 2016. *Handbook of Academic Integrity*. Springer. https://link.springer.com/referencework/10.1007/978-981-287-098-8

Chen, G., and N. Lai. 2024. “Enhancing Interdisciplinary English Learning: A Systematic Review of the HyFlex Model’s Strategies, Challenges, and Implications.” *Asian Journal of Education and Social Studies* 50 (6): 1–10. DOI: 10.9734/ajess/2024/v50i61388

Crompton, H., and D. Burke. 2023. “Artificial Intelligence in Higher Education: The State of the Field.” *International Journal of Educational Technology in Higher Education* 20 (1): 22. https://doi.org/10.1186/s41239-023-00392-8

Czerniewicz, L., N. Agherdien, J. Badenhorst, D. Belluigi, T. Chambers, M. Chili, M. de Villiers, et al. 2020. “A Wake-Up Call: Equity, Inequality and Covid-19 Emergency Remote Teaching and Learning.” *Postdigital Science and Education* 2 (3): 946–967. https://doi.org/10.1007/s42438-020-00187-4

Dabis, A., and C. Csáki. 2024. “AI and Ethics: Investigating the First Policy Responses of Higher Education Institutions to the Challenge of Generative AI.” *Humanities and Social Sciences Communications* 11: 1006. DOI: 10.1057/s41599-024-03526-z.

Department of Higher Education and Training. 2014. *Policy for the Provision of Distance Education in South African Universities in the Context of an Integrated Post-School System*. Government Notice 535, Government Gazette 37811, 7 July. Pretoria: Government Printer

Eaton, S. E. 2021. *Plagiarism in Higher Education: Tackling Tough Topics in Academic Integrity*. ABC-CLIO.

Evangelista, E. D. L. 2025. “Ensuring Academic Integrity in the Age of ChatGPT: Rethinking Exam Design, Assessment Strategies, and Ethical AI Policies in Higher Education.” *Contemporary Educational Technology* 17 (1): ep559. https://doi.org/10.30935/cedtech/15775

Gaya, H. J., and E. E. Smith. 2016. “Developing a Qualitative Single Case Study in the Strategic Management Realm: An Appropriate Research Design?” *International Journal of Business Management and Economic Research* 7 (2): 529–538.

Guadu, A., D. Dibekulu, and A. W. Menberu. 2026. “Academic Integrity in the Age of AI: Challenges and Strategies for African Higher Education.” *Discover Education* 5: 21. DOI: 10.1007/s44217-025-01031-2

Howell, B. 2024. *Regulating Artificial Intelligence in a World of Uncertainty*. American Enterprise Institute.

Lawrence-Slater, M. 2001. “Assessing Hybrid Classes in a Flexible Learning Environment.” In *31st Annual Frontiers in Education Conference: Impact on Engineering and Science Education. Conference Proceedings*, vol. 2, F3C-1–F3C-6. IEEE

Luo, J. 2024. “A Critical Review of GenAI Policies in Higher Education Assessment: A Call to Reconsider the ‘Originality’ of Students’ Work.” *Assessment & Evaluation in Higher Education* 49 (5): 651–664. https://doi.org/10.1080/02602938.2024.2309963

Maguire, M., and B. Delahunt. 2017. “Doing a Thematic Analysis: A Practical, Step-by-Step Guide for Learning and Teaching Scholars.” *All Ireland Journal of Higher Education* 9 (3): 3351–33514.

McCabe, D. L., K. D. Butterfield, and L. K. Treviño. 2012. *Cheating in College: Why Students Do It and What Educators Can Do About It*. Johns Hopkins University Press.

Mumford, E. 2006. "The Story of Socio-Technical Design: Reflections on Its Successes, Failures and Potential." *Information Systems Journal* 16 (4): 317–342. https://doi.org/10.1111/j.1365-2575.2006.00221.x

Oates, B. J. 2006. *Researching Information Systems and Computing*. Sage Publications.

Owan, V. J., K. B. Abang, D. O. Idika, E. O. Etta, and B. A. Bassey. 2023a. "Exploring the Potential of Artificial Intelligence Tools in Educational Measurement and Assessment." *Eurasia Journal of Mathematics, Science and Technology Education* 19 (8): em2307. https://doi.org/10.29333/ejmste/13428

Owan, V. J., J. A. Ekpenyong, and M. E. Asuquo. 2023b. "Artificial Intelligence in Higher Education: Emerging Practices, Ethical Challenges, and Policy Implications." *Education and Information Technologies* 28: 11831–11852. https://doi.org/10.1007/s10639-023-11674-6.

Oyedemi, T., and S. Mogano. 2018. "The Digitally Disadvantaged: Access to Digital Communication Technologies among First Year Students at a Rural South African University." *Africa Education Review* 15 (1): 175–191. DOI: 10.1080/18146627.2016.1264866.

Rudolph, J., S. Tan, and S. Tan. 2023. "ChatGPT: Bullshit Spewer or the End of Traditional Assessments in Higher Education?" *Journal of Applied Learning and Teaching* 6 (1): 342–363. https://doi.org/10.37074/jalt.2023.6.1.9

Sullivan, M., A. Kelly, and P. McLaughlan. 2023. "ChatGPT in Higher Education: Considerations for Academic Integrity and Student Learning." *Journal of Applied Learning and Teaching* 6 (1): 31–40. DOI: 10.37074/jalt.2023.6.1.17.

Sutherland-Smith, W. 2023. "Generative AI and Academic Integrity: Rethinking Detection and Prevention." *International Journal for Educational Integrity* 19 (1): 1–15.

Swart, O., and C. C. Shuttleworth. 2021. "The New Face of Alternative Assessment in Accounting Sciences: Technology as an Anthropomorphic Stakeholder." *South African Journal of Higher Education* 35 (3): 200–219. https://doi.org/10.20853/35-3-3900.

Swartz, R. 2023. "Academic Integrity in the Age of AI: Equity Considerations for South African Universities." *Critical Studies in Teaching and Learning* 11 (1): 45–62.

Swartz, R., M. Ivancheva, L. Czerniewicz, and N. P. Morris. 2019. "Between a Rock and a Hard Place: Dilemmas Regarding the Purpose of Public Universities in South Africa." *Higher Education* 77 (4): 567–583. https://doi.org/10.1007/s10734-018-0291-9

Swiecki, Z., H. Khosravi, G. Chen, R. Martinez-Maldonado, J. M. Lodge, S. Milligan, N. Selwyn, and D. Gašević. 2022. "Assessment in the Age of Artificial Intelligence." *Computers and Education: Artificial Intelligence* 3: 100075. https://doi.org/10.1016/j.caeai.2022.100075

Trist, E. L., and K. W. Bamforth. 1951. "Some Social and Psychological Consequences of the Longwall Method of Coal-Getting." *Human Relations* 4 (1): 3–38. https://doi.org/10.1177/001872675100400101

Walters, P., and R. Kop. 2009. "Heidegger, Digital Technology, and Postmodern Education: From Being in Cyberspace to Meeting on MySpace." *Bulletin of Science, Technology & Society* 29 (4): 278–286. DOI: 10.1177/0270467609336305.