When Students Can Generate but Cannot Explain: A Case Study of Presentation-Based Assessment in Graphic Design Education at a Ghanaian Technical University

Al-hassan Bawa, Tamale Technical University, Ghana, (https://orcid.org/0000-0003-2187-0227); Nombeko Penelope Mpako, University of South Africa, South Africa, (https://orcid.org/0000-0002-2661-5463)

Abstract

The ubiquity of generative artificial intelligence (AI) in higher education presents both opportunities and significant pedagogical challenges. In graphic communication design, students increasingly rely on AI to conduct research, prepare presentations, generate design concepts, and produce videos—tasks traditionally requiring competence in computer-aided design (CAD) software. While these tools can enhance productivity, some students use them in ways that bypass learning, submitting AI-generated work without understanding underlying design principles or technical vocabulary. This study reports a qualitative case study conducted at Tamale Technical University in Northern Ghana, where 38 third-year graphic design students completed two presentation-based assessments requiring oral defense of their work. The study addressed three questions: (1) To what extent do students use generative AI in ways that hinder learning? (2) What patterns of terminology deficit, inconsistency, and anxiety emerge during oral defense? (3) How do students perceive the effectiveness of presentation-based assessment? Findings indicate that 81.6% of students used AI extensively, 89.5% could not define basic terminology from their own submissions, and 71.1% showed inconsistencies between submitted work and oral explanation. Despite initial anxiety, 71.1% reported that presentation-based assessment improved their learning. The study offers context-specific recommendations for design educators, including structuring interactive presentations, formulating probing questions, and clearly distinguishing between permitted and unethical AI use. It suggests that oral defense can reveal learning gaps, while treating claims about detecting unethical AI use cautiously and acknowledging alternative explanations such as language anxiety and prior instruction.

Keywords: Generative AI, graphic communication design, assessment integrity, oral presentation, viva voce, ethical AI use, technical university education, Ghana

Introduction

The rapid proliferation of generative artificial intelligence (AI) tools has transformed higher education, particularly in practice-based disciplines such as graphic communication design. Tools such as ChatGPT, Midjourney, DALL-E, and Canva's AI suite enable students to generate research summaries, slide decks, design concepts, and video content with minimal effort (Bozkurt et al., 2023; Southworth et al., 2023). While these technologies can enhance creativity and productivity, they also pose challenges to academic integrity, assessment design, and the development of foundational skills (Cotton et al., 2024; Sullivan et al., 2023).

This tension is especially evident in technical universities, where emphasis on hands-on competence and software proficiency conflicts with students' increasing reliance on AI as a

shortcut. In graphic communication design, students are expected to master computer-aided design (CAD) tools such as Adobe Illustrator, Photoshop, and InDesign. However, some bypass the deliberate practice these tools require, using AI to produce polished outputs without understanding principles of typography, layout, color theory, or visual hierarchy (Chiu, 2024; Meron & Araci, 2023; Verganti et al., 2020). This results in what Swiecki et al. (2022) term the "illusion of competence," where students appear capable through submitted artefacts but lack the conceptual and procedural knowledge to explain or defend their work.

The problem is intensified when assessment relies solely on written submissions. AI can generate coherent text, professional designs, and plausible captions with little trace of its use (Perkins et al., 2024), allowing gaps in learning to remain hidden. Students may also overestimate their competence, having never been required to articulate design decisions or justify creative choices independently (Mollick & Mollick, 2023).

This study reports a qualitative case study at Tamale Technical University in Northern Ghana, where a consistent pattern emerged: students submitted AI-generated design projects and presentations but were unable to explain basic terminology or defend design decisions during in-class presentations. The study is situated in a context of uneven digital infrastructure, where students have unequal access to personal computers but widespread access to AI tools via smartphones (Adarkwah, 2021; Modiba et al., 2025).

The article has three aims. First, it examines how presentation-based assessment reveals learning gaps in graphic communication design. Second, it critiques reliance on submission-only assessment, arguing that it may encourage superficial AI dependence. Third, it offers context-specific recommendations for educators navigating the integration of generative AI. The central argument is that in an environment where AI can generate outputs, oral defense provides a mechanism for making student understanding visible.

Research Questions

This study was guided by the following research questions:

1. To what extent do third-year graphic communication design students use generative AI in ways that may bypass learning, as evidenced in presentation performance?
2. What patterns of terminology deficit, inconsistency between submitted work and oral explanation, and anxiety emerge during presentation-based assessment?
3. How do students perceive the effectiveness of presentation-based assessment as a learning tool?

Literature Review

Oral and Viva Voce Assessment as an Authenticity Mechanism

Oral assessment, including viva voce ("living voice") examinations, has a long history in higher education as a means of authenticating student work and probing understanding. Unlike written submissions, oral assessments require students to respond spontaneously, defend their reasoning,

and demonstrate knowledge without external aids (Dawson, 2021). Sotiriadou et al. (2020) argue that viva voce enhances authenticity by making student thinking visible and reducing opportunities for academic misconduct. Similarly, institutional guidance presents oral defense as a strategy to minimize cheating and verify that submitted work reflects students' own understanding (Macquarie University, 2023).

Research shows that oral assessment can reveal gaps between submitted work and actual understanding. Joughin (2010) identifies three core dimensions: content (what is assessed), interaction (dialogic engagement), and authenticity (alignment with professional practice). In design education, the "crit" (critique) has long functioned as a formative oral assessment where students present work and receive feedback (Blair, 2006; Ezennia et al., 2025). However, its use as a summative mechanism for ensuring assessment integrity in the AI era remains underexplored.

In response to generative AI, scholars increasingly advocate a return to oral assessment. Lodge et al. (2023) and Ogunleye et al. (2024) suggest that live performance-based assessments are more resilient to AI misuse because they require spontaneous demonstration of understanding. Su et al. (2026) similarly highlight the role of authentic assessment in addressing generative AI challenges. This study builds on this emerging literature by examining a specific implementation of presentation-based assessment in graphic design education.

Generative AI Use and Academic Integrity

The prevalence of student AI use is well documented. Cotton et al. (2024) found that 68% of university students in a UK sample had used ChatGPT for assessed work, often without considering it cheating, while Perkins et al. (2024) reported rates exceeding 80% in some disciplines. Scholars, however, distinguish between forms of AI use. Mollick and Mollick (2023) argue that AI can function as "co-intelligence" for ideation and drafting when students critically engage with outputs. Unethical use, by contrast, involves submitting AI-generated work without understanding, attribution, or critical evaluation (Perkins et al., 2024).

This study adopts a nuanced position: AI use is not inherently unethical. It becomes problematic when students cannot explain or defend AI-generated content, present it as their own original work, or bypass learning processes altogether. This distinction is central to interpreting the study's findings.

Terminology Deficit and Academic Literacy

A smaller body of literature examines the relationship between AI use and disciplinary vocabulary acquisition. Scholars note that AI-generated text may include vocabulary exceeding students' comprehension, creating what Lodge et al. (2023) term the "illusion of competence." However, empirical evidence remains limited. This study addresses this gap by providing concrete data on terminology deficits among AI-using students.

Theoretical Framework

This study is informed by two complementary perspectives: constructivist learning theory and situated assessment theory. Constructivism holds that learners actively construct knowledge through experience, reflection, and social interaction rather than passively receiving information (Piaget, 1973; Vygotsky, 1978). From this perspective, learning involves developing mental models that can be applied in new contexts, not merely producing correct outputs (Jonassen & Rohrer-Murphy, 1999). In design education, this is reflected in studio-based learning, where students learn through doing, reflection, and dialogue with peers and instructors (Schön, 1987).

Applied to AI use, constructivism suggests that when students rely on AI to generate designs or presentations without engaging in underlying cognitive processes, they bypass durable learning. Explaining one's work during a presentation becomes a constructive act: students must organize knowledge, translate tacit understanding into language, and respond to challenges. This aligns with Vygotsky's (1978) concept of the Zone of Proximal Development (ZPD), where learning occurs through guided interaction. The Q&A session functions as a ZPD activity, with lecturer questioning scaffolding students' articulation of design reasoning.

Situated assessment theory, drawing on situated learning (Lave & Wenger, 1991), argues that assessment should reflect authentic professional practice. Competence is demonstrated in context through performance, not abstract testing (Gulikers et al., 2004). In graphic communication design, professional practice requires presenting work, explaining decisions, responding to feedback, and revising accordingly (Sawyer, 2017; Schön, 2010). Presentation-based assessment mirrors this process. Together, these frameworks explain why such assessment reveals learning gaps that submission-only approaches may obscure.

Methodology

Research Design

This study employed a qualitative case study design (Yin, 2018), which is appropriate for investigating a contemporary phenomenon within its real-life context. The phenomenon under investigation was student performance during presentation-based assessment in a graphic communication design course. A case study design allowed for in-depth exploration of students' behaviours, explanations, and learning outcomes as they occurred naturally within the classroom setting (Tisdell et al., 2025). This paper is positioned as an empirical case study combined with pedagogical reflection.

Research Context

The study was conducted at Tamale Technical University (TaTU) in the Northern Region of Ghana. TaTU is one of Ghana's technical universities, offering competency-based programmes emphasising practical skills and industrial readiness. The Department of Graphic Communication Design enrols approximately 120 students across four levels (100, 200, 300 and 400). The study focused on a third-year cohort (n = 38) enrolled in a required course, "Advertising I," during the 2024-2025 academic year. The institutional context is significant for two reasons. First, students have unequal access to personal computers but near-universal access to smartphones, making AI tools such as ChatGPT and Canva's AI features widely accessible via mobile devices (Adarkwah,

2021). Second, class sizes are moderate (30-40 students), making presentation-based assessment logistically feasible, unlike very large class size.

Participants

Participants were 38 third-year students (33 males, 5 females, aged 21-25 years) pursuing a Bachelor of Technology in Graphic Communication Design. Purposive sampling was used because all students enrolled in the course were required to complete the same assessment tasks (Patton, 2015). Students had previously completed introductory courses, but prior assessments had been submission-only. All 38 students provided informed consent. Ethical approval was obtained from the TaTU Research Ethics Committee. Students were assured that participation or non-participation in the study would not affect their grades.

Researcher Positionality

The principal author served as both the course lecturer and the primary researcher. This dual role introduces potential observer and response bias. To mitigate this, a reflexivity journal was maintained, grading was completed before analysis, peer debriefing was conducted with a colleague (the second author) not involved in teaching, and supportive questioning strategies were employed. These measures are further addressed in limitations.

Assessment Design and Procedure

The course "Advertising I" was redesigned to incorporate two major assessment tasks, each comprising a submitted artefact and an in-class presentation with Q&A.

Assessment Task 1: Design Research Poster

Students were required to research a contemporary design topic (e.g., "The Role of Color Psychology in Brand Identity," "Typography Trends in West African Advertising") and produce a research poster using computer-aided design software (Adobe InDesign or Illustrator). The poster had to include visual elements, text, and data visualization. After submission, each student presented their poster to the class using a projector for 8–10 minutes, followed by 5–7 minutes of Q&A from the lecturer and peers.

Assessment Task 2: Brand Identity Project

Students were required to develop a complete brand identity for a hypothetical local business (e.g., a restaurant, tailor shop, or technology hub). Deliverables included a logo, color palette, typography system, stationery mockups, and a 30-second promotional video. Students submitted digital files and then presented their work in 12–15 minutes, explaining their design rationale, software techniques, and creative decisions, followed by 7–10 minutes of Q&A.

Presentation Guidelines

Students were informed at the start of the semester that presentations would be graded on: (a) clarity of explanation, (b) appropriate use of design terminology, (c) ability to answer questions, and (d) demonstrated understanding of software processes. They were explicitly told that AI-generated content was permitted as a research or ideation tool but that they would be required to

explain and defend every element of their submission. The rubric (see Appendix A) allocated 40% of the total grade to presentation and Q&A performance.

Data Collection

Data were collected over a 14-week semester (January–April, 2025) using multiple sources to ensure triangulation (Denzin & Lincoln, 2018).

1. Observation and Field Notes

The researcher (also the course lecturer) took detailed field notes during each presentation session, recording: (a) terminology that students used but could not define when questioned, (b) inconsistencies between submitted work and oral explanations, (c) specific questions that students struggled to answer, and (d) observable signs of anxiety or avoidance during Q&A. A total of 38 observation records were produced.

2. Video Recordings

With student consent, all presentations were video recorded. Recordings were reviewed after each session to verify field notes and capture verbatim exchanges during Q&A. Only researcher-reviewed recordings were used; no recordings were shared externally.

3. Semi-Structured Follow-Up Interviews

Following the two assessment tasks, 12 students (31.6% of the cohort) were purposively selected for follow-up interviews based on observed difficulties during presentations. Selection criteria included: (a) inability to define three or more terms from their own slides, (b) inconsistency between submitted work and oral explanation, or (c) admitted AI use without understanding. Interviews lasted 20–30 minutes and explored students' AI use habits, understanding of design concepts, and perceptions of the presentation requirement. Interviews were audio-recorded and transcribed verbatim.

4. Document Analysis

Submitted design files (posters, logos, mockups, videos) were analyzed for evidence of AI generation, including: unusually polished output inconsistent with prior work, generic visual styles typical of AI tools, and metadata indicating AI software use. However, document analysis alone was not used to determine AI use; rather, it was compared with presentation performance

Data Analysis

Data analysis followed thematic analysis procedures outlined by Braun and Clarke (2006, 2021), comprising six phases: familiarization, initial coding, theme generation, theme review, theme definition, and write-up.

Phase 1: Familiarization: The researcher transcribed interview recordings, reviewed field notes, and watched video recordings multiple times to achieve deep familiarity with the data.

Phase 2: Initial Coding: Using NVivo 14 software, the researcher generated initial codes from the data. Examples of codes included: "cannot define contrast," "admits using ChatGPT for text," "inconsistent color rationale," "anxiety during Q&A," and "copied AI-generated phrase."

Phase 3: Theme Generation: Codes were grouped into potential themes. For instance, "cannot define contrast," "cannot explain kerning," and "uses 'hierarchy' incorrectly" were grouped under a theme labelled "Terminology Deficit."

Phase 4: Theme Review: Themes were reviewed against the coded data and the full dataset to ensure they accurately represented student experiences.

Phase 5: Theme Definition: Final themes were defined, named, and described with clear boundaries.

Phase 6: Write-Up: Themes were written up with illustrative quotations and observational evidence.

To enhance trustworthiness, two strategies were employed: peer debriefing (a colleague reviewed coding decisions) and member checking (five interviewed students reviewed their transcripts and confirmed accuracy) (Lincoln & Guba, 1985).

Determining Evidence of AI Use

Evidence of AI use was determined through triangulation of multiple indicators: (a) linguistic markers (phrases typical of ChatGPT such as "in the dynamic landscape of"), (b) stylistic markers in visuals (typical of Midjourney or DALL-E), (c) metadata where available, (d) student admissions during interviews or Q&A, and (e) comparison with prior student work. No single indicator was sufficient; at least two indicators were required for classification as "extensive AI use." This study does not claim perfect detection accuracy but rather reports observable patterns.

Ethical Considerations

Several ethical safeguards were implemented. First, students were informed that the research was separate from grading and that declining participation would not affect their assessment. Second, all identifying information was removed from reported data; pseudonyms (e.g., S1, S2) are used in the Findings section. Third, during presentations, the researcher took care to phrase questions supportively rather than punitively, aiming to diagnose understanding rather than embarrass students. Fourth, students who struggled were offered additional tutoring sessions after data collection.

Findings

The analysis of observation records, video recordings, follow-up interviews, and submitted design work revealed five themes related to students' AI use and performance during presentation-based assessment: (1) Widespread but Undisclosed AI Use, (2) Terminology Deficit and Superficial Vocabulary, (3) Inconsistency Between Submitted Work and Oral Explanation, (4) Anxiety and Avoidance During Q&A, and (5) Perceived Effectiveness of Presentation-Based Assessment.

Theme 1: Widespread but Undisclosed AI Use
Of the 38 students, observational evidence suggested that 31 (81.6%) used generative AI extensively, though only 8 (21.1%) acknowledged this during presentations. Indicators included text with characteristic ChatGPT phrasing (e.g., "In the dynamic landscape of contemporary visual communication…"), visual styles typical of Midjourney or DALL·E, and AI-generated voiceovers.

Field notes illustrate this pattern:

Student P7 reads: "The heuristic evaluation of semiotic modalities reveals polysemic interpretations." When asked to define "polysemic," the student pauses for 12 seconds, responds "many meanings?" but cannot provide an example, later admitting: "The AI wrote that sentence." (Field notes, Week 3)

Interview data confirmed widespread reliance. Student S4 noted:

"I use ChatGPT for almost everything… But when you asked me about 'kerning,' I had no idea. The AI put that word in my notes." (S4, interview)

Similarly, Student S9 stated:

"I used Midjourney for logos… When asked about color choices, I couldn't explain. The AI chose them." (S9, interview)

Three students (7.9%) submitted work that appeared entirely AI-generated, with no evidence of software use or process knowledge. When questioned, they could not describe any design steps. The issue is not AI use per se, but students' inability to explain AI-generated content, indicating a learning gap.

Theme 2: Terminology Deficit and Superficial Vocabulary
The most consistent finding was students' inability to define or correctly use basic design terminology from their own work. A total of 34 students (89.5%) could not define at least three terms during Q&A. Table 1 presents the most frequently misused or unexplained terms.

Table 1: Design Terminology Students Could Not Define or Explain

Term	Number of students unable to define (n=38)	Example of incorrect/vague explanation
Kerning	31 (81.6%)	"It's like spacing" (could not distinguish from tracking)
Hierarchy (visual)	28 (73.7%)	"The important things are bigger" (no mention of contrast, position, or gaze)
Contrast	24 (63.2%)	"Different colors" (no mention of value, size, or type contrast)
Polysemy	22 (57.9%)	"Many meanings" (could not apply to a visual example)

Semiotics	19 (50.0%)	"Study of signs" (could not differentiate icon/index/symbol)
Gestalt	17 (44.7%)	"Something about grouping" (could not name any principle)
Typographic hierarchy	26 (68.4%)	"Big headings, small text" (no mention of weight, color, or placement)
Hexadecimal	15 (39.5%)	"Color code" (could not explain what the digits represent)

Field notes captured a representative exchange:

Student R12 presents a poster about color theory. Slide reads: “Utilizing complementary color schemes establishes visual tension while maintaining harmonic equilibrium.” Lecturer asks: “What is a complementary color scheme?” Student responds: “Colors that work well together.” Lecturer: “Can you give me an example?” Student: “Blue and green?” Lecturer: “Are those complementary?” Student: “I’m not sure. The AI wrote that.” (Field notes, Week 5)

Follow-up interviews suggest two main causes of terminology deficits: (a) reliance on AI-generated text that students did not fully engage with, and (b) prior coursework that did not require oral explanation. As Student S7 noted:

“In first and second year, we just submitted our designs… No one asked us to explain anything. Now I realize I don’t know the words for what I made.” (S7, interview)

Theme 3: Inconsistency Between Submitted Work and Oral Explanation
A third theme was the inconsistency between submitted work and students’ ability to explain it. Of the 38 students, 27 (71.1%) demonstrated at least one significant mismatch between artefact and explanation.

Examples included:

- **Color rationale mismatches:** A student used purple and orange for a funeral services logo but could not justify their combined meaning. (Field notes, Week 8)
- **Software technique claims:** A student claimed use of the Pen tool but could not demonstrate it, later admitting reliance on an AI vector generator. (Field notes, Week 10)
- **Typographic decisions:** A student selected three fonts but could not name or justify them. (Field notes, Week 4)
- **Video production claims:** A student attributed a video to Adobe Premiere Pro but could not explain its creation; metadata indicated AI generation. (Field notes, Week 12)

Student S11 reflected:

“I felt embarrassed… The AI made the poster… I didn’t want to admit that in front of the class.” (S11, interview)

Theme 4: Anxiety and Avoidance During Q&A
Observational data indicated high anxiety during Q&A, especially when students were asked to explain terminology or decisions. Indicators included long pauses (>5 seconds; n = 29, 76.3%), avoiding eye contact (n = 26, 68.4%), tentative responses (n = 24, 63.2%), and physical withdrawal from the podium (n = 18, 47.4%).

A representative incident illustrates both anxiety and learning potential:

Student P19 is asked to explain menu hierarchy. After a long pause, the student whispers, "I don't know. I just copied what the AI made." The lecturer responds supportively, guiding a two-minute discussion. The student later states: "Now I understand." (Field notes, Week 9)

While such admissions enabled productive learning moments, most students initially attempted answers before acknowledging AI use or remaining silent.

Interview data confirmed this pattern. Student S2 stated:

"My heart was beating so fast… I knew I had used AI… I was hoping you wouldn't ask." (S2, interview)

In contrast, Student S14 reported a shift in behavior:

"At first I hated the presentations… but then I realized I had to learn the work. For the second project, I made sure I understood everything." (S14, interview)

Theme 5: Perceived Effectiveness of Presentation-Based Assessment
Despite initial resistance, 27 students (71.1%) reported that presentations improved their learning. Reflections included:

"Before this course, I would just generate and submit. Now I know I have to understand what I submit." (S22)
"The questions made me realize I didn't know design vocabulary." (S18)
"All courses should have presentations… otherwise students will just use AI and learn nothing." (S31)

However, 8 students (21.1%) viewed the requirement negatively, citing stress and language barriers. Some accommodations were provided, including the option to use English and Dagbani, though uptake was limited.

From the lecturer's perspective, presentation-based assessment revealed learning gaps that submission-only assessment obscured:

"If I had just marked these posters, every student would have passed… The presentations showed that most students had no idea what they had submitted. The output was AI's competence, not theirs." (Researcher journal, Week 6)

Discussion

Alternative Explanations

Before interpreting the findings, alternative explanations must be considered. Student anxiety during Q&A may reflect general public speaking anxiety or fear of negative evaluation rather than AI misuse. Terminology deficits may stem from weak prior instruction or English as a second language, and inconsistencies may reflect poor study habits or time pressure. These possibilities are acknowledged. The interpretation that follows therefore identifies patterns consistent with AI-related learning gaps without making causal claims.

Interpretation of Findings

The finding that 89.5% of students could not define basic design terminology from their own presentations (Theme 2) challenges constructivist notions of meaningful learning. Constructivism emphasizes active cognitive construction rather than the production of correct outputs (Piaget, 1973). In this study, students produced visually polished artefacts—posters, slide decks, and brand identities—but lacked the underlying knowledge required to explain or apply them. This reflects what Swiecki et al. (2022) describe as the “illusion of competence,” where AI-generated artefacts are mistaken for evidence of learning.

From a Vygotskian perspective (Vygotsky, 1978), the Q&A session functioned as a scaffolded learning interaction. When students acknowledged gaps in understanding (e.g., inability to define “kerning”), the lecturer provided targeted support within the Zone of Proximal Development. The incident described in Theme 4, where a student admitted “I don’t know” and then engaged in a guided discussion, illustrates this process. Crucially, such scaffolding depends on making misunderstanding visible—something submission-only assessment does not require. Presentation-based assessment therefore transforms assessment from evaluation into a site of learning.

The inconsistency between submitted work and oral explanation (Theme 3) can be interpreted through situated assessment theory (Gulikers et al., 2004). Submission-only assessment is decontextualized, evaluating artefacts in isolation rather than in the context of professional practice. In contrast, design work in professional settings requires explanation, justification, and responsiveness to feedback. Students in this study could produce artefacts but could not perform these discursive practices.

This aligns with Lave and Wenger’s (1991) concept of legitimate peripheral participation. Students relying on AI without understanding remain on the periphery of the design community, lacking the language and reasoning required for full participation. Presentation-based assessment, by requiring students to articulate and defend their work, either develops these capacities or exposes their absence. The finding that 71.1% of students reported improved learning (Theme 5) suggests that confronting this gap can promote more authentic engagement.

Comparison with Existing Literature

The prevalence of undisclosed AI use (81.6%) aligns with recent international findings. Cotton, Cotton, and Shipway (2024) reported that 68% of university students in a UK sample used ChatGPT for assessed work, often without considering it cheating, while Perkins et al. (2024) found rates exceeding 80% in some disciplines. This study extends these findings to a Ghanaian technical university context, suggesting that AI misuse is not confined to Western settings but represents a broader global challenge requiring context-sensitive pedagogical responses.

The terminology deficit identified in this study adds a new dimension to the literature. Prior research has focused on AI's capacity to generate coherent text (Rudolph et al., 2023) and the difficulty of detecting AI use (Weber-Wulff et al., 2023), but has paid less attention to whether students understand the language they present. Findings here suggest that students may reproduce AI-generated text containing vocabulary beyond their comprehension. This has implications for academic literacy, as AI may both mask language deficiencies and hinder the development of disciplinary vocabulary through authentic use.

The finding that 71.1% of students reported positive perceptions of presentation-based assessment, despite its stressfulness, aligns with Sotiriadou et al. (2020) on the authenticity benefits of oral assessment. Su et al. (2026) similarly note that students often initially resist oral assessment due to anxiety but later value its contribution to learning. This study supports that pattern and suggests that even in contexts where English is not the first language, most students recognize its pedagogical value. The limited uptake of local language options (Dagbani) indicates a potential area for further consideration in designing inclusive assessment practices.

Implications for Practice

The most immediate implication is that lecturers in design and related disciplines should move beyond submission-only assessment, as AI can produce competent-looking artefacts that students cannot explain or defend. Practical implementation of presentation-based assessment includes announcing requirements early, designing rubrics that allocate 30–50% of marks to presentation and Q&A performance, and preparing diagnostic questions focused on terminology, design rationale, and software processes. Lecturers should foster a supportive environment in which students who admit AI use are guided rather than penalized, accommodate language diversity (e.g., allowing local languages), and manage larger classes through group presentations or rotating schedules.

At the program level, departments should review assessment practices across the curriculum, as a single presentation-based course cannot offset reliance on submission-only assessment elsewhere. Program coherence is needed so that students progressively encounter oral defense requirements. At the institutional level, Tamale Technical University and similar institutions should provide policy guidance clarifying that AI may be used as a tool, but students remain responsible for understanding their work. Institutions should also embed oral defense in capstone projects, offer faculty development on AI-resilient assessment design, and invest in basic presentation infrastructure.

At the sector level, the Ghana Tertiary Education Commission (GTEC) should address assessment design alongside student conduct. Policies focused solely on detecting or punishing AI misuse are unlikely to succeed given the limitations of detection tools. Instead, guidance should promote assessment methods that make understanding visible, such as oral presentations, live demonstrations, and process portfolios. In contexts marked by digital inequality, presentation-based assessment also offers a more equitable alternative to technologically intensive proctoring systems, requiring minimal infrastructure.

Limitations

First, the researcher also served as the course lecturer, introducing potential observer bias. Although peer debriefing and member checking were used, this dual role may have influenced student responses during Q&A, either increasing anxiety or encouraging greater honesty. Future studies should involve independent observers or multiple lecturers.

Second, no control group was included. A quasi-experimental design comparing presentation-based and submission-only sections would strengthen causal inference.

Third, the study spanned a single semester. The long-term effects of presentation-based assessment on student behavior and AI use remain unknown. Longitudinal research is needed to examine whether students modify their practices in subsequent courses.

Fourth, the study was conducted at a single institution in northern Ghana. While this provides contextual depth, generalizability is limited. However, the underlying theoretical mechanisms—constructivist learning and situated assessment—may be transferable to similar contexts.

Finally, evidence of AI use was based on observable indicators rather than forensic detection, and alternative explanations (e.g., anxiety, language proficiency, prior instruction) cannot be ruled out. Accordingly, claims are modest: the study identifies patterns consistent with AI-related learning gaps but does not establish causation.

Conclusion

This case study examined presentation-based assessment in a graphic communication design course at a Ghanaian technical university. Findings indicate that 81.6% of students used AI extensively, 89.5% could not define basic terminology from their own submissions, and 71.1% showed inconsistencies between submitted work and oral explanation. Despite initial anxiety, 71.1% reported that presentation-based assessment improved their learning.

The contribution of this paper is modest and context-specific: it provides an empirical case study showing how oral defense can reveal learning gaps that submission-only assessment obscures. It extends established literature on viva voce assessment (Sotiriadou et al., 2020; Joughin, 2010) to the generative AI context. The study does not claim that presentation-based assessment is novel or that it definitively detects AI misuse; rather, it demonstrates how oral defense can make understanding visible and support scaffolded learning.

The study offers practical implications for design educators, including distinguishing ethical from unethical AI use, accommodating language diversity, and sequencing oral defense across programs. The implication is not to prohibit AI, but to design assessments that require students to demonstrate understanding through explanation and defense.

References

Adarkwah, M. A. (2021). "I'm not against online teaching, but what about us?": ICT in Ghana post COVID-19. *Education and Information Technologies, 26*(2), 1665–1685.

Blair, B. (2006). *Perception, interpretation, impact: An examination of the learning value of formative feedback to students through the design studio critique* (Doctoral dissertation, Institute of Education, University of London).

Bozkurt, A., Xiao, J., Lambert, S., Pazurek, A., Crompton, H., Koseoglu, S., … Jandric, P. (2023). Speculative futures on ChatGPT and generative artificial intelligence (AI). *Asian Journal of Distance Education, 18*(1), 53–130.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101.

Braun, V., & Clarke, V. (2021). *Thematic analysis: A practical guide*. SAGE.

Chiu, T. K. F. (2024). Future research recommendations for transforming higher education with generative AI. *Computers and Education: Artificial Intelligence, 6*, 100197.

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International, 61*(2), 228–239.

Dawson, P. (2021). *Defending assessment security in a digital world*. Routledge.

Denzin, N. K., & Lincoln, Y. S. (Eds.). (2018). *The SAGE handbook of qualitative research* (5th ed.). SAGE.

Ezennia, I. S., & Agbonome, P. C. (2025). Advances in design studio critique techniques: A narrative review on innovative critique models and lessons for improving architecture education. *Discover Education, 4*(1), 369.

Gulikers, J. T. M., Bastiaens, T. J., & Kirschner, P. A. (2004). A five-dimensional framework for authentic assessment. *Educational Technology Research and Development, 52*(3), 67–86.

Jonassen, D. H., & Rohrer-Murphy, L. (1999). Activity theory as a framework for designing constructivist learning environments. *Educational Technology Research and Development, 47*(1), 61–79.

Joughin, G. (2010). *A short guide to oral assessment*. Leeds Met Press in association with University of Wollongong.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.

Lodge, J. M., Howard, S., Bearman, M., & Dawson, P. (2023). *Assessment reform for the age of artificial intelligence*. Australian Government Tertiary Education Quality and Standards Agency (TEQSA). https://www.teqsa.gov.au/guidesresources/resources/corporate-publications/assessment-reform-age-artificial-intelligence

Macquarie University. (2023). *Authentic assessment through viva voce*. TEACH Resource. https://teche.mq.edu.au/2023/03/authentic-assessment-through-viva-voce/

Meron, Y., & Araci, Y. T. (2023). Artificial intelligence in design education: Evaluating ChatGPT as a virtual colleague. *Design Science, 9*, e30.

Modiba, F. S., Van den Berg, A., & Mago, S. (2025). Opportunities and challenges of generative AI supporting research in African classrooms. *South African Journal of Higher Education, 39*(3), 173–193.

Mollick, E. R., & Mollick, L. (2023). *Using AI to implement effective teaching strategies in classrooms*. Wharton School Research Paper.

Ogunleye, B., Zakariyyah, K. I., Ajao, O., Olayinka, O., & Sharma, H. (2024). Higher education assessment practice in the era of generative AI tools. *Journal of Applied Learning & Teaching, 7*(1), 46–56.

Patton, M. Q. (2015). *Qualitative research and evaluation methods* (4th ed.). SAGE.

Perkins, M., Roe, J., Postma, D., McGaughran, J., & Hickerson, D. (2024). Detection of GPT-4 generated text in higher education. *Journal of Academic Ethics, 22*(1), 89–113.

Piaget, J. (1973). *To understand is to invent: The future of education*. Grossman Publishers.

Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments? *Journal of Applied Learning and Teaching, 6*(1), 342–363.

Sawyer, R. K. (2017). Teaching creativity in art and design studio classes. *Educational Research Review, 22*, 99–113.

Schön, D. A. (1987). *Educating the reflective practitioner*. Jossey-Bass.

Schön, D. A. (2010). Educating the reflective practitioner. *Australian Journal of Adult Learning, 50*(2), 448–451.

Sotiriadou, P., Logan, D., Daly, A., & Guest, R. (2020). The role of authentic assessment to preserve academic integrity and promote skill development and employability. *Studies in Higher Education, 45*(11), 2132–2148.

Southworth, J., Migliaccio, K., Glover, J., Glover, J. N., Reed, D., McCarty, C., Thomas, A., Bracey, E., Heller, R., & Odom, J. (2023). Developing a model for AI across the curriculum. *Computers and Education: Artificial Intelligence, 4*, 100127.

Su, F., O'Dea, M., Wood, M., & Seidl, E. (2026). Embracing a new world: Authentic assessment designs in an age of GenAI. *Perspectives: Policy and Practice in Higher Education*, 1–12.

Sullivan, M., Kelly, A., & McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity. *Journal of Applied Learning and Teaching, 6*(1), 31–40.

Swiecki, Z., Khosravi, H., Chen, G., Martinez-Maldonado, R., Lodge, J. M., Milligan, S., Buckingham Shum, S., & Gasevic, D. (2022). Assessment in the age of artificial intelligence. *Computers and Education: Artificial Intelligence, 3*, 100075.

Tisdell, E. J., Merriam, S. B., & Stuckey-Peyrot, H. L. (2025). *Qualitative research: A guide to design and implementation*. Wiley.

Verganti, R., Vendraminelli, L., & Iansiti, M. (2020). Innovation and design in the age of artificial intelligence. *Journal of Product Innovation Management, 37*(3), 212–227.

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

Weber-Wulff, D., Anohina-Naumeca, A., Bjelobaba, S., Foltýnek, T., Guerrero-Dib, J., Popoola, O., … Waddington, L. (2023). Testing of detection tools for AI-generated text. *International Journal for Educational Integrity, 19*(1), 1–39.

Yin, R. K. (2018). *Case study research and applications* (6th ed.). SAGE.

Appendix A: Presentation Assessment Rubric (Sample)

Criteria	**Excellent (4)**	**Proficient (3)**	**Basic (2)**	**Unsatisfactory (1)**

Explanation of design rationale	Clearly explains all major design decisions	Explains most decisions, minor gaps	Explains only basic decisions	Cannot explain decisions
Use of design terminology	Uses 5+ terms correctly	Uses 3-4 terms correctly	Uses 1-2 terms correctly	Uses no terminology
Response to Q&A	Answers all questions confidently	Answers most questions	Answers some questions	Cannot answer most questions
Demonstration of software process	Clearly describes steps and techniques	Describes steps, some details missing	Vague description	Cannot describe process
Clarity and organisation	Well-structured, within time	Mostly clear, minor issues	Some disorganisation	Unclear, disorganise