ISSN: 2473-4901
v12 n6772

From Threat to Opportunity: AI-Assisted Oral Assessment as a Scalable Assessment Design

Ae-Sook Kim
ae-sook.kim@quinnipiac.edu

Hans Nguyen
hans.nguyen@quinnipiac.edu

Yan Jin
yan.jin@quinnipiac.edu

Quinnipiac University
Hamden, CT 06518 USA

Abstract

Generative AI has challenged traditional assessment approaches in higher education, creating an urgent need for assessments that can not only provide valid evidence of learning outcomes but also promote learning. This study examines whether AI-assisted oral assessment can preserve these dual purposes while providing a scalable alternative to traditional approaches. Using a randomized controlled trial design with Institutional Review Board (IRB) exempt approval, the study was conducted in two business analytics courses, including both undergraduate and graduate students, at a private university in the Northeastern United States. Descriptive statistics, *t*-tests, ANOVA, and qualitative content analysis were employed to analyze the data. Overall, students who participated in AI-assisted oral assessment reported higher scores in perceived learning outcomes and assessment experiences. Students' qualitative feedback further suggested that it promoted self-reflection and metacognitive awareness, strengthened critical thinking, and improved communication skills. These findings suggest that AI-assisted oral assessment can preserve the learning benefits of oral assessment while offering a more scalable approach to implementation. Future studies could explore this topic across different educational contexts and assessment designs to enhance our understanding of how AI can help, rather than hurt, assessment.

Keywords: AI-assisted assessment, oral assessment, authentic assessment

From Threat to Opportunity: AI-Assisted Oral Assessment as a Scalable Assessment Design

Ae-Sook Kim, Hans Nguyen, and Yan Jin

1. INTRODUCTION

The rapid adoption of generative artificial intelligence (Generative AI) has fundamentally challenged academic assessments in business analytics and information systems education in higher education. Students are increasingly relying on AI for code generation, data analysis, and problem-solving. Tasks that traditionally served as evidence of student learning, including multiple-choice questions, short-answer responses, case analyses, and even reflections, can now be completed with minimal student effort using AI tools such as ChatGPT and Claude (Cotton et al., 2024; Farazouli et al., 2024; Newton & Xiromeriti, 2024). Consequently, concerns about the validity of traditional assessment methods have grown substantially. In a recent report by the American Association of Colleges and Universities (AAC&U, 2025), 78% of college faculty reported experiencing an increase in academic integrity issues, primarily attributed to Generative AI. These developments have intensified calls to redesign assessment in the age of Generative AI (Batat, 2025; Dubay & Richards, 2024), including calls for such changes in information systems education (Van Slyke et al., 2023).

Assessments serve two complementary functions in higher education. On the one hand, they provide valid evidence of student learning (summative function); on the other hand, they promote learning by encouraging reflection and self-improvement (formative function) (Biggs, 1996; Black & Wiliam, 1998; Earl, 2012; Sadler, 1989). Therefore, the challenge facing higher education is not simply to redesign assessment to continue providing valid evidence of student learning, but also to preserve its capacity to promote learning in the age of Generative AI.

In response to these challenges, educators have increasingly turned to authentic assessment approaches that require students to demonstrate rather than merely recall knowledge (Conlon & Wilson, 2026; Kofinas et al., 2025). Unlike traditional approaches that often emphasize the final answer, authentic assessments emphasize the thinking process by requiring students to explain their reasoning and apply knowledge in realistic contexts (Kickbusch et al., 2025; Wiggins, 1990). One prominent form of authentic assessment is oral assessment, in which students demonstrate their understanding through verbal responses to interactive questioning (Fenton, 2025). The recent emergence of Generative AI has renewed interest in this approach due to its potential to preserve the dual purposes of assessment—providing valid evidence of student learning and promoting learning. First, its interactive and real-time nature requires students to generate unscripted responses on the spot, making it easier to verify genuine understanding while reducing opportunities to overly rely on Generative AI (Droulers et al., 2026; Fenton, 2025; Ward et al., 2024). Second, by requiring students to justify their reasoning, it encourages students to identify gaps in their understanding and engage in deeper cognitive processing, thereby promoting learning (Chi & Wylie, 2014; Rawls et al., 2015).

Despite these pedagogical advantages, oral assessment remains difficult to implement at scale because it requires substantial instructor time and effort to develop customized questions, conduct one-on-one conversations, and evaluate students' responses (Droulers et al., 2026; Fenton, 2025). Consequently, although oral assessment is widely recognized as a valuable approach, its usage is often limited to small classes or high-stakes examinations, situations where additional time investment is feasible. Recent advances in artificial intelligence have created new opportunities to address this longstanding scalability constraint. First, rather than relying on predetermined question banks, Generative AI can generate contextually relevant follow-up questions to students' written answers, enabling instructors to evaluate not only what students know but also how they think (Ba et al., 2026; Kasneci et al., 2023; McGrath et al., 2025). Second, automatic speech recognition (ASR) technology, commonly referred to as voice-to-text technology, can accurately transcribe students' spoken responses for evaluation, substantially reducing the heavy burden associated with conducting and evaluating oral responses (Chao et al., 2026). Together, these advances make it increasingly feasible to incorporate oral assessment into routine classroom practice while preserving its pedagogical benefits.

The integration of artificial intelligence (AI) into teaching and learning is not new in educational research. This growing stream of research has examined the use of AI as a tool to provide automated and immediate feedback (e.g., Mahapatra, 2024), personalized recommendations (e.g., Quan, 2025), and dynamic tutoring (e.g., Urban et al., 2024). These functions are particularly useful to boost student engagement, enhance cognitive learning processes and overall learning outcomes, as suggested by prior literature reviews and conceptual work (Debay and Richards, 2024; Estaphan et al., 2025; Fenton, 2025; Kickbusch et al., 2025; Luo et al., 2025), as well as a recent meta-analysis by Han et al. (2026). However, discussions of Generative AI in assessment have largely emphasized its potential to undermine academic integrity and assessment validity (AAC&U, 2025; Cotton et al., 2024).

A small but growing body of research has started exploring the positive role of Generative AI in assessment. For instance, Generative AI has been used to support test preparation by generating practice questions, as well as to improve accessibility by providing text-to-speech, translation, content summarization functions, and adjustable display formats (e.g., font size) to better support diverse learners (Luo et al., 2025; McGrath et al., 2025; Zhao et al., 2025). While these applications may enhance students' assessment experiences, they do not directly strengthen the core educational dual purposes of assessment. To our knowledge, no empirical study has examined whether Generative AI can strengthen assessment itself by enabling authentic assessment approaches (e.g., oral assessment) to be implemented at scale. Accordingly, this study contributes to the literature by conceptualizing AI-assisted oral assessment as a scalable assessment design and providing preliminary evidence of its learning benefits.

This study investigates the effectiveness of AI-assisted oral assessment while also considering its potential as a scalable assessment approach. Specifically, we address the following research question: *Does the use of AI-assisted oral assessment improve students' perceived learning outcomes in business analytics education?*

2. CONCEPTUAL FRAMEWORK FOR AI-ASSISTED AUTHENTIC ORAL ASSESSMENT

Authentic assessment has emerged as a widely endorsed approach in business education, including analytics and marketing, because it evaluates what students can do with their knowledge in realistic contexts, rather than what they can merely recall on a test (Cotton et al., 2024; Newton & Xiromeriti, 2024). Authentic assessment can take multiple forms, including oral assessment, simulations, role-play, and other performance-based tasks, such as analyzing a case, presenting a recommendation, or producing a professional-quality artifact (e.g., dashboards, reports, or presentations) that mirror workplace demands. In these tasks, students are expected to demonstrate critical thinking, applied judgment, and decision-making in context (Archbald & Newmann, 1988; Conlon & Wilson, 2026; Dahl et al., 2018). From this perspective and in line with Wiggins (1990), authentic assessment is best understood as a direct evaluation of student performance on tasks that require students to carry out meaningful intellectual work, rather than recalling information from memory. As the ultimate goal is to improve task performance, not just to monitor it, authentic assessment provides students with the opportunity to explain their reasoning and support their answers with appropriate evidence, rather than simply selecting the correct option. These tasks also mirror the complex ambiguities and ill-structured priorities of professional life, requiring students to navigate trade-offs, justify assumptions, and adapt their responses to stakeholder needs. In doing so, authentic assessment creates a tighter alignment between the intended learning outcomes and the tasks and criteria used to evaluate them, consistent with constructive alignment principles (Biggs, 1996).

The significant role of authentic assessment becomes even more critical in the age of Generative AI. In traditional education, outcome-focused assessments such as multiple-choice questions or short written responses could still function as reasonable proxies for learning because students typically had to do the underlying work themselves to arrive at a correct answer. In AI-era classrooms, however, students can easily generate polished explanations or solutions with AI tools, which weakens the connection between the submitted answer and the learning process that supposedly produced it. In other words, non-authentic assessments that only capture final outcomes may no longer provide a reliable picture of what students actually understand. Thus, authentic assessment shifts from an optional enhancement to a necessary approach to ensure that students, even when using AI, remain actively involved in

explaining their reasoning, making informed judgments, and applying course concepts in ways that reflect genuine learning (Kickbusch et al., 2025; Kofinas et al., 2025).

Within this broader move toward authentic assessment, oral assessment represents a particularly strong form of authentic assessment because it requires students to articulate, justify, and refine their thinking under realistic, time-limited conditions (Kofinas et al., 2025). It introduces cognitive challenges to students as they must respond under time pressure, revise their reasoning when prompted, and demonstrate evaluative judgment through explanation and self-correction.

Despite these benefits, traditional oral assessments require substantial instructor time for question generation, individual exam administration, and evaluation, and they may produce inconsistent scoring when human judgment varies across students (Wiggins, 1990). These characteristics make oral assessment difficult to implement in large classes. Recent advances in Generative AI and automated speech recognition decouple these labor costs from class size, enabling the pedagogical benefits of oral assessment, such as articulating reasoning and metacognitive awareness, to be delivered efficiently at scale. By automating the resource-intensive component, this approach preserves the authentic, dialogic features of oral assessment while improving scalability without compromising validity. Figure 1 summarizes this conceptual framework by contrasting non-authentic, authentic, and AI-assisted authentic assessment in terms of their validity and scalability in AI-era education.

	Non-Authentic Assessment	Authentic Assessment	AI-Assisted Authentic Assessment
Ability to reveal and promote learning	**Low** (emphasizes outcomes over learning processes & vulnerable to AI-generated responses)	**High** (provides direct observations of student learning & promotes learning through reflection, and metacognition)	**High** (Preserves the ability to observe & promote learning)
Implementation Scalability	**High** (easy to implement at scale)	**Low** (resource intensive)	**High** (scalable due to technological support

Figure 1. Conceptual Framework

3. METHODOLOGY

Given the novelty of AI-assisted oral assessment, this study is exploratory in nature. Rather than testing formal hypotheses, we compare the experimental conditions on the three outcome constructs and use the two baseline constructs to examine the comparability of the two groups.

3.1 Study Design

To examine the impact of AI-assisted oral assessment on student perceived learning outcomes, this study employed a randomized controlled experiment. Students were randomly assigned to either a control (no oral follow-up question after submitting a short answer written response) or a treatment (having an oral follow-up question) group during their weekly knowledge check assessments.

Students from two classes at a private university in the United States were invited to participate in the survey. Both were 7-week accelerated asynchronous online courses in analytics, one at the undergraduate level and the other at the graduate level.

The survey was designed to capture five main constructs: two baseline constructs for students' self-reported academic standing and familiarity with AI tools, and three outcome constructs including two for perceived learning outcomes and one for assessment experience. In particular, the learning

outcomes were measured in two ways: students' overall perception of topic mastery, and students' self-rated confidence in performing tasks across Bloom Taxonomy levels (understand, analyze, apply, evaluate, and create). Each construct was a composite score of the three to five survey instruments. Each survey instrument was measured with a 7-point Likert scale, 1 meaning strongly disagree, 4 meaning neither agree nor disagree, and 7 meaning strongly agree. We included the full questionnaire in Appendix A.

3.2 Procedures

To avoid any unintended harm to human subjects, this study, including the intervention and post-treatment survey, was approved by the University's Institutional Review Board (IRB) under the exempt review category.

This research leveraged the AI-assisted oral assessment feature provided by a collaborating partner specializing in technology integration for higher education. The intervention was integrated into two assessments in Modules 4 and 5 consecutively, in the middle of the course. For each assessment, students in both control and treatment groups saw the same assignment description in their learning management system (LMS) and then clicked on a link leading to our collaborating partner's website. This link directed the two groups to their respective sets of questions that were almost identical, consisting of the same multiple-choice questions and a short answer question. The only difference occurred after the completion of the short-answer question. While students in the control group exited the assignment, students in the treatment group received an additional AI-assisted oral follow-up question, prompting them to defend, clarify, or extend their written response. Examples are presented in Appendix B. Students had to record their response in 90 seconds and submit it. The countdown automatically started when AI generated a follow-up question and presented it to students. Students' responses were automatically analyzed and graded by AI.

After concluding Module 5, students were invited to participate in the online survey administered via Qualtrics. Students were informed that their responses were anonymous and voluntary, and they should provide explicit consent for participation. As an incentive, minor extra credit was offered. The survey was closed after 7 days.

3.3 Methods

On the one hand, we utilized quantitative analysis with students' post-intervention survey data to provide descriptive statistics, independent samples t-tests, and two-way analysis of variance (ANOVA) for key constructs using R programming language. On the other hand, we explored qualitative insights from treatment group students' response to two open-ended questions (N=29), using an inductive thematic analysis approach, allowing themes to emerge from the data. Each response was carefully reviewed, and any recurring themes across responses were recorded. To enhance validity and reliability, initial coding was refined through multiple iterations to ensure both internal coherence within each theme and clear distinction between themes. This iterative review process was conducted in collaboration with ChatGPT-5.5 (OpenAI, 2026), an approach that has been supported in the literature for enhancing face validity and augmenting content analysis (Jalali & Akhavan, 2024; Wachinger et al., 2024). The researcher independently reviewed each student's written feedback and generated an initial set of themes. ChatGPT was used to propose themes from the same responses, which were then compared with the researcher's coding. The themes were subsequently reviewed, refined, and revised through multiple rounds of interaction with ChatGPT. Final coding decisions were made solely by the researcher based on critical evaluation of AI-generated suggestions throughout the process. As one response can reflect multiple themes, theme percentages summed to more than 100%.

3.4 Participants

Fifty students out of fifty-six students participated in the survey. However, during the quality screening processes, one response was removed due to its invariant response patterns, raising reliability concerns owing to its inattentive response. As a result, 49 responses were included in the subsequent analysis, representing an 87.5% response rate (see Table 1). The high response rate likely minimized non-response bias, suggesting that the sample was broadly representative of the student population in both classes. In total, 27 valid participants were undergraduate and 22 were graduate students.

Academic Level	# of Participants	Total # Students	Response Rate
Undergraduate	27	31	87.1%
Graduate	22	25	88.0%
Total	**49**	**56**	**87.5%**

Table 1. Survey Response Rate

The survey participants consisted of 59% treatment and 41% control group students, with 55% undergraduates, 65% females, and 63% business majors. Almost all non-business majors were undergraduate students in healthcare analytics from diverse schools including Health Sciences, Nursing, and the College of Arts and Sciences. About 70% of the undergraduate participants had non-business majors, and about 89% of them were female. The average age of the participants was 22.7, ranging from 19 to 50. One outlier was an adult learner enrolled in the online degree program (see Table 2).

Experiment	N	%
Control	20	40.8
Treatment	29	59.2
Academic	**N**	**%**
Undergraduate	27	55.1
Graduate	22	44.9
Gender	**N**	**%**
Male	17	34.7
Female	32	65.3
Academic Level	**N**	**%**
Sophomore	7	14.3
Junior	6	12.2
Senior	16	32.7
Graduate	20	40.8
Majors	**N**	**%**
Business Major	31	63.3
Non-business Major	18	36.7
TOTAL	**49**	**100.0%**

Table 2. The Survey Participants' Characteristics

4. RESULTS

4.1 Self-Reported Baseline Characteristics, Learning Outcomes, and Experiences

Cronbach's alpha values for the five constructs ranged from 0.8 to 0.94, indicating internal consistency. The baseline self-reported academic standing, two learning outcomes constructs, and students' assessment experience scores ranged from 5.7 to 5.8, which were somewhat positive (refer to Table 3). The baseline AI familiarity, however, scored relatively low (mean = 5.0). Students, in particular, rated "neutral" on their frequent use of AI tools for their coursework.

Construct (# of Items)	Survey Items	Mean(S.D.)		Experiment Mean(S.D.)		Academic Level Mean(S.D.)		
		N=49	Cronbach's α	Control (n=20)	Treatment (n=29)	Undergraduate (n=27)	Graduate (n=22)	Cohen's d (Effect Size)
C1. Self-Reported Academic Standing (N=5)	I feel confident in my overall academic ability.	5.96 (1.32)		5.80 (1.64)	6.07 (1.07)	5.63 (1.62)	6.36 (0.66)*	-0.57 (medium)
	I have strong analytical reasoning skills.	5.63 (1.27)		5.55 (1.36)	5.69 (1.23)	5.26 (1.43)	6.09 (0.87)*	-0.69 (medium)
	I communicate clearly and effectively (spoken).	5.78 (1.28)		5.50 (1.43)	5.97 (1.15)	5.44 (1.53)	6.18 (0.73)*	-0.60 (medium)
	I communicate clearly and effectively (written).	5.94 (1.23)		5.90 (1.48)	5.97 (1.05)	5.59 (1.50)	6.36 (0.58)*	-0.65 (medium)
	I perform well under time pressure.	5.33 (1.52)		5.15 (1.66)	5.45 (1.43)	4.96 (1.70)	5.77 (1.15)+	-0.55 (medium)
	Composite Score	5.73 (1.18)	0.93	5.58 (1.41)	5.83 (1.00)	5.38 (1.39)	6.15 (0.65)*	-0.69 (medium)
C2. AI Familiarity (Baseline) (N=3)	I understand how to use AI tools appropriately.	5.65 (1.36)		5.35 (1.57)	5.86 (1.19)	5.04 (1.51)	6.41 (0.59)***	-1.16 (large)
	I frequently use AI tools for my coursework.	4.33 (1.57)		4.35 (1.57)	4.31 (1.61)	3.89 (1.53)	4.86 (1.49)*	-0.65 (medium)
	I have a positive attitude toward using AI tools.	5.12 (1.56)		5.00 (1.69)	5.21 (1.50)	4.85 (1.70)	5.45 (1.34)	–
	Composite Score	5.03 (1.27)	0.80	4.90 (1.35)	5.13 (1.22)	4.59 (1.34)	5.58 (0.94)**	-0.83 (large)
C3. Learning Outcomes (N=3)	I have become confident in my understanding of the topic.	5.82 (1.03)		6.05 (0.89)	5.66 (1.11)	5.63 (1.15)	6.05 (0.84)	–
	I have become knowledgeable about the topic.	5.94 (0.99)		6.20 (0.70)	5.76 (1.12)+	5.74 (1.13)	6.18 (0.73)	–
	I am confident in explaining the topic clearly to others.	5.27 (1.30)		5.35 (1.23)	5.21 (1.37)	5.00 (1.27)	5.59 (1.30)	–
	Composite Score	5.67 (1.02)	0.90	5.87 (0.80)	5.54 (1.14)	5.46 (1.12)	5.94 (0.83)+	–
C4. Bloom's Taxonomy Learning Outcomes (N=5)	Helped me understand key concepts.	5.78 (1.25)		5.55 (1.43)	5.93 (1.10)	5.70 (1.41)	5.86 (1.04)	–
	Helped me analyze reasoning and recognize gaps.	5.80 (1.31)		5.75 (1.45)	5.83 (1.23)	5.56 (1.53)	6.09 (0.92)	–
	Allowed me to justify and defend my understanding.	5.96 (1.19)		5.90 (1.48)	6.00 (0.96)	5.63 (1.42)	6.36 (0.66)*	-0.64 (medium)
	Pushed me to brainstorm new ideas.	5.59 (1.38)		5.50 (1.67)	5.66 (1.17)	5.30 (1.54)	5.95 (1.09)+	-0.48 (small)
	Helped me apply content to different situations.	5.80 (1.27)		5.65 (1.60)	5.90 (1.01)	5.52 (1.45)	6.14 (0.94)+	-0.49 (small)
	Composite Score	5.78 (1.15)	0.94	5.67 (1.41)	5.86 (0.94)	5.54 (1.34)	6.08 (0.78)+	–
C5. Assessment Experience (N=4)	The knowledge check was a fair assessment experience to evaluate	5.57 (1.4)		5.60 (1.31)	5.55 (1.48)	5.15 (1.56)	6.09 (0.97)*	-0.71 (medium)
	The knowledge check was well aligned with course content.	6.06 (1.14)		6.05 (1.23)	6.07 (1.10)	5.81 (1.30)	6.36 (0.85)+	-0.49 (small)
	The knowledge check motivated deeper learning over memorization.	5.65 (1.11)		5.60 (1.23)	5.69 (1.04)	5.33 (1.14)	6.05 (0.95)*	-0.67 (medium)
	The knowledge check allowed me to demonstrate what I really know.	5.67 (1.18)		5.45 (1.36)	5.83 (1.04)	5.33 (1.24)	6.09 (0.97)*	-0.67 (medium)
	*The knowledge check caused me stress.	3.51 (2.01)		3.00 (1.75)	3.86 (2.13)	3.59 (1.89)	3.41 (2.20)	–
	Composite Score	5.74 (1.02)	0.86	5.68 (1.04)	5.78 (1.01)	5.41 (1.04)	6.15 (0.83)**	-0.78 (medium)

Note: 1. * This item was excluded from the composite score due to its reverse nature.
2. T-Test results, statistical significance: + p< 0.1, * p < .05, ** p < .01, *** p < .001. Cohen's d reported for significant comparisons (small ≥ .20, medium ≥ .50, large ≥ .80).

Table 3. Self-Reported Outcomes and T-Test Results by Experiment and by Academic Level

Even though students in the treatment group showed a higher score in most survey items, no significant difference was found between control and treatment groups across five constructs. On the contrary, undergraduate and graduate students showed meaningful differences across three constructs. On average, graduate students had higher scores than undergraduate students in baseline academic standing and AI familiarity. In terms of AI familiarity, graduate students were more confident in their ability to use AI tools appropriately and reported more frequent use of AI tools for their coursework. The corresponding effect sizes were medium to large, indicating that graduate students scored 0.69 and 0.83 standard deviations higher than undergraduate participants on these two measures, respectively. Additionally, graduate students had higher scores on one outcome construct, assessment experience, with an effect size of 0.78, indicating that their scores were 0.78 standard deviation higher than those of undergraduate students. However, no substantial mean differences were observed for the remaining two outcome constructs, although one item, "The recent knowledge check allowed me to justify and defend my understanding," showed a notable difference. Graduate students, on average, scored 0.64 standard deviations higher than undergraduate students. The density plots in Figure 2 visually showed differences in the distributions of composite scores by academic level, with graduate students more concentrated at higher scores (>=6) than their counterparts.

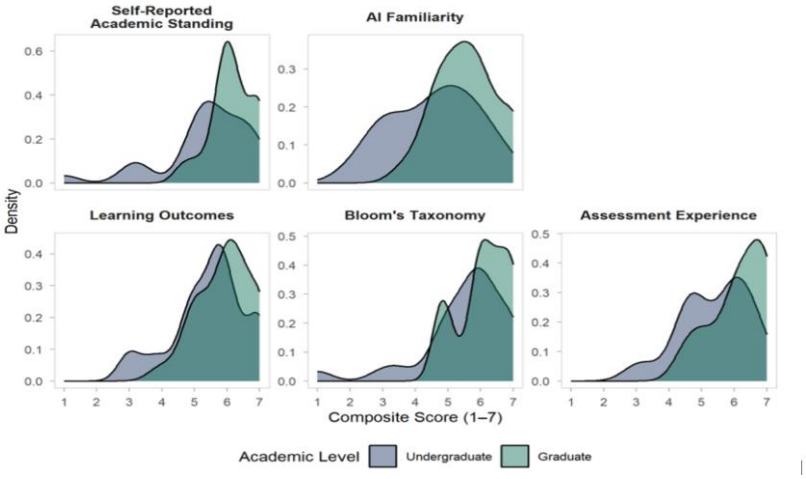

Figure 2. Density Distributions of Five Constructs by Academic Level

A two-way ANOVA was conducted to explore whether the effects of experimental conditions varied by academic level. The results did not provide clear evidence of interaction effects across learning outcome

ISSN: 2473-4901
v12 n6772

and assessment experience constructs. In general, the gap in students' perceived learning outcomes between academic groups was consistent regardless of experimental conditions, with graduate students reporting higher scores across conditions. The upward trend from undergraduate to graduate students was consistent across the three outcome constructs (see Figure 3).

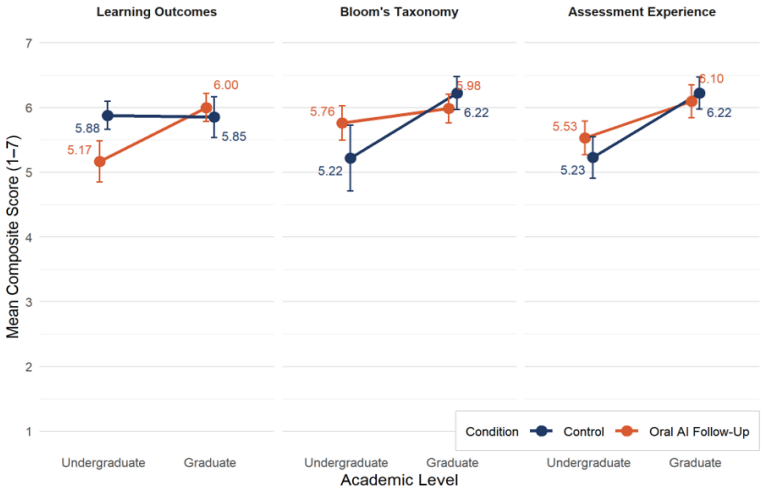

Figure 3. Interaction Plots: Treatment Condition × Academic Level on Three Outcome Constructs

Although the interaction effects were not pronounced, a consistent directional crossover pattern was observed across all three constructs. The oral AI follow-up group was associated with higher scores among undergraduate students for the assessment experience and Bloom's Taxonomy learning outcome constructs, but lower scores for the other learning outcome construct; these patterns were reversed among graduate students.

4.2 Learning Processes and Outcomes from Thematic Analysis

Regarding the question, "Describe one example where an AI-generated follow-up question led you to change, refine, or extend your original written answer. What did you learn from that experience?", thematic analysis identified four distinct themes/: (1) refine responses for greater specificity, detail, and evidence, (2) enhance understanding through feedback, (3) recognition of knowledge gaps and self-assessment, and (4) critical thinking and re-analysis (see Table 4).

Theme	N	%
Refined responses for greater specificity, detail, and evidence	17	60.7
Enhanced understanding through feedback	11	39.3
Recognition of knowledge gaps and self-assessment	8	28.6
Critical thinking and re-analysis	7	25.0
Total # of Valid Responses	**28**	**100**

Notes: Three responses were classified as miscellaneous: two noted appreciations for the revision opportunity as a motivational feature, and one described stress related to the oral recording component.

Table 4. Learning Processes Associated with AI-Assisted Follow-Up Questions

Over 60% of students indicated that AI-assisted follow-up questions helped refine their answers for greater specificity, detail, and supporting evidence. One student explained that "the AI-generated feedback led me to refine and extend my originally written answer to be more specific. From that experience, I learned the importance of using specific examples to help get my point across." It encouraged them to support their claim with evidence and data to justify their conclusions. Meanwhile, about 40% of students acknowledged that it enhanced their understanding, as one reflected, "If I got it wrong they told me what I was missing and this helped me learn more about the concept and made me

understand it more" and another student explained that the follow-up questioning process helped them understand "why not all charts can qualify for any scenarios," deepening their conceptual understanding.

Additionally, over 25% of students thought it helped them reflect on their mistakes, reconsider and adjust their initial approaches. One participant explained, "I am able to make a guided analysis as to where I went wrong or a mistake I made, and reflect on that to improve for next time." Other students similarly reported "rethinking my approach" and learning to recognize weaknesses in their initial responses.

Students also reported that it promoted critical thinking and reanalysis, challenging them to think beyond their first answer and examine ideas from new angles. Students described that it "pivots my thinking to take a new approach", "re-analyzes responses," and evaluates "actionable insights based on the provided data." Altogether, it was evident that AI-assisted follow-up questions functioned as a form of formative feedback that prompted students to refine their thinking and deeply engage with course content and their learning.

Similarly, thematic analysis of responses to the question "Overall, how did the written test plus AI-generated follow-up questions affect your learning, if at all?" identified four distinct themes . Students demonstrated that AI-assisted follow-up oral questions facilitated (1) deeper understanding and knowledge reinforcement, (2) increased reflection and self-awareness, (3) improved communication and response quality, and (4) enhanced critical thinking and deeper processing (see Table 5).

Theme	N	%
Deeper understanding and knowledge reinforcement	15	51.7
Increased reflection and self-awareness	7	24.1
Improved communication and response quality	6	20.7
Enhanced critical thinking and deeper processing	5	17.2
Total # of Valid Responses	**29**	**100**

Notes: Six responses were classified as miscellaneous. They reflected minimal perceived impact on learning, general positive perceptions of the assessment experience, or were unrelated to the learning outcomes under investigation.

Table 5. Learning Outcomes Associated with AI-Assisted Follow-Up Questions

More than half of the students reported that it enhanced their learning, describing "a better understanding," an opportunity to "learn more deeply", and an ability to "retain what I was learning." One student explained, it "helped me learn more deeply by turning quick, partial answers into clearer, more accurate explanations… pushed me to explain my reasoning, connect ideas, and add specific examples from the data or output…could correct mistakes while the material was still fresh and see how revising my answer improved my score, which made the concepts easier to remember and apply in the future." Meanwhile, about a quarter of respondents suggested that it increased their awareness of their mistakes, knowledge gaps, and opportunities for improvement. AI-assisted follow-up questions helped them identify weaknesses and gaps, which students described as "where I was weak", "what I did wrong", and "highlight what was missing." These realizations subsequently prompted them to "improve", "rethink", "revisit my answer", or "correct mistakes." One student stated that it "greatly improved my learning because I was able to reflect on my incorrect answers and state how I can improve or explain what went wrong. This allowed me to truly learn from my mistake and remember how to correct it for the next time."

Additionally, students described learning to provide clearer explanations, using phrases such as "hitting the right point", "expand on" responses, and "refine answer to be clear and actionable." One student indicated that it "helped me learn more deeply by turning quick, partial answers into clearer, more accurate explanations…the follow-up questions pushed me to explain my reasoning, connect ideas, and add specific examples from the data or output."

Lastly, students reported that the follow-up question pushed them "to think deeper" and "to explain my reasoning, connect ideas." One student stated that it "challenged my original thought process and forced

me to think even more about my explanation and answers." Although these did not emerge as distinct themes, students also mentioned increased engagement with course content and their ability to apply learned concepts to real-world situations.

5. DISCUSSION

The findings of this study suggest that AI-assisted follow-up questions can positively contribute to students' learning outcomes. The mean scores of the perceived learning items ranged from 5.27 to 5.96, all well above the midpoint of 4.0, indicating that students generally reported positive perceptions of their learning. Those outcomes include understanding and knowledgeable of the topic, analyzing reasoning and recognizing gaps, justifying and defending their understanding, and applying content to different situations in the context of business analytics education. Additionally, a medium-to-large gap was found in learning outcomes between undergraduate and graduate students. This gap is consistent with the existing literature supporting the "Knowledge-is-Power" hypothesis, which suggests that prior knowledge has a profound positive effect on learning (Dochy et al., 1999). Graduate students, on average, reported more favorable learning outcomes and assessment experience, as well as higher levels of higher-order thinking, than undergraduates.

Meanwhile, thematic analysis of student written feedback further provided solid evidence that AI-assisted follow-up questions functioned as a formative learning mechanism, prompting students to re-examine their initial responses, engage in self-assessment, and develop greater metacognitive awareness of their knowledge gaps. More than half of the students reported that engaging with AI follow-up questions deepened and solidified their understanding, suggesting that interactive AI engagement actively contributes to their learning. This aligns with Kickbusch et al. (2025)'s characterization of authentic assessment as fostering "critical reflection, iterative feedback, and collaborative knowledge building" (p. 3), capabilities that are increasingly critical for navigating complex digital environments in the AI era.

Despite the positive learning outcomes reported, no meaningful difference was found between groups with (treatment) and without (control) AI-assisted follow-up questions. This should be interpreted with caution, given several methodological constraints. First, the relatively small sample size may have contributed to the lack of a clear difference, making it difficult to detect the small-to-medium gap between groups. Second, the substantial gap in 'prior knowledge' between graduate and undergraduate students, reflected in differences in the baseline academic standing and AI familiarity, may have contributed to the existing variance in students' perceived learning outcomes, in turn obscuring treatment effects. The crossover patterns between experimental conditions and academic levels implied that the lack of a clear overall treatment effect may be due to unequal or reverse effects across groups. For example, the control group outperformed the oral AI follow-up group among undergraduates, whereas this pattern reversed among graduate students. These opposing directional effects across groups cancelled each other out in the overall treatment mean, resulting in little overall difference between the treatment and control groups. Reliably detecting interaction effects requires a sufficiently large sample, with sufficient statistical power and the sample size in the current study was too limited to provide definitive evidence of such effects.

From an authentic assessment perspective, the survey results indicated that students experienced AI-assisted oral assessment as a meaningful learning activity. It offers a practical way to shift from outcome-only assessment toward more process-oriented evidence of learning, in line with calls to redesign authenticity around visible processes and judgment in AI-mediated contexts (Kickbusch et al., 2025). The findings of this study also illustrated that AI-assisted oral assessment preserves key benefits of authentic assessment by making students' learning processes more visible while supporting improved learning outcomes. At the same time, this approach leverages AI to increase efficiency and alleviate scalability constraints that often limit oral assessment in business and analytics courses (Conlon & Wilson, 2026; Fenton, 2025). For example, administering even a brief three-minute individual oral exam would require approximately two hours of instructor time for a class of 40 students, before accounting for question development and grading. With AI-assisted oral assessment, instructor effort would shift from administering and grading oral exams for individual students to higher-level tasks such as setting up oral exams, designing the rubric, and validating AI-generated scores, so that oral assessment becomes feasible in larger classes where it would otherwise be impractical. We should also note that AI-assisted oral assessment does not replace the need for careful assignment design and human instructor

judgement, but it offers a feasible way to make dialogic, process-focused tasks more widely implementable in AI-era education.

6. CONCLUSION AND FUTURE WORK

Higher education is facing significant assessment challenges. While Generative AI leads to increasing concerns about the validity of traditional assessments, it may also create new opportunities to rethink how student learning is evaluated and supported. This study explored one such possibility with the use of AI-assisted oral assessment, a scalable form of oral assessment. In our qualitative analysis, students consistently reported that the oral follow-up process helped them refine their thinking, recognize knowledge gaps, justify their reasoning, and engage more deeply with course concepts. These findings suggest that AI-assisted oral defense represents a promising assessment innovation deserving further attention. This study contributes to the emerging literature on AI in education by uncovering AI's positive role in assessment. Rather than viewing Generative AI solely as a threat to assessment validity, our findings suggest that integrating AI into the assessment process may support learning by encouraging students to elaborate on their reasoning, identify knowledge gaps, and connect ideas. More broadly, our findings suggest the potential of Generative AI in addressing a longstanding challenge in assessment: balancing pedagogical value with scalability. By combining AI-generated follow-up questions with the oral defense approach, this study conceptualizes a scalable form of oral assessment and offers early evidence of its promise to address AI-introduced assessment challenges. In doing so, our research responds to recent calls in extant literature to enhance assessment experiences (Fenton, 2025; Kofinas et al., 2025) and to understand the increasing relevance of AI in education (Batat, 2025; Dubay & Richards, 2024; Fenton, 2025).

There are several practical considerations for implementing AI-assisted oral assessment. To maximize its learning potential, two areas warrant particular attention: assessment transparency and assessment design. The clarity and transparency regarding the grading rubric and the process used to generate follow-up questions may enhance students' understanding of the assessment and their confidence in its fairness, thereby better supporting students' learning. Additionally, extension of the 90-second response time limit may reduce unnecessary time pressure for students, allowing them to formulate more thoughtful and well-reasoned oral responses.

This study has some limitations that could be avenues for future investigations. Students' exposure to the oral defense is limited to only two weekly assignments (out of seven), and 90 seconds for each response, which may have contributed to the small effect size observed. In addition, the learning outcomes were primarily based on a self-reported survey, which is prone to biases in students' self-perceptions. Sample size proved to be another limitation, which could also explain the lack of clear differences despite the observed mean differences between the control and treatment groups for multiple measures. Nevertheless, we believe the findings provide a promising foundation for future research on the novel integration of Generative AI in assessment. Future research could build on this work to directly measure the impact on instructors' time and effort, as well as incorporate objective performance measures (e.g., test scores), more significant interventions (e.g., a longer intervention period, such as a full semester), larger sample sizes, and subjects beyond analytics or business to further understand the impact of this novel AI-assisted oral assessment approach.

7. REFERENCES

AAC&U (2025). *The AI challenge: How college faculty assess the present and future of higher education in the age of AI*. Retrieved May 10, 2026 from https://imaginingthedigitalfuture.org/wp-content/uploads/2026/01/Elon-AACU-faculty-AI-survey-full-report-1-21-26.pdf

Archbald, D. A., & Newmann, F. M. (1988). *Beyond standardized testing: Assessing authentic academic achievement*. National Association of Secondary School Principals.

Ba, S., Shi, X., Wu, S., & Lu, G. (2026). Artificial intelligence agents in computer-supported collaborative learning: A systematic literature review. *Computers and Education: Artificial Intelligence*, *10*. https://doi.org/10.1016/j.caeai.2026.100579

Batat, W. (2025). Holixec education: From atomistic and artificial intelligence (AI) to holixec intelligence (HI) for human-centric leadership in the phygital age. *Journal of Macromarketing*, *45*(3), 528–541. https://doi.org/10.1177/02761467251350356

Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education, 32*(3), 347-364. https://doi.org/10.1007/BF00138871

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 7–74. https://doi.org/10.1080/0969595980050102

Chao, F.-A., Yan, B.-C., & Chen, B. (2026). Probing the hidden talent of ASR foundation models for L2 English oral assessment. In *ICASSP 2026-2026 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 4056–4060). https://doi.org/10.48550/arXiv.2510.16387

Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, *49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823

Conlon, B., & Wilson, S. (2026). Enhancing statistics education through student-generated data and authentic assessment. *Assessment & Evaluation in Higher Education*, *51*(4), 791–810. https://doi.org/10.1080/02602938.2025.2553894

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. https://doi.org/10.1080/14703297.2023.2190148

Dahl, A. J., Peltier, J. W., & Schibrowsky, J. A. (2018). Critical thinking and reflective learning in the marketing education literature: A historical perspective and future research needs. *Journal of Marketing Education, 40*(2), 101–116. https://doi.org/10.1177/0273475317752452

Dochy, F., Segers, M., & Buehl, M. (1999). The relation between assessment practices and outcomes of studies: The case of research on prior knowledge. *Review of Educational Resear*ch, *69*(2), 145–186. https://doi.org/10.2307/1170673

Droulers, M., Krautloher, A., & Shaeri, S. (2026). Interactive oral assessment: Co-existence of formative and summative purposes. *Teaching in Higher Education*, *31*(1), 67–85. https://doi.org/10.1080/13562517.2025.2549945

Dubay, C. M., & Richards, M. B. (2024). Leveraging artificial intelligence in project-based service learning to advance sustainable development: A pedagogical approach for marketing education. *Marketing Education Review*, *34*(4), 307–323. https://doi.org/10.1080/10528008.2024.2411975

Earl, L. M. (2012). *Assessment as learning: Using classroom assessment to maximize student learning*. Corwin Press.

Estaphan, S., Kramer, D., & Witchel, H. J. (2025). Navigating the frontier of AI-assisted student assignments: challenges, skills, and solutions. *Advances in Physiology Education*, *49*(3), 633–639. https://doi.org/10.1152/advan.00253.2024

Farazouli, A., Cerratto-Pargman, T., Bolander-Laksov, K., & McGrath, C. (2024). Hello GPT! Goodbye home examination? An exploratory study of AI chatbots impact on university teachers' assessment practices. *Assessment and Evaluation in Higher Education*, *49*(3), 363–375. https://doi.org/10.1080/02602938.2023.2241676

Fenton, A. (2025). Reconsidering the use of oral exams and assessments: An old way to move into a new future. *Educational Researcher*, *54*(7), 430–436. https://doi.org/10.3102/0013189X251333638

Han, F., Deng, M., Yan, T., & Wang, H. (2026). AI-based educational interventions for enhancing cognitive learning processes in students with disabilities: A meta-analysis. *Learning and Individual Differences*, 127. https://doi.org/10.1016/j.lindif.2026.102876

Jalali, M. S., & Akhavan, A. (2024). Integrating AI language models in qualitative research: Replicating interview data analysis with ChatGPT. *System Dynamics Review*, *40*(3), e1772. https://doi.org/10.1002/sdr.1772

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., et al. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, *103*. https://doi.org/10.1016/j.lindif.2023.102274

Kickbusch, S., Ashford-Rowe, K., Kemp, A., Boreland, J., & Huijser, H. (2025). Beyond detection: Redesigning authentic assessment in an AI-mediated world. *Education Sciences*, *15*(11), 1–17. https://doi.org/10.3390/educsci15111537

Kofinas, A. K., Tsay, C. H. H., & Pike, D. (2025). The impact of generative AI on academic integrity of authentic assessments within a higher education context. *British Journal of Educational Technology*, *56*(6), 2522–2549. https://doi.org/10.1111/bjet.13585

Luo, J., Zheng, C., Yin, J., & Teo, H. H. (2025). Design and assessment of AI-based learning tools in higher education: A systematic review. *International Journal of Educational Technology in Higher Education*, *22*(1). https://doi.org/10.1186/s41239-025-00540-2

Mahapatra, S. (2024). Impact of ChatGPT on ESL students' academic writing skills: A mixed methods intervention study. Smart *Learning Environments, 11*(9). https://doi.org/10.1186/s40561-024-00295-9

McGrath, C., Farazouli, A., & Cerratto-Pargman, T. (2025). Generative AI chatbots in higher education: A review of an emerging research area. *Higher Education*, *89*(6), 1533–1549. https://doi.org/10.1007/s10734-024-01288-w

Newton, P., & Xiromeriti, M. (2024). ChatGPT performance on multiple choice question examinations in higher education. A pragmatic scoping review. *Assessment and Evaluation in Higher Education, 49(6),* 781–798. https://doi.org/10.1080/02602938.2023.2299059

OpenAI. (2026). *ChatGPT-5.5*. https://chatgpt.com/

Quan, Y. (2025). The role of AI-driven feedback in fostering growth mindset and engagement: A self-determination theory perspective. *Learning and Motivation, 92*(1), 1–14. https://doi.org/10.1016/j.lmot.2025.102192

Rawls, J., Wilsker, A., & Rawls, R. S. (2015). Are you talking to me? On the use of oral examinations in undergraduate business courses. *Journal of the Academy of Business Education*, *16*, 22–33.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*(2), 119–144. https://doi.org/10.1007/BF00117714

Su, F., O'Dea, M., Wood, M., & Seidl, E. (2026). Embracing a new world: Authentic assessment designs in an age of generative artificial intelligence (GenAI). *Perspectives: Policy and Practice in Higher Education*, *3108*, 1–12. https://doi.org/10.1080/13603108.2025.2601741

Urban, M., Děchtěrenko, F., Lukavský, J., Hrabalová, V., Svacha, F., Brom, C., & Urban, K. (2024). ChatGPT improves creative problem-solving performance in university students: An experimental study. *Computers and Education, 215*. https://doi.org/10.1016/j.compedu.2024.105031

Van Slyke, C., Johnson, R. D., & Sarabadani, J. (2023). Generative artificial intelligence in information systems education: Challenges, consequences, and responses. *Communications of the Association for Information Systems, 53*, 1–21. https://doi.org/10.17705/1CAIS.05301

Wachinger, J., Bärnighausen, K., Schäfer, L. N., Scott, K., & McMahon, S. A. (2024). Prompts, pearls, imperfections: Comparing ChatGPT and a human researcher in qualitative data analysis. *Qualitative Health Research*. https://doi.org/10.1177/10497323241244669

Ward, M., O'Riordan, F., Logan-Fleming, D., Cooke, D., Concannon-Gibney, T., Efthymiou, M., & Watkins, N. (2024). Interactive oral assessment case studies: An innovative, academically rigorous, authentic assessment approach. *Innovations in Education and Teaching International, 61(5),* 930-947. https://doi.org/10.1080/14703297.2023.2251967

Wiggins, G. (1990). The case for authentic assessment. *Practical Assessment, Research & Evaluation, 2*(1), 2. https://doi.org/10.7275/ffb1-mm19

Zhao, X., Cox, A., & Chen, X. (2025). The use of Generative AI by students with disabilities in higher education. *Internet and Higher Education*, *66*, 1–11. https://doi.org/10.1016/j.iheduc.2025.101014

APPENDIX A:

Survey Questions

Student Consent of Study Participation

1. I have read the study information and understand that my participation in this survey is voluntary and will not affect my grade on this course.

☐ Yes, I agree to participate
☐ No, I do not agree to participate
(If "No," you may stop the survey.)

2. [Identifier] Please type in your 6-digit identifier here _____

Demographic Questions

Please answer the following questions. You may select "Prefer not to say" if you prefer not to disclose it.

3. Age

☐ _____years old
☐ Prefer not to say

4. Gender

☐ Female
☐ Male
☐ Non-binary / another gender
☐ Prefer not to say

5. First-generation college student (neither parent/guardian completed a 4-year college degree)

☐ Yes
☐ No
☐ Prefer not to say

6. Academic level

☐ First-year undergraduate
☐ Second-year undergraduate
☐ Third-year undergraduate
☐ Fourth-year or above undergraduate
☐ Graduate / professional student

7. Major / primary field of study: _____

Main Survey Questions

Response Scale: Please indicate your agreement using this scale.

1 = Strongly disagree 2 3 4 = Neither agree nor disagree 5 6 7 = Strongly agree

Item #	Post Survey Questions	Likert Scale Response (1 to 7)						
	Academic Standing							
1	I feel confident in my overall academic ability.	1	2	3	4	5	6	7
2	I have strong analytical reasoning skills (for example, breaking down problems, interpreting data, and drawing logical conclusions).	1	2	3	4	5	6	7
3	I communicate clearly and effectively in spoken situations.	1	2	3	4	5	6	7
4	I communicate clearly and effectively in written work.	1	2	3	4	5	6	7
5	I perform well under time pressure.	1	2	3	4	5	6	7
	Generative AI Familiarity							
6	I understand how to use AI tools appropriately.	1	2	3	4	5	6	7
7	I frequently use AI tools for my coursework.	1	2	3	4	5	6	7
8	I have a positive attitude toward using AI tools in my learning.	1	2	3	4	5	6	7
	Learning Outcomes – Overall							
9	I have become confident in my understanding of this topic.	1	2	3	4	5	6	7
10	I have become knowledgeable about this topic.	1	2	3	4	5	6	7
11	I have become confident in explaining this topic clearly to others.	1	2	3	4	5	6	7
	Bloom's Taxonomy Learning Outcomes							
12	The assessment helped me understand key concepts of the topic.	1	2	3	4	5	6	7
13	The assessment helped me analyze my reasoning and recognize gaps in my knowledge.	1	2	3	4	5	6	7
14	The assessment allowed me to justify and defend my understanding and reasoning.	1	2	3	4	5	6	7
15	The assessment pushed me to brainstorm new ideas.	1	2	3	4	5	6	7
16	The assessment helped me practice applying the course content knowledge to different situations.	1	2	3	4	5	6	7
	Assessment Experience							
17	The assessment was a fair way to evaluate my understanding of the material.	1	2	3	4	5	6	7
18	The assessment was well aligned with what we covered in the course.	1	2	3	4	5	6	7
19	The assessment was a stressful experience.	1	2	3	4	5	6	7
20	The assessment motivated me to prepare for deeper learning rather than just memorizing facts.	1	2	3	4	5	6	7
21	The assessment allowed me to demonstrate what I really know.	1	2	3	4	5	6	7
Open-Ended 1	Describe one example where an AI-generated follow-up question led you to change, refine, or extend your original written answer. What did you learn from that experience?							
Open-Ended 2	Overall, how did the written test plus AI-generated follow-up questions affect your learning, if at all?							
Open-Ended 3	If this quiz format were used again, what one change would most improve your learning experience or learning outcomes?							

APPENDIX B:
Examples of the open-ended question, a student response, and the follow-up question based on the response

A. The open-ended question

A marketing team ran an A/B test on two email subject lines to increase click-through rate (CTR).

- Variant A (control) was sent to 3,200 users and had a CTR of 5.1%.
- Variant B (treatment) was sent to 3,050 users and had a CTR of 5.4%.
- The statistical analysis reported a p-value of 0.028.

You are asked whether the company should roll out Variant B to all customers, or keep testing.

Please include:

1. A clear decision: (1) end campaign & roll out Variant B or (2) keep testing.
2. Discussion of how the p-value, the absolute CTR difference, and the sample sizes together affect how confident you are in the result.
3. Consider practical marketing consequences (e.g., projected ROI, customer experience, and the risk/cost of acting on a small or uncertain improvement).

Note: for this question, there will be a follow-up question that you need to respond orally (i.e., speaking). You will have **90 seconds starting automatically**. You do NOT have to speak for full 90 seconds. Take 15 - 30 seconds to think before responding, and stop whenever you are done. You will be graded based on the verbal content you speak and will not be punished for the time of silence.

B. An exemplary response from students

"I would lean toward keeping the test running rather than immediately rolling out Variant B.
Even though the p-value of 0.028 suggests the result is statistically significant at the 5% level, the actual lift in CTR is only 0.3% (5.4% vs. 5.1%), which is pretty small. With sample sizes around ~3,000 per group, the test has decent power, but not enough to fully rule out variability—especially for such a marginal gain.
From a marketing standpoint, this comes down to practical impact vs. statistical significance. A 0.3% lift might not meaningfully impact revenue or engagement unless scaled across a very large audience. There's also risk in rolling out a "winner" that doesn't consistently outperform over time or across different segments.
Given that, I'd want to either:
Run the test longer to confirm consistency, or test additional subject line variations to see if we can drive a more meaningful lift
So overall, while the data is promising, it's not strong enough (from a business impact perspective) to confidently justify a full rollout yet."

C. The follow-up question based on that response

Great analysis. Given the "marginal gain," what specific business factors would justify rolling out Variant B immediately? Rubric: 1pt for conceptual understanding of business context, 1pt for logical reasoning.