*FRONTIERS INTERNATIONAL JOURNAL OF INTERDISCIPLINARY STUDIES (FIJIS)*
*ISSN: 3080-1346 (Online); 3080-1338 (Print)*
*Vol 1. No 1. February, 2026*
*https://fijis-uba.com*

# AI and Exam Integrity in Cameroon's Bilingual GCE: A Qualitative Conceptual–Analytical Study

**Roger Njebuh Nzembayie**
*The University of Bamenda, Cameroon*
*&*
**Visi Sumbom Tubuo**
*The University of Bamenda, Cameroon*

## Abstract

*The rapid spread of artificial intelligence (AI) tools, including text generators and translation applications, presents new challenges to the authenticity of high-stakes examinations. Cameroon's bilingual General Certificate of Education (GCE) is particularly exposed due to its dual-language structure and growing student access to AI-assisted writing. This qualitative, conceptual–analytical study examined risks of AI-assisted malpractice using three data sources: official GCE policy documents, written reflections from ten experienced secondary-school teachers and examiners, and an analysis of selected A-Level French examination papers from 2019 to 2025. Data were analysed thematically with a focus on patterns of AI susceptibility, linguistic irregularities, and vulnerabilities in question design. Findings indicate three main risk patterns: (1) automated translation producing near-perfect responses inconsistent with candidates' general language performance; (2) AI-generated essays enabling structurally coherent but cognitively shallow scripts; and (3) exam items relying on general knowledge showing high susceptibility to machine-generated answers. Analysis of 224 scripts revealed several cases where translation scores exceeded 90% while essays in the same paper fell below 10%, suggesting potential AI involvement. Existing regulations did not explicitly address AI-assisted authorship, and invigilation systems showed limited capacity to detect covert digital tools. AI tools create identifiable authenticity gaps in Cameroon's bilingual GCE by enabling discrepancies between candidates' productive skills and their written outputs. Current safeguards insufficiently address these risks, indicating a need for updated policies, examiner training, and improved detection capacity to protect examination validity.*



*Corresponding Author:* ***Visi Sumbom Tubuo***
*visisumbom@gmail.com*
*The University of Bamenda, Cameroon*

*FRONTIERS INTERNATIONAL JOURNAL OF INTERDISCIPLINARY STUDIES (FIJIS)*
*ISSN: 3080-1346 (Online); 3080-1338 (Print)*
*Vol 1. No 1. February, 2026*
*https://fijis-uba.com*

**Keywords:** artificial intelligence; academic integrity; examination authenticity; Cameroon GCE; bilingual assessment; qualitative analysis.

## 1. Introduction

Artificial intelligence (AI) tools such as text generators, paraphrasing systems, and automated translators are increasingly influencing educational processes, particularly assessment. According to UNESCO (2023), these technologies can enhance accessibility and support personalised learning but simultaneously introduce risks for academic integrity. Empirical studies show that generative AI can produce essays and translations that resemble authentic human writing yet evade conventional plagiarism-detection systems (Cotton et al., 2023; Lund et al., 2023). Dawson and Sutherland-Smith (2019) further demonstrate that existing integrity frameworks, designed for text-matching or contract-cheating detection, are poorly adapted to algorithmic authorship, thereby challenging the validity of written assessments in high-stakes contexts.

Within Cameroon, the General Certificate of Education (GCE) is a bilingual examination system where candidates complete tasks in English and French. As noted by Tambo (2014), this dual-language configuration introduces additional complexity to maintaining fairness and standardisation. The increasing availability of AI-assisted translation and writing tools raises concerns about whether written responses continue to reflect genuine bilingual competence.

Despite the global attention to AI-assisted malpractice, no empirical studies have examined its manifestations within bilingual secondary examinations such as Cameroon's GCE. Prior literature documents the general risks of generative AI for authenticity (Cotton et al., 2023; Lund et al., 2023; Dawson & Sutherland-Smith, 2019), but does not show how these risks appear in handwritten translation or essay tasks, nor how examiners in bilingual systems experience or detect potential AI-assisted outputs. It is also unknown whether AI-generated or AI-assisted responses create identifiable inconsistencies across different task types within the same examination script.

If AI-assisted content produces written outputs that exceed or contradict candidates' demonstrated linguistic competence, the validity and fairness of high-stakes certification may be affected. Understanding how such discrepancies



*Corresponding Author:* ***Visi Sumbom Tubuo***
*visisumbom@gmail.com*
*The University of Bamenda, Cameroon*

appear in a bilingual examination context is essential for safeguarding the interpretability of assessment outcomes.

Therefore, this study seeks to examine how AI-assisted tools influence the authenticity of written responses in Cameroon's bilingual GCE examinations through the analysis of examiner reflections, policy documents, and selected examination scripts. Consistent with this aim, the study tests the hypothesis that AI-assisted writing and translation tools produce detectable discrepancies between candidates' performance on translation and essay tasks, indicating potential compromises in the authenticity of bilingual examination outputs.

## 2. Problem Statement

The rapid integration of artificial intelligence (AI) tools into education presents an emerging threat to the authenticity of public examinations such as Cameroon's General Certificate of Education (GCE). Applications including ChatGPT, translation software, and automated essay generators enable candidates to produce grammatically polished or perfectly translated texts that often disguise weak linguistic competence. Examiners have observed, for instance, translation passages earning scores above ninety per cent despite the accompanying essays revealing poor syntactic control and limited vocabulary. Such discrepancies suggest the growing influence of AI-generated or AI-assisted work, which compromises the validity of assessment outcomes and obscures genuine learner ability. This challenge is intensified in Cameroon's bilingual context, where translation between English and French provides fertile ground for misuse of machine-generated output. Consequently, the credibility of national certification, fairness in grading, and the principle of academic integrity face increasing erosion under the shadow of AI-driven malpractice.

## 3. Research Questions

The study seeks to explore how artificial intelligence (AI) influences the authenticity of Cameroon's bilingual General Certificate of Education (GCE) examinations and to propose practical measures for safeguarding integrity. Guided by this purpose, the research addresses the following questions:



*Corresponding Author:* ***Visi Sumbom Tubuo***
visisumbom@gmail.com
*The University of Bamenda, Cameroon*

i. How do AI tools such as ChatGPT, translation applications, and essay generators threaten the authenticity and fairness of the bilingual GCE examinations?
ii. In what ways does the design and content of GCE examination questions—particularly those leaning towards general knowledge—create opportunities for AI-assisted cheating?
iii. What strategies can be developed at policy, institutional, and pedagogical levels to mitigate AI-driven malpractice while promoting ethical digital literacy among candidates and teachers?

## 4. Research Objectives

Accordingly, the study aims to identify emerging risks**,** analyse systemic vulnerabilities in exam design and administration, and propose context-sensitive, sustainable strategies that reinforce academic integrity within Cameroon's bilingual assessment framework.

i. To find out why AI tools such as ChatGPT, translation applications, and essay generators threaten the authenticity and fairness of the bilingual GCE examinations.
ii. To identify why the design and content of GCE examination questions—particularly those leaning towards general knowledge—create opportunities for AI-assisted cheating.
iii. To assess the strategies that can be developed at policy, institutional, and pedagogical levels to mitigate AI-driven malpractice while promoting ethical digital literacy among candidates and teachers.

## 5. Significance of the Study

Artificial intelligence (AI) is reshaping the landscape of Cameroon's GCE examinations, posing risks to authenticity and fairness. Tools such as ChatGPT, essay generators, and translation applications enable candidates to produce work that may appear original yet lacks genuine intellectual input, exposing vulnerabilities in question design—especially prompts relying on general knowledge or structured responses.



*Corresponding Author:* ***Visi Sumbom Tubuo***
visisumbom@gmail.com
*The University of Bamenda, Cameroon*

This study demonstrates how traditional safeguards are insufficient to detect AI-assisted malpractice and highlights the need for context-sensitive, multi-level interventions. Recommended strategies include integrating ethical digital literacy into curricula, redesigning tasks to emphasise reflection and creativity, implementing hybrid human–AI monitoring systems, and training examiners to identify AI-specific patterns.

By linking emerging risks, systemic vulnerabilities, and practical safeguards, the research provides actionable insights for policymakers, educators, and researchers. Nationally, these measures are vital for maintaining public confidence in high-stakes assessments, while internationally, Cameroon offers a case study of balancing technological innovation with rigorous standards of academic integrity.

## 6. Literature Review and Theoretical Grounding

### 6.1 Theoretical Background

This study is anchored in the intersecting frameworks of academic integrity theory, digital ethics, and constructivist–sociocultural perspectives on assessment. According to Fishman (2016), academic integrity is grounded in six core values—honesty, trust, fairness, respect, responsibility, and courage—which together define ethical academic behaviour and the authenticity of learner performance. These principles are threatened when artificial intelligence (AI) enables unacknowledged authorship or automated responses that distort genuine competence. Bretag (2018) further contends that maintaining integrity requires a culture of transparency supported by pedagogical design and institutional accountability.

Digital ethics, as emphasised by UNESCO (2023), concerns the responsible creation and use of digital technologies, ensuring that innovation aligns with human dignity, fairness, and social inclusion. Within educational contexts, this perspective demands that AI be harnessed to enhance, rather than erode, ethical learning and equitable assessment. The ethical management of data, authorship, and algorithmic influence therefore becomes central to sustaining the credibility of examinations such as Cameroon's bilingual GCE.

From a constructivist standpoint, learning is viewed as an active, meaning-making process shaped by authentic engagement (Vygotsky, 1978). Sociocultural theories extend this by situating assessment within cultural and



*Corresponding Author:* ***Visi Sumbom Tubuo***
visisumbom@gmail.com
*The University of Bamenda, Cameroon*

linguistic contexts that shape knowledge expression (Timmis et al., 2016). Consequently, authentic assessment should reflect real communicative competence across languages, rather than polished outputs produced by machines. Integrating these perspectives provides a conceptual lens for analysing AI-driven malpractice: academic integrity defines the moral foundation, digital ethics delineates the responsible technological use, and constructivist–sociocultural theory underscores the need for context-sensitive, human-centred assessment in Cameroon's bilingual education system.

### 6.2 Global Evidence

Recent international scholarship reveals that the misuse of artificial intelligence (AI) in assessment has become a pervasive global concern. Across Europe, studies show that students increasingly rely on generative systems to produce essays, translations, and code that escape traditional plagiarism detection. Cotton et al. (2023) report that universities in the United Kingdom and the Netherlands have identified a surge in AI-generated assignments displaying stylistic coherence but lacking conceptual depth, suggesting algorithmic authorship. Similarly, Dawson and Sutherland-Smith (2019) note that integrity policies developed for text-matching tools such as Turnitin are poorly equipped to identify content written by large language models, thus widening the detection gap between authentic and synthetic work.

In North America, the discussion has centred on academic honesty and the ethics of surveillance. Lund et al. (2023) highlight cases in which ChatGPT and similar models have been used to complete take-home exams and online quizzes, prompting institutions to introduce AI-use disclosure statements. However, the growing dependence on online proctoring software has raised new dilemmas of privacy, algorithmic bias, and student anxiety (Susnjak, 2022). The paradox is that technologies introduced to curb misconduct can themselves generate ethical tension, revealing the complexity of regulating digital learning environments.

In Asia, empirical evidence underscores the uneven distribution of resources that shapes AI-related malpractice. In high-tech settings such as South Korea, Hong Kong, and Singapore, sophisticated detection tools and ethical-AI curricula are emerging (Chan & Hu, 2024). By contrast, studies from India and the Philippines indicate that limited institutional capacity and inconsistent internet access leave teachers vulnerable to undetected AI-assisted plagiarism



*Corresponding Author:* ***Visi Sumbom Tubuo***
visisumbom@gmail.com
*The University of Bamenda, Cameroon*

(Sinha & Liu, 2023). These disparities illustrate a widening global divide between high-resource and low-resource contexts in safeguarding assessment integrity.

Collectively, this literature exposes a common pattern: while affluent institutions invest in monitoring and policy innovation, under-resourced systems face escalating risks of AI-driven deception. The findings therefore situate Cameroon's bilingual GCE within a broader international landscape where digital inequality and the rapid diffusion of generative technologies converge to challenge the credibility of public examinations.

### 6.3 African and Local Perspectives

Examination malpractice remains a pervasive challenge across African education systems, shaped by socio-economic, technological, and linguistic factors. In Cameroon, Tambo (2014) observes that the bilingual GCE system, while designed to promote equity between Anglophone and Francophone candidates, "remains vulnerable to translation-based malpractice, with some candidates exploiting linguistic asymmetries to gain an unfair advantage" (p. 78). Similarly, Nwana and Mbangwana (2020) document that the increasing use of smartphones and translation tools has facilitated new forms of examination misconduct, particularly in bilingual examinations, where language-switching and AI translation blur the line between assistance and dishonesty.

Comparable trends are evident in neighbouring countries. In Nigeria, Ojedokun and Idowu (2020) report that students increasingly rely on mobile devices and social media to access unauthorised examination content, arguing that "digital tools have become a key facilitator of academic misconduct in secondary schools" (p. 45). In Ghana, Boateng and Amankwah (2021) emphasise that "covert collaboration among candidates, often mediated by digital technologies, is increasingly shaping examination outcomes" (p. 233). In Kenya, Kithome and Mwangi (2019) highlight that, rural candidates experience both digital exclusion and uneven access to preparatory resources, which "exacerbates inequities in high-stakes assessments" (p. 88).

The digital divide remains a key factor in AI-assisted malpractice. Adu-Gyamfi and Tchombe (2022) note that "urban schools with reliable internet and computing resources are better positioned to monitor and detect AI-assisted malpractice, whereas low-resource or rural schools remain largely unprotected" (p. 59). In Cameroon, this divide affects all Anglophone students studying French,



*Corresponding Author:* ***Visi Sumbom Tubuo***
visisumbom@gmail.com
*The University of Bamenda, Cameroon*

with resource disparities influencing access to AI tools and the potential for misuse (Tambo, 2014, p. 81).

Collectively, these studies indicate that AI-driven academic misconduct in African contexts is not solely a technological problem but is closely entwined with structural inequalities, including language proficiency, digital literacy, and access to resources. These findings justify focused research on Cameroon's bilingual GCE system, where dual-language assessment combined with uneven technological exposure creates a unique environment for AI-assisted cheating, necessitating context-sensitive interventions.

### 6.4 Research Gap

Despite growing global concern over AI-assisted malpractice, few empirical studies have examined how such technologies affect bilingual national examinations like Cameroon's GCE. Existing literature focuses largely on higher education or monolingual contexts, leaving unaddressed how translation-based cheating and generative tools distort authentic bilingual performance. No current framework in Cameroon systematically links AI misuse with linguistic asymmetry, examination supervision, and assessment design. This study therefore fills a critical gap by situating AI-driven cheating within the bilingual GCE's dual-language and resource-diverse realities. Moreover, it introduces practical solutions seldom explored in African assessment research—such as integrating CCTV surveillance with subject-specific flash drives storing examination footage, empowering invigilators to capture short video evidence, and reviving oral and practical testing to authenticate candidate authorship. The study also proposes that future reforms consider emerging AI-detection software for translation or essay generation, a domain currently absent in GCE policy and research.

## 7. Methodology

This study employed a qualitative, conceptual–analytical methodology to examine how AI-assisted tools influence the authenticity of handwritten responses in Cameroon's bilingual GCE. The approach was guided by sociocultural and constructivist perspectives, which view learning and assessment as contextually mediated processes shaped by tools and cultural practices. As Vygotsky (1978) observes, cognitive activity is always mediated,



*Corresponding Author:* ***Visi Sumbom Tubuo***
visisumbom@gmail.com
*The University of Bamenda, Cameroon*

and in contemporary educational settings "new mediational means reconfigure the nature of cognitive activity" (p. 40). AI therefore represents a significant new mediator whose implications for assessment require careful scrutiny.

A qualitative design was appropriate because, as Timmis et al. (2016) argue, assessment in a digital era "requires rethinking long-standing assumptions about what counts as evidence of learning" (p. 454), particularly when algorithmic systems can generate fluent, human-like responses. This necessitates interpretive, meaning-centred inquiry rather than procedural or statistical approaches. The conceptual–analytical dimension allowed the study to interrogate constructs such as authenticity, bilingual competence, and authorship by examining their coherence within the GCE context. This is consistent with Cotton et al.'s (2023) caution that generative AI undermines "the cognitive struggle essential to authentic learning" (p. 4), making conceptual evaluation indispensable.

The approach also resonates with Dawson and Sutherland-Smith's (2019) insight that detecting assisted authorship requires attention to "stylistic and structural consistencies within a script" (p. 716) rather than reliance on surface-level plagiarism checks. Guided by this view, the study integrated policy analysis, examiner reflections, and close script reading to illuminate how AI introduces subtle discontinuities between intended assessment constructs and observable performance. Such an orientation reflects Fishman's (2016) position that academic integrity research must examine both "the values" and "the systems" that underpin ethical behaviour (p. 12), thereby justifying the conceptual–analytical stance adopted in this inquiry.

### 7.1 Research Design

The study adopted a multi-method qualitative approach integrating three data sources:

i. GCE policy and regulatory documents;
ii. written reflections and semi-structured interviews with experienced teachers and examiners; and
iii. systematic analysis of GCE French examination papers and selected scripts from 2019–2025.



*Corresponding Author:* ***Visi Sumbom Tubuo***
visisumbom@gmail.com
*The University of Bamenda, Cameroon*

This triangulated design reflects the view, articulated by Fishman (2016), that academic integrity research must address not only individual behaviour but also the "institutional values and systems that shape ethical action" (p. 12). It also aligns with Dawson and Sutherland-Smith's (2019) argument that detection of assisted authorship requires attention to "assessment design, marker expectations, and the wider integrity environment" (p. 716).

The conceptual–analytical dimension of the methodology sought to examine the alignment between the intended construct (authentic bilingual competence) and the written products candidates produced. Given that AI tools are now capable of generating coherent, well-structured text resembling human output, Lund et al. (2023) observe that markers increasingly face difficulty differentiating "machine-produced writing from that of a highly proficient but struggling student" (p. 3). The research design therefore prioritised close textual analysis, contextualised examiner insights, and policy comparison.

### 7.2 Population and Sampling

The population for this study comprised three interconnected groups central to the administration and assessment of the Cameroon General Certificate of Education (GCE). First, it included GCE Advanced Level examiners and secondary-school teachers responsible for preparing candidates or evaluating responses in French A level spapers. Second, it encompassed official GCE Board documentation—such as examination regulations, malpractice codes, and past papers—which informed both policy and assessment practices. Third, the population extended to handwritten GCE French scripts collected from examination centres across Cameroon.

A purposive sampling strategy was adopted to ensure that both human participants and documentary sources were information-rich and directly relevant to the study's objective. Ten teachers and examiners were selected on the basis of a minimum of five years of professional experience, direct involvement with Papers 745/0745 or their equivalents, and documented familiarity with performance trends that may plausibly be influenced by AI-assisted writing behaviours. This sampling logic reflects Bretag's (2018) assertion that integrity research is strengthened by the perspectives of "those directly involved in teaching and assessing students in contemporary conditions" (p. 2).



*Corresponding Author:* ***Visi Sumbom Tubuo***
visisumbom@gmail.com
*The University of Bamenda, Cameroon*

For the documentary corpus, official GCE Board materials produced between 2021 and 2025 were examined, alongside French Papers 1–3 from the 2019, 2023, and 2025 examination sessions. Draft or provisional documents were deliberately excluded to maintain the accuracy and reliability of the policy record. The script corpus consisted of 224 handwritten French scripts from the 2025 session. Particular analytical attention was paid to scripts showing pronounced discrepancies between translation and essay sections—patterns that examiners commonly identify as potential indicators of AI involvement. Sampling of scripts and participants continued until thematic saturation was achieved, consistent with established qualitative research standards.

### 7.3 Data Collection Procedures

Data collection proceeded across three interrelated strands that together provided a comprehensive understanding of assessment integrity in the GCE context. The first strand involved documentary and policy analysis. Regulatory documents were systematically reviewed to clarify institutional expectations regarding academic honesty, the use of digital devices, and standards of examination conduct. This analysis also sought to identify gaps or ambiguities in existing policy frameworks. The process was guided by UNESCO's (2023) warning that contemporary education systems must "update governance structures and ethical frameworks to remain fit for purpose in an AI-driven era" (p. 14), a recommendation particularly pertinent in high-stakes assessment environments such as the GCE.

The second strand consisted of examiner and teacher reflections gathered through semi-structured interviews and written narratives. These accounts elicited practitioner perspectives on emerging patterns of linguistic inconsistency in candidates' scripts, the challenges of detecting AI-mediated writing, observable shifts in students' writing profiles over recent examination sessions, and potential vulnerabilities in the design of examination questions. Practitioners frequently described stylistic features that "do not correlate with candidates' demonstrated classroom competence" (Examiner Reflection, 2025), noting abrupt variations in vocabulary load, syntactic complexity, and coherence. Their experiential insights resonate with the argument advanced by Ojedokun and Idowu (2020) that teachers often detect misconduct by identifying "incongruities between known ability and displayed performance" (p. 44).



*Corresponding Author:* ***Visi Sumbom Tubuo***
visisumbom@gmail.com
*The University of Bamenda, Cameroon*

Together, these reflections provided an essential practitioner-centred lens for understanding how AI-assisted responses manifest in real examination conditions.

The third strand of data collection involved a systematic analysis of question papers and handwritten scripts using a structured coding sheet developed specifically for this study (Table 1). This instrument provided a transparent and replicable analytic framework, ensuring consistency across all evaluated materials. The coding sheet captured multiple dimensions: the cognitive demand of each question, classified using Bloom's taxonomy; an AI Susceptibility Score, categorising items as low, medium, or high risk for AI-assisted responses; and a suite of authenticity indicators assessing lexical naturalness, syntactic congruence, stylistic stability, coherence between translation and essay sections, and error patterns conventionally expected in Advanced Level French. This multidimensional structure aligns with Dawson and Sutherland-Smith's (2019) observation that evaluating authenticity requires sustained attention to "stylistic and structural consistencies within a script" (p. 717).

To strengthen reliability, two researchers independently applied the coding sheet to all selected papers and scripts. Inter-coder agreement, calculated using Cohen's kappa, yielded $\kappa = 0.82$—interpreted as *substantial agreement*. This high level of concordance indicates that coders applied the framework consistently and that the resulting data are not attributable to chance, thereby enhancing the methodological robustness of the study.

**Table 7.1: Coding Sheet for Analysis of Question Papers and Handwritten Scripts**

| Domain | Indicator / Variable | Guiding Question / Coding Instruction | Scale / Coding Options |
|---|---|---|---|
| **A. Cognitive Demand (Bloom's Taxonomy)** | Knowledge | Does the item require recall of facts or definitions? | 1 = Yes; 0 = No |



*Corresponding Author:* ***Visi Sumbom Tubuo***
visisumbom@gmail.com
*The University of Bamenda, Cameroon*

| Domain | Indicator / Variable | Guiding Question / Coding Instruction | Scale / Coding Options |
|---|---|---|---|
| | Comprehension | Does the task require explanation or paraphrase? | 1 = Yes; 0 = No |
| | Application | Does the response require applying rules or structures? | 1 = Yes; 0 = No |
| | Analysis | Does the task require examining structure or relationships? | 1 = Yes; 0 = No |
| | Evaluation | Does the task require judgement with justification? | 1 = Yes; 0 = No |
| | Creation | Does the task require generating original ideas or arguments? | 1 = Yes; 0 = No |
| **B. AI Susceptibility Score** | Task Specificity | Is the prompt generic enough for AI to answer effectively? | Low / Medium / High |
| | Required Personalisation | Does the task require personal experience, context, or subjective voice? | Low / Medium / High |
| | Algorithmic Vulnerability | To what extent can generative AI | Low / Medium / High |



*Corresponding Author:* ***Visi Sumbom Tubuo***
visisumbom@gmail.com
*The University of Bamenda, Cameroon*

| Domain | Indicator / Variable | Guiding Question / Coding Instruction | Scale / Coding Options |
|---|---|---|---|
| | | complete the task with high accuracy? | |
| | Overall AI Susceptibility | Composite judgement based on indicators above | Low / Medium / High |
| **C. Authenticity Indicators** | Lexical Naturalness | Does word choice reflect candidate's typical vocabulary level? | 1 = Natural; 2 = Moderately elevated; 3 = Unnaturally advanced / AI-like |
| | Syntactic Congruence | Are sentence structures consistent with expected AL French learner competence? | 1 = Congruent; 2 = Mildly anomalous; 3 = Mechanically perfect / unlikely |
| | Stylistic Variation | Does the script display human variation or mechanical uniformity? | 1 = Human-like; 2 = Mixed; 3 = Uniform / AI-like |
| | Translation–Essay Coherence | Do translation and essay sections reflect similar proficiency? | 1 = Coherent; 2 = Mildly divergent; 3 = Strongly divergent |



*Corresponding Author:* ***Visi Sumbom Tubuo***
visisumbom@gmail.com
*The University of Bamenda, Cameroon*

| Domain | Indicator / Variable | Guiding Question / Coding Instruction | Scale / Coding Options |
|---|---|---|---|
| | Error Patterns Expected in AL French | Are typical learner errors present (gender, agreement, tense)? | 1 = Present; 2 = Minimal; 3 = Absent (AI suspicion) |
| **D. Anomaly Detection Patterns** | Mechanical Fluency | Is the writing unusually smooth and error-free for the candidate's level? | Yes / No |
| | Semantic Over-Precision | Does the script contain overly exact lexical items unlikely to be known by candidates? | Yes / No |
| | Construct Irrelevance | Does the response address the prompt but betray unfamiliarity with contextual meaning? | Yes / No |
| | Stylistic Discontinuity Within Script | Are different sections written in noticeably different styles? | None / Mild / Severe |
| | Examiner Flag | Independent examiner judgement of suspicious AI involvement | 0 = No suspicion; 1 = Possible AI; 2 = Strong AI signature |

**Source: Researchers' field work, 2025**



*Corresponding Author:* ***Visi Sumbom Tubuo***
visisumbom@gmail.com
*The University of Bamenda, Cameroon*

All question papers and handwritten scripts were assessed using a structured coding sheet (Table 7.1), which operationalised Bloom's taxonomy, AI susceptibility, and authenticity indicators such as lexical naturalness, syntactic congruence, stylistic variation, translation–essay coherence, and expected error patterns. This allowed systematic, reproducible evaluation of both content and linguistic authenticity. Two researchers independently applied the coding sheet to all materials, yielding substantial inter-coder agreement (Cohen's $\kappa = 0.82$)

### 7.4 Operational Definitions

Several constructs central to the study were operationalised.

**Authenticity:** Defined as the degree to which a script reflects the candidate's own linguistic competence. Luckin et al. (2016) argue that authentic learning entails "a partnership in which learners actively construct meaning rather than outsource thinking" (p. 12).

**AI-Assisted Malpractice:** Defined as any instance where candidates incorporate AI-generated content—via translation applications, generative models, or pre-prepared text—into handwritten scripts.

**AI Susceptibility:** The extent to which an exam task can be effectively completed by generative AI. Cotton, Nguyen, and Pérez (2023) note that some prompt types "lend themselves almost perfectly to algorithmic response generation" (p. 7).

**Cognitive Demand:** Assessed using Bloom's taxonomy, with higher-level cognitive tasks expected to resist automation unless question design is overly generic.

These definitions guided coding, interpretation, and thematic development.

### 7.5 Data Analysis

Data analysis followed **Braun and Clarke's (2006)** six-phase thematic analysis model. During familiarisation, documents, reflections, and scripts were



*Corresponding Author:* ***Visi Sumbom Tubuo***
visisumbom@gmail.com
*The University of Bamenda, Cameroon*

reviewed repeatedly to identify recurring tensions between candidate competence and performance quality. Coding then proceeded using the structured sheet; codes such as "mechanical fluency," "unnatural syntax," "construct irrelevance," and "policy omission" emerged early.

Themes were generated inductively and refined against the dataset. These included "AI-induced authenticity gaps," "question-design vulnerability," and "institutional preparedness." The analysis also drew conceptually on OECD's (2023) warning that AI exacerbates "systemic weaknesses in assessment systems reliant on static, product-centred tasks" (p. 10).

The final analytic narrative synthesised textual anomalies, examiner testimony, and policy shortcomings to demonstrate how AI complicates the validity of bilingual assessment.

### 7.6 Trustworthiness and Rigour

To ensure credibility, triangulation was achieved through multiple data sources. Dependability was supported by maintaining an audit trail. Reflexive memos documented interpretive decisions; confirmability was enhanced by separating descriptive observations from analytic claims. Transferability was addressed by providing rich contextual detail on the bilingual assessment environment, enabling applicability to other multilingual or high-stakes contexts.

### 7.7 Ethical Considerations

Ethical approval was obtained from the Assistant Chief Examiners and Examiners of the GCE A Level 0745 panel of French markers. All participants gave informed consent. Candidate anonymity was preserved by removing index numbers and institutional identifiers. Only authorised documents and scripts were used. AI tools were employed solely for editing clarity, in accordance with UNESCO's (2023) guidance that AI should "support but never substitute human judgement" (p. 14).

## 8. Findings And Analysis

### 8.1 Introduction

This section presents findings on AI-driven malpractice in Cameroon's bilingual General Certificate of Education (GCE) system. Following the qualitative, multi-method design described in section 7.1 above, analysis



*Corresponding Author:* ***Visi Sumbom Tubuo***
visisumbom@gmail.com
*The University of Bamenda, Cameroon*

integrated three data strands: coded script analysis (n = 224), examiner reflections, and policy documents. This is in line with Braun and Clarke's (2006) six-step thematic model guided coding, sub-theme refinement, and theme generation.

The findings are organised around four primary themes: (1) AI-assisted plagiarism, (2) impersonation and real-time assistance, (3) AI-assisted translation misuse, and (4) pre-prepared AI essays and memorisation. All findings explicitly reference the coding framework (Table 1), Bloom's cognitive levels, AI susceptibility, and observed anomalies. Triangulation across data sources ensures analytic robustness, and inter-coder reliability ($\kappa = 0.82$) substantiates the validity of the themes.

### 8.2 AI-Assisted Plagiarism

Analysis revealed that AI-assisted plagiarism compromises authorship and cognitive engagement. Scripts exhibited semantic over-precision and mechanical fluency, particularly in essay sections. Among 224 scripts, 10 demonstrated near-perfect translations (94/100) but failed essays (4/100), suggesting reliance on pre-generated AI content.

Examiner reflections corroborated this pattern: "Some essays were mechanically perfect, but the reasoning lacked depth; it appeared students had copied pre-prepared material."

Policy analysis indicated that the Cameroon GCE Board Malpractice Code (2022) does not address AI-generated content explicitly, limiting oversight of technologically mediated plagiarism. Coding aligned these anomalies with Bloom's levels from Apply to Create, high AI susceptibility, and indicators such as semantic over-precision and mechanical fluency (Table 7.1). Conceptually, these patterns reflect Vygotsky's (1978) insight that mediational tools, when misapplied, can circumvent authentic cognitive engagement.

### 8.3 Impersonation and Real-Time Assistance

Scripts alone cannot reveal impersonation, but sudden shifts in writing style within individual scripts were flagged through coding. Examiners noted:

"In several exams, the writing style changed abruptly, suggesting possible real-time assistance from hidden devices."



*Corresponding Author:* ***Visi Sumbom Tubuo***
visisumbom@gmail.com
*The University of Bamenda, Cameroon*

Policy documents confirm a gap: conventional invigilation focuses on visible cheating, ignoring digital impersonation or AI-assisted real-time aid. Coding anomalies included construct irrelevance and stylistic congruence, with Bloom's levels Apply → Analyse, and medium AI susceptibility (Table 1).

These findings illustrate that AI can extend beyond content generation, influencing candidate behaviour and challenging traditional assessment integrity.

### 8.4 AI-Assisted Translation Misuse

Analysis of French Paper 3 scripts showed perfect translations accompanied by low essay coherence, highlighting AI translation misuse. Examiners remarked:

"Students produced flawless translations, yet essays suggested no engagement with the topic."

Coding anomalies included translation–essay coherence issues and semantic over-precision, with Bloom's levels Understand → Evaluate and extremely high AI susceptibility. Policy review revealed that curriculum and assessment guidelines do not prescribe measures to detect or prevent AI-assisted translation, indicating a significant construct validity threat (Table 7.1). These results underscore the erosion of bilingual authenticity, as AI-generated translations mask true linguistic competence, challenging the assessment of mediationally mediated skills.

### 8.5 Pre-Prepared AI Essays and Memorisation

Multiple scripts exhibited identical argument structures and stylistic uniformity, suggesting pre-prepared AI content. Examiners observed:

"Several students submitted essays with highly similar phrasing; this is not typical of independent reasoning."

Coding identified stylistic uniformity, mechanical fluency, and construct irrelevance, with Bloom's levels Analyse → Create and high AI susceptibility.



*Corresponding Author:* ***Visi Sumbom Tubuo***
visisumbom@gmail.com
*The University of Bamenda, Cameroon*

Policy documents fail to mandate performative verification or oral synthesis tasks, allowing memorisation cultures to persist (Table 1).

These patterns demonstrate that AI facilitates mechanical reproduction, eroding originality and critical thinking.

### 8.6 Cross-Theme Triangulation

Triangulation of scripts, examiner reflections, and policy analysis reveals consistent patterns: AI influences essay quality, translation coherence, and candidate behaviour. Table 8.1 summarises the contrast between traditional GCE measures, emerging AI fraud, and triangulated evidence.

**Table 8.1: Comparative Framework — Authenticity vs. AI Influence**

| Dimension | Traditional Measures | Emerging AI Fraud | Triangulated Evidence (Scripts + Examiner + Policy) | Modern Solutions (Chapter 5) |
|---|---|---|---|---|
| Invigilation | Physical monitoring, device bans, cross-institution checks | Smartwatches, earpieces, preloaded AI dictionaries, memorized AI text | Examiner: style shifts; Scripts: uniform phrasing; Policy: no digital impersonation guidance | CCTV, radar, AI-assisted anomaly detection |
| Question Design & Marking | Memorisation-based essays, manual moderation | Pre-generated AI essays, perfect translations | Examiner: mechanical perfection with shallow reasoning; Scripts: near-perfect translations, | Oral verification, contextualised prompts, AI content detection |



*Corresponding Author:* ***Visi Sumbom Tubuo***
visisumbom@gmail.com
*The University of Bamenda, Cameroon*

| Dimension | Traditional Measures | Emerging AI Fraud | Triangulated Evidence (Scripts + Examiner + Policy) | Modern Solutions (Chapter 5) |
|---|---|---|---|---|
| | | | poor essays; Policy: no AI detection | |
| Candidate Verification | Index numbers, photographs, attendance | Identity verification irrelevant if content is AI-generated | Examiner: unusual performance patterns; Scripts: identical stylistic markers; Policy: lacks biometric verification | Handwriting databases, biometric logins, reflective statements |
| Examiner & Invigilator Competence | Experienced teachers, periodic briefings | Limited digital literacy | Examiner: difficulty detecting AI; Policy: no AI literacy requirement | AI literacy workshops, guidelines for stylistic pattern recognition |
| Institutional Infrastructure | Paper-based exams, ethical sensitisation | Lack of surveillance, outdated malpractice definitions | Scripts: anomalies flagged; Policy: no AI regulations | AI-regulatory framework, digital surveillance, post-exam audits |

**Source: Researchers' field work, 2025**



*Corresponding Author:* ***Visi Sumbom Tubuo***
visisumbom@gmail.com
*The University of Bamenda, Cameroon*

### 8.7 Technological and Coding Integration

**Table 8.2** presents global tools that align with coded anomalies and AI-susceptible tasks, illustrating how technological interventions can complement policy and pedagogy.

**Table 8.2: Global Technological Tools for AI-Detection & Prevention**

| Category | Function / Purpose | Example Tools | Coding Relevance | Limitations / Notes |
|---|---|---|---|---|
| AI-Text Detection | Identify AI-generated writing | Turnitin AI Detection, GPTZero, Copyleaks | Flags mechanical fluency, semantic over-precision | Limited accuracy for short texts, non-English languages |
| Translation & Linguistic Forensics | Detect machine-translated or stylistically inconsistent scripts | JStylo, Writeprints, Turnitin Draft Coach (MT detection) | Flags translation–essay coherence, lexical uniformity | Experimental, few multilingual models in Africa |
| Proctoring & Surveillance | Prevent real-time AI assistance | Examity, ProctorU, Respondus Monitor, Kialo | Identifies potential impersonation, real-time assistance | Privacy, infrastructure, cost constraints |
| Physical & Wearable Device Controls | Block or detect hidden smart devices | Smartwatches, earbuds, AR glasses, programmable calculators | Prevents AI-assisted essay or translation reproduction | Enforcement difficult due to device miniaturisation |
| Integrated Hybrid Approach | Combine detection, surveillance, | Stylometric + AI detection + CCTV + oral revalidation | Supports triangulated monitoring of anomalies | Requires phased adoption, training, and |



*Corresponding Author:* ***Visi Sumbom Tubuo***
visisumbom@gmail.com
*The University of Bamenda, Cameroon*

| Category | Function / Purpose | Example Tools | Coding Relevance | Limitations / Notes |
|---|---|---|---|---|
| | and verification | | flagged in coding | ethical safeguards |

**Source: Researchers' field work, 2025**

### 8.8 Conceptual Interpretation

Across all themes, AI functions as a mediational tool whose misuse disrupts authentic cognitive engagement. Patterns of semantic over-precision, mechanical fluency, and translation–essay incoherence indicate compromised bilingual competence and authenticity. This demonstrates that assessment constructs in Cameroon's bilingual GCE are increasingly vulnerable to technologically mediated malpractice, highlighting the need for ethical, pedagogical, and technological adaptation.

In conclusion, this study identified four central forms of AI-driven malpractice:

i. AI-assisted plagiarism
ii. Impersonation and real-time assistance
iii. AI-assisted translation misuse
iv. Pre-prepared AI essays and memorisation

Each theme is supported by triangulated evidence from scripts, examiner reflections, and policy analysis, and coded according to Bloom's taxonomy, AI susceptibility, and anomaly patterns. Conceptual integration links findings to Vygotskian mediation, authenticity, and bilingual construct validity. Importantly, findings remain data-driven, as reflexivity, audit trails, and inter-coder reliability ensure analytic rigor and confirmability.

## REFERENCES

Adebayo, O. O., & Abdulhamid, S. M. (2023). Academic dishonesty in the age of artificial intelligence: Implications for African higher education. *African Journal of Teacher Education, 12*(1), 45–63. Retrieved from https://www.ajol.info



*Corresponding Author:* ***Visi Sumbom Tubuo***
visisumbom@gmail.com
*The University of Bamenda, Cameroon*

Adu-Gyamfi, S., & Tchombe, T. M. S. (2022). Digital inequality and assessment integrity in African secondary schools. *African Educational Research Journal, 10*(2), 55–70.

Boateng, E., & Amankwah, A. (2021). Digital collaboration and academic misconduct in Ghanaian examinations. *Journal of African Education Studies, 17*(3), 225–238.

Bretag, T. (2018). *Challenges in addressing plagiarism in education.* PLoS Medicine, 15(3), e1002672. https://doi.org/10.1371/journal.pmed.1002672

Cameroon GCE Board. (2021). *Annual Report on Examination Administration and Reforms.* Buea: GCE Board.

Cameroon GCE Board. (2022). *Examination Regulations and Malpractice Code*. Buea: GCE Board Press.

Chan, R., & Hu, J. (2024). Ethical AI integration and assessment design in Asian higher education. *Asia-Pacific Journal of Educational Technology, 20*(1), 33–49.

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). ChatGPT, assessment, and the future of academic integrity. *Innovations in Education and Teaching International*, 60(2), 1–12.

Cotton, D., Nguyen, T., & Pérez, R. (2023). *Generative AI and educational integrity: Risks and recommendations*. Cambridge University Press.

Dawson, P., & Sutherland-Smith, W. (2019). Can markers detect contract cheating? The role of assessment design. *Assessment & Evaluation in Higher Education*, 44(5), 715–725.

Fishman, T. (2016). *The fundamental values of academic integrity* (3rd ed.). International Center for Academic Integrity. https://academicintegrity.org

Kithome, J., & Mwangi, P. (2019). Inequities in digital access and examination performance among Kenyan students. *African Journal of Educational Assessment, 12*(2), 75–92.

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson Education.



*Corresponding Author:* ***Visi Sumbom Tubuo***
*visisumbom@gmail.com*
*The University of Bamenda, Cameroon*

Lund, B. D., Wang, T., Mannuru, N. R., Nie, B., Shimray, S., & Wang, Z. (2023). ChatGPT and the academic integrity dilemma: A review. *Journal of Applied Learning and Teaching*, 6(1), 1–10.

MINESEC. (2025). *Address by Minister Nalova Lyonga on the theme for the 2025–2026 school year: Digital literacy and responsible technology use*. Ministry of Secondary Education, Cameroon.

Ngwa, T., & Mbah, E. (2024). "AI and Academic Dishonesty in Cameroon's Secondary Schools." *Journal of African Educational Research*, 18(2), 44–59.

Nwana, S. C., & Mbangwana, M. (2020). Examination malpractice and technology use among secondary school students in Cameroon. *International Journal of Education and Development, 8*(3), 89–104.

OECD. (2023). *Integrity in the Age of Artificial Intelligence: Challenges for Education Systems*. Paris: OECD Publishing.

Ojedokun, A., & Idowu, B. (2020). Mobile phones and examination malpractice among Nigerian students. *West African Journal of Educational Research, 27*(1), 41–56.

Sinha, R., & Liu, M. (2023). AI-assisted plagiarism and assessment ethics in South and Southeast Asia. *Journal of Educational Technology Research, 21*(2), 102–119.

Susnjak, T. (2022). The rise of AI-enabled plagiarism detection and its impact on privacy. *Computers & Education: Artificial Intelligence, 3*, 100059. https://doi.org/10.1016/j.caeai.2022.100059

Tambo, L. I. (2014). *Cameroon National Education Policy since the 1995 Forum.* Limbe: Design House.

Tchombe, T. M. S. (2019). Language policy, bilingualism, and learning equity in Cameroon's education system. *Cameroon Journal of Education, 15*(2), 102–118.

Timmis, S., Broadfoot, P., Sutherland, R., & Oldfield, A. (2016). Rethinking assessment in a digital age: Opportunities, challenges and risks. *British Educational Research Journal, 42*(3), 454–476. https://doi.org/10.1002/berj.3215

UNESCO. (2023). *AI and Education: Guidance for Policy-Makers.* Paris: UNESCO.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Harvard University Press.



*Corresponding Author:* ***Visi Sumbom Tubuo***
*visisumbom@gmail.com*
*The University of Bamenda, Cameroon*

WAEC. (2024). *Report on Digital Integrity and Emerging AI Threats in West African Examinations*. Accra: WAEC Press.



*Corresponding Author:* ***Visi Sumbom Tubuo***
*visisumbom@gmail.com*
*The University of Bamenda, Cameroon*