Harnessing Objective Questions to Enhance Writing Pedagogy and AI Literacy in the Digital Age

Yuan Tao[1], Xian Liao[2]

Chinese Language Centre, The Hong Kong Polytechnic University

Faculty of Humanities, The Education University of Hong Kong

Author Note

https://orcid.org/0009-0005-5089-8255

https://orcid.org/0000-0002-4470-8436

We declare no conflicts of interest.

Correspondence concerning this article should be addressed to ctctytao@polyu.edu.hk

The integration of artificial intelligence (AI) technologies, particularly Artificial Intelligence-Generated Content (AIGC), has revolutionized language education. In traditional L1 writing pedagogy, writing has long been regarded as a highly subjective, constructive domain. Historically, writing instruction has relied heavily on subjective assessments such as essays and open-ended responses to nurture students’ reflective processes, argumentative thinking, and creative problem-solving (Graham & Perin, 2007). For example, when analyzing literary works, educators encourage students to explore unique themes and viewpoints, emphasizing original interpretation.

While generative AI offers unprecedented convenience in content creation, it has introduced profound pedagogical challenges. Excessive reliance on AIGC tools frequently obscures gaps in students' real capabilities and threatens to erode higher-order cognitive engagement. This phenomenon prompts a critical reflection: traditional subjective writing tasks alone may be insufficient or too cognitively demanding to rapidly equip students with the critical discernment necessary for the AI era. Conversely, objective questions (OQs)—long overlooked in writing pedagogy and dismissed as limited to assessing basic lexical-grammatical knowledge (Haladyna & Rodriguez, 2013)—hold untapped potential as formative scaffolding tools.

This research posits that while OQs cannot replace subjective writing, they can serve as a vital complementary mechanism. By integrating OQs into process and genre-based writing approaches, this study aims to address the unique challenges posed by AIGC, provide effective scaffolding for mastering complex structural writing skills in L1 Chinese, and expand AI literacy frameworks to include ethical evaluation and bias detection.

Literature Review: Objective Questions, Writing Pedagogy, and AI Literacy

Writing Pedagogy and the Role of Objective Assessment

Contemporary writing pedagogy is heavily informed by process and genre approaches (Flower & Hayes, 1981; Hyland, 2007). The process approach emphasizes recursive stages of planning, drafting, and revising, viewing writing as a cognitive problem-solving activity. Conversely, the genre approach focuses on the communicative purpose, social context, and structural conventions of specific text types, providing learners with explicit rhetorical scaffolding. Within these established paradigms, however, objective questions (OQs) are rarely utilized as instructional tools. In the domain of writing assessment, proficiency is typically evaluated through direct, performance-based measures such as essays and portfolios. When OQs are mentioned in writing assessment literature—often termed "indirect assessment"—they are frequently criticized as lacking construct validity. Consequently, they are traditionally relegated to testing basic grammar, syntax, and vocabulary rather than authentic writing ability (Haladyna & Rodriguez, 2013, p. 278). As a result, mainstream writing pedagogy has largely ignored the potential of OQs to foster logical analysis and structural organization.

This traditional skepticism is perhaps best summarized by Haladyna and Rodriguez (2013, p. 265), who point out, "By definition, what is measured using the traditional SR format does not seem to be writing." However, the current landscape has fundamentally changed; because AIGC frequently generates the initial text, it requires humans to evaluate and make selections. Consequently, writing ability has often shifted toward the capacity for judgment and selection.

Recognizing this shift, it becomes clear that while traditional writing pedagogy has often leaned toward a "performance-only" approach—frequently rejecting selected-response items as inauthentic—modern assessment frameworks demonstrate their critical value. As Deane, Sabatini, and Fowles (2012) illustrate in their Cognitively-Based Assessment of, for, and as Learning (CBAL) framework, selected-response items can be highly effective when used as "lead-in tasks." For instance, objective questions requiring students to evaluate peer summaries or determine the validity and strength of specific evidence successfully measure higher-order reading and critical thinking skills. These cognitive prerequisites are indispensable for complex writing but are often difficult to isolate and measure through extended performance tasks alone (Deane et al., 2012, pp.

88-91). Further supporting the use of such tools in language education, Church (2023) discusses the development and results of formative, multiple-choice question (MCQ)-style online tests aimed at developing critical thinking and higher-order aspects of academic writing. This research validates the effectiveness of MCQs in scaffolding academic writing instruction.

Because of the historical gap in writing scholarship regarding these tools, it is also necessary to look toward interdisciplinary education to understand the true capacity of objective assessments. Studies in fields such as medical education demonstrate that carefully constructed multiple-choice questions (MCQs) can, in fact, effectively assess higher-order thinking skills, such as application and analytical reasoning (Javaeed, 2018; Palmer & Devitt, 2007). More recently, scholarship in English for Specific Purposes (ESP) has built upon this, showing that scenario-based MCQs can successfully tap into Bloom's higher-order cognitive processes (Lenchuk & Ahmed, 2021).

Despite these proven benefits in other disciplines and emerging writing frameworks, there remains a pressing need to apply these insights to L1 Chinese writing pedagogy. By embedding OQs within L1 Chinese writing instruction, educators can create an intermediate cognitive step that isolates and reduces the cognitive load required to understand complex structural conventions before students are expected to produce their own texts.

AI Literacy in the Language Classroom

In the context of AI-driven education, writing pedagogy faces unprecedented challenges. As Alhajji (2024) notes, the widespread adoption of artificial intelligence has fundamentally transformed how students engage with writing tasks. However, this transformation highlights a critical assessment gap: writing performance with AI assistance often differs significantly from independent writing ability, allowing AI to obscure gaps in students' real capabilities. Despite this shift, traditional evaluation metrics—such as linguistic quality, coherence, and organization—

remain the core indicators of writing. These traditional dimensions have not fully adapted to the AI era, lacking criteria to assess the reasonable use of AI tools or the effectiveness of human-AI collaboration (Jin et al., 2025).

As Artificial Intelligence-Generated Content (AIGC) becomes increasingly accessible, a secondary issue has emerged: many students passively consume AI outputs without evaluating their quality, logical rigor, ethical implications, or inherent biases (Zhai et al., 2024). While AI offers substantial benefits to higher education, such as personalized learning and enhanced data analysis, excessive reliance on these tools bypasses in-depth cognitive engagement. This poses a significant risk of eroding students' higher-order cognitive skills, including logical reasoning, critical thinking, and independent problem-solving (Cotton et al., 2024; Lodge et al., 2023).

To address the limitations of traditional assessment and the risks of passive AI consumption, this study situates its pedagogical model within established AI literacy frameworks. Ng et al. (2021) conceptualize AI literacy through four dimensions: knowing and understanding AI, applying AI, evaluating AI, and AI ethics. While traditional writing metrics focus on text production, integrating AI literacy requires educators to expand their pedagogical constructs. By incorporating tasks that target critical evaluation—such as bias detection, identifying hallucinations, and ethical decision-making—educators can effectively measure and foster responsible human-AI collaboration alongside traditional writing skills (Jin et al., 2025; Ng et al., 2021).

Theoretical Framework and Innovative Approach

This research connects these gaps through three theoretical lenses:

1. **Constructivist Learning and Genre Theory**: OQs are designed to target higher-order genre analysis. By comparing human-written and AIGC texts, students actively construct knowledge of genre conventions rather than passively receiving them (Hyland, 2007).

2. **Cognitive Load Theory** (Sweller et al., 2019): OQs act as scaffolding tools, isolating complex structural elements (e.g., thesis formulation, MECE frameworks) to reduce intrinsic cognitive burden before full-scale writing.

3. **Comprehensive AI Literacy Frameworks**: The study maps OQs to established AI literacy constructs (Ng et al., 2021), expanding beyond mere output evaluation to encompass ethical use and bias detection.

Unlike prior approaches, this study implements three innovative strategies: (1) integrating writing-oriented thinking frameworks (e.g., MECE, SWOT) into OQ design; (2) designing scenario-based OQs that simulate real-world AIGC evaluation tasks; and (3) combining OQs with subjective writing tasks to create a balanced learning cycle.

Methodology

Participants and Context

To evaluate the efficacy of the proposed model, an 8-week empirical study was conducted at a university in Hong Kong. Participants were 75 undergraduate students enrolled in a university-level L1 Chinese business writing course. The sample comprised native Chinese-speaking students from diverse academic disciplines. (Note: While the instruction and assessment were conducted in Chinese, the conceptual framework and examples presented in this paper have been translated into English for wider dissemination).

The demographic profile of the participants and the course context are summarized in Table 1.

Table 1

Participant Profile and Course Context

Category	Description
Target Participants	75 Undergraduate Students
Age & Year of Study	Average age 19–20; Year 2, Semester 2
Background	Local students (JUPAS enrolled)
Language Profile	Native Cantonese speakers; highly accustomed to Standardized Written Chinese (conceptually approximate to Putonghua).
Course Context	Aims to enhance Chinese communication skills tailored to workplace requirements in the construction and environment sectors.
Core Objectives	By the end of the course, students are expected to master: <1>. **Accuracy:** Precise professional Chinese expression. <2>. **Cognitive Structuring:** Effective organization of content and logical reasoning. <3>. **Genre Mastery:** Application of various written genres for professional and academic purposes.

Test-Item Development and Validation

A comprehensive bank of 40 scenario-based Objective Questions (OQs) was developed. These items integrated writing frameworks (e.g., MECE, SWOT, 5W2H1E, SMART) and AIGC evaluation scenarios (e.g., identifying logical fallacies and biases in AI-generated text). These items consisted of 30 multiple-choice (MC) questions and 10 True/False (T/F) questions. To ensure validity, the items were reviewed by two writing assessment experts and piloted with a non-participating cohort (N=30). Reliability testing yielded a strong Cronbach's alpha of .84, indicating high internal consistency.

Research Procedure and Instructional Workflow

The intervention followed a structured, five-step cyclical workflow (see Figure 1) over the 8-week semester. Throughout the semester, the course covered essential workplace writing genres (such as official correspondence, memos, proposals, and analytical reports) integrated with advanced cognitive frameworks.

The macro-level workflow was structured as follows:

1. **Pre-assessment:** Students completed a baseline subjective writing task (outline of a customer complaint reply letter) and an AIGC evaluation test to identify knowledge gaps.
2. **Framework-Based Skill Building:** OQs integrated with thinking frameworks were introduced. For example, students answered MECE-related questions to learn idea organization without overlap.
3. **Scenario-Based AIGC Evaluation Tasks:** Students engaged with OQs simulating real-world AI outputs, practicing bias detection, evidence evaluation, and logical critique.
4. **Subjective Writing Application:** Students applied the scaffolded skills to subjective writing tasks (e.g., a customer complaint reply letter).
5. **Iterative Refinement and Reflection:** Teacher feedback on both OQ performance and subjective writing closed the learning loop.

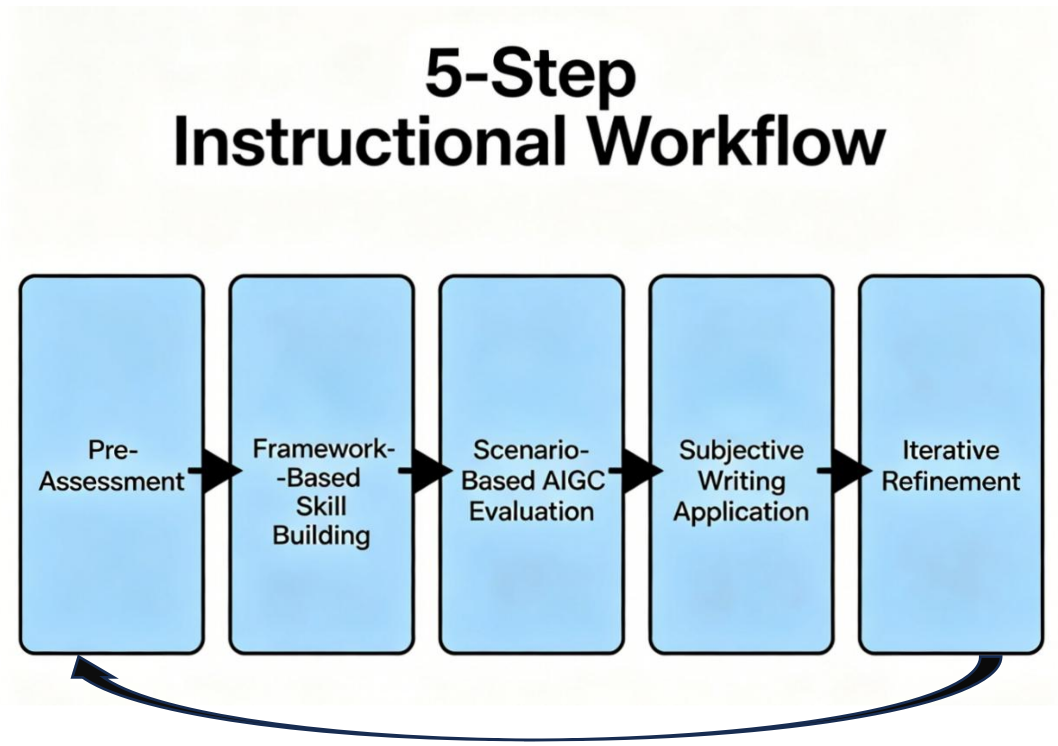
Micro-Level Lesson Implementation:

While the five steps dictate the overall semester progression, individual class sessions were designed to seamlessly blend cognitive planning, objective testing, and AI-assisted writing. For instance, during a session focused on writing a "Customer Complaint Reply Letter," the instructional flow was executed as follows:

- **Step 1 (Conceptualization):** Students were first required to brainstorm and outline the structural framework of a reply letter independently.
- **Step 2 (Targeted Practice):** Students completed a set of specific OQs directly related to the stylistic and logical requirements of complaint reply letters.
- **Step 3 (AI-Assisted Drafting & Human Editing):** Armed with the structural framework and knowledge from the OQs, students designed prompt inputs for the AI. They then critically evaluated and manually revised the AI-generated text to finalize their writing assignment.

Figure 1

Instructional Workflow



Scoring Procedures and Data Analysis

Essays were double-blind marked using a standardized analytic writing rubric focusing on logical structure and use of evidence. Inter-rater reliability was high (Cohen's $\kappa = .82$). AIGC evaluation skills were measured via the standardized OQ test. Pre- and post-test scores were analyzed using paired-samples t-tests to determine statistical significance. Qualitative data from student reflections ($N = 75$) and classroom observations were also collected.

Analysis of Actual Application Effect

Quantitative Results

Statistical testing revealed significant learning gains following the 8-week intervention. For logical coherence and structure in subjective essays, students' average scores improved from 65.0 ($SD = 8.2$) to 82.0 ($SD = 6.4$), indicating a statistically significant increase, $t(119) = 14.3$, $p < .001$. In the AIGC evaluation domain (identifying logical fallacies, biases, and structural weaknesses in AI text), accuracy improved dramatically from a pre-test average of 28.0 ($SD = 12.1$) to 79.0 ($SD = 9.8$), $t(119) = 22.6$, $p < .001$. Furthermore, accuracy in applying frameworks (e.g., MECE/SWOT) increased by 41.8%, and effective evidence use in writing increased by 31.6%.

Qualitative Results

Student reflections corroborated the quantitative findings. Approximately 79% of students explicitly noted that Objective Questions (OQs) lowered the barrier to understanding abstract structural rules, stating that OQs "made complex writing rules easier to understand."

Specifically, students highlighted how analytical OQs enhanced their cognitive planning before and during AI engagement.

Student 1:

「本次訓練中，分析能力類題目對我的 AI 輔助寫作提升最顯著。透過案例拆解、多維度評估與邏輯梳理，我學會更系統地提煉關鍵問題、分類論點，並讓 AI 輔助生成結構化、有層次的內容，避免零散與主觀。」

(Translation: "During this training, the analytical skills questions provided the most significant improvement *to my AI-assisted writing. Through case deconstruction, multi-dimensional evaluation, and logical organization, I learned to systematically extract key issues, categorize arguments, and use AI to generate structured, multi-layered content, thereby avoiding fragmentation and subjectivity.")*

Student 2:

「分析類題目對我寫作能力的提升最為顯著。這類題目讓我利用 AI 進行結構化拆解與邏輯檢視，讓我學會如何利用提示詞挖掘深度論據，將平鋪直敘轉化為多層次的嚴謹推論。」

(Translation: "Analytical questions had the most obvious impact on improving my writing skills. These questions enabled me to use AI for structural deconstruction and logical review, teaching me how to utilize prompts to unearth in-depth arguments and transform flat narratives into multi-layered, rigorous reasoning.")

Furthermore, discussions and OQ practices focused on AI bias and logical evaluation helped students transition from passive AIGC consumers to critical editors. Many noted that they learned to consciously "check AI outputs for specific details and logic," rather than accepting AI-generated drafts at face value. The scenario-based OQs trained them to identify subtle hallucinations, logical fallacies, and inappropriate tones in AI-generated workplace correspondence (see Appendix for specific MC and T/F question examples used for this training).

However, critical challenges emerged: some students reported that OQs felt "mechanic" and did not assist in developing their personal writer identity. Additionally, designing high-quality OQs targeting higher-order thinking was time-consuming for educators.

Discussion: The Limits and Potential of Objective Questions

While the study contributes meaningful pedagogical insights, it is crucial to moderate the scope of claims regarding OQs as a panacea for AI-related writing challenges. The findings indicate that OQs are highly effective for scaffolding analytical thinking, structural organization, and foundational AI evaluation. They successfully break down the cognitive load required to master genre conventions.

This observation aligns seamlessly with Cognitive Load Theory (Sweller et al., 2019), which suggests that isolating complex rules into manageable tasks reduces extraneous cognitive burden, thereby freeing up students' working memory for higher-order schema acquisition. Furthermore, treating OQs as a form of instructional scaffolding resonates with Vygotsky’s concept of the Zone of Proximal Development (ZPD) (Vygotsky, 1978, p. 88). It corroborates recent literature on AI-assisted education, which emphasizes that providing explicit, structured affordances before open-ended AI engagement enables learners to more effectively navigate and control the complexities of human-AI co-creation (Yan et al., 2024). However, OQs do not—and cannot—adequately address holistic writing competencies such as source synthesis, argument construction, creative expression, or writer identity. Writing remains fundamentally a constructive, subjective act. As noted in the qualitative findings, OQs cannot measure a student's unique writing style or vivid narrative imagery. Attempting to design OQs that "test creativity" is a paradoxical endeavor that risks overemphasizing technical proficiency.

Therefore, educators should utilize OQs to handle the heavy lifting of structural and logical scaffolding, thereby freeing up cognitive resources and classroom time. Teachers can then focus their personalized feedback on students' creative processes, voice, and complex synthesis in subjective tasks. Furthermore, integrating AI literacy requires establishing standardized rubrics for AIGC evaluation. Educators must guide students not only in assessing the structural validity of AI outputs but also in questioning the ethical provenance, potential biases, and cultural assumptions embedded within generative AI.

Conclusion and Future Directions

Integrating objective questions into L1 Chinese writing pedagogy in the AIGC era offers a pragmatic, theoretically grounded approach to scaffolding complex structural skills and developing AI literacy. The empirical data from this study demonstrates that when combined with established writing frameworks and AI evaluation scenarios, OQs significantly enhance logical coherence and digital discernment. By reducing time spent on grading basic structural issues, this methodological shift actually enhances the human element of teaching, allowing educators to focus on higher-order creative feedback.

Future development priorities should focus on building an integrated teaching ecosystem where OQ banks are dynamically updated to reflect evolving AI models and curriculum standards. Additionally, systematic teacher training is essential to help educators seamlessly bridge the gap between OQ-based structural scaffolding and authentic, subjective writing practice. By recognizing both the distinct advantages and the inherent limitations of objective questions, educators can leverage them effectively to prepare students for the rigorous demands of business writing in an AI-driven landscape.

References

Alhajji, R. (2024). Assessing the influence of AI on modern student writing standards: An educators' perspective. *Research Journal in Advanced Humanities, 5*(3), 112-125. https://doi.org/10.58256/5tpepv73

Church, Anneen. (2023). Optimising academic writing assessment during Covid-19: The development multiple choice tests to develop writing without writing. *Perspectives in Education*. 41（3）. 23-38. https://doi.org/10.38140/pie.v41i3.6804

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

Deane, P., Sabatini, J., & Fowles, M. (2012). Rethinking K-12 writing assessment to support best instructional practices. In Writing assessment in the 21st century (pp. 83-102). https://doi.org/10.37514/PER-B.2012.0452.2.05

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication, 32*(4), 365–387. https://doi.org/10.2307/356600

Ginto, C., Wang, L., & Patel, S. (2025). Cognitive offloading or cognitive overload? How AI alters the mental architecture of coping. Frontiers in Psychology, 16. https://doi.org/10.3389/fpsyg.2025.1699320

Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent *students*. *Journal of Educational Psychology, 99*(3), 445–476. https://doi.org/10.1037/0022-*0663*.99.3.445

Haladyna, T. M., & Rodriguez, M. C. (2013). *Developing and validating test items*. Routledge. https://doi.org/10.4324/9780203850381

Hyland, K. (2007). Genre pedagogy: Language, literacy and L2 writing instruction. *Journal of* Second *Language Writing, 16*(3), 148–164. https://doi.org/10.1016/j.jslw.2007.07.005

Javaeed, A. (2018). Assessment of higher ordered thinking in medical education: Multiple choice questions and modified essay questions. *MedEdPublish, 7*(2), 60. https://doi.org/10.15694/mep.2018.0000128.1

Jin, Y., Yang, K., Martinez-Maldonado, R., Gašević, D., & Yan, L. (2025). The Agency Gap: How Generative AI Literacy Shapes Independent Writing after AI Support. *arXiv preprint*. https://arxiv.org/abs/2507.04398

Lenchuk, I., & Ahmed, A. (2021). Tapping into Bloom taxonomy’s higher-order cognitive processes: The case for multiple choice questions as a valid assessment tool in the ESP classroom. *Arab World English Journal (AWEJ), Special Issue on Covid 19 Challenges*, 160–171. https://doi.org/10.24093/awej/covid.12

Lodge, J. M., Thompson, K., & Corrin, L. (2023). Mapping out a research agenda for generative artificial intelligence in tertiary education. *Australasian Journal of Educational Technology, 39*(2), 1-16. https://doi.org/10.14742/ajet.8695

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence, 2*, 100041. https://doi.org/10.1016/j.caeai.2021.100041

Palmer, E. J., & Devitt, P. G. (2007). Assessment of higher order cognitive skills in undergraduate education: Modified essay or multiple choice questions? *BMC Medical Education, 7*(1), 49. https://doi.org/10.1186/1472-6920-7-49

Sweller, J., van Merriënboer, J. J., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review, 31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5

Weigle, S. C. (2002). *Assessing writing*. Cambridge University Press. https://doi.org/10.1017/CBO9780511732997

Yan, D., Wang, M., & Chiu, T. K. F. (2024). Scaffolding human-AI co-writing: An investigation of prompt engineering frameworks for academic writing. Computers and Education: Artificial Intelligence, 6, 100201. https://arxiv.org/html/2512.18388v2

Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: A systematic review. *Smart Learning Environments, 11*, 28. https://doi.org/10.1186/s40561-024-00316-7

Appendices

Appendix 1. Sample MC for Argumentative Writing Logic

(Referenced content translated into English for illustration)

Scenario: "There is a growing movement to promote plant-based diets. Proponents argue that these diets are more sustainable as they require fewer resources to produce compared to meat-based diets. They also point out the health benefits associated with consuming more plants, such as a reduced risk of heart disease. On the other hand, some people claim that a balanced diet can include meat and that sudden shifts to plant-based diets may lead to nutritional deficiencies."

Question: Which statement best represents the main argument of the passage?

A. Plant-based diets are the only sustainable option.

B. There are both pros and cons to plant-based diets.

C. Meat should be completely removed from diets.

D. People who oppose plant-based diets are misinformed.

Correct Answer: B *(Rationale: The passage presents both supporting arguments (sustainability, health benefits) and counterarguments (nutritional deficiencies) of plant-based diets, indicating that there are pros and cons. Option A is incorrect because it uses absolute language ("only") not supported by the text. Option C contradicts the passage's acknowledgment of balanced diets including meat. Option D is unsupported as the passage does not dismiss opponents as misinformed. This question tests students' ability to identify the core argument by analyzing supporting and counterarguments, a key skill for argumentative writing.)*

Appendix 2. Sample MC for AIGC Evaluation

AI-Generated Content:

Review A: "This smartphone is amazing. It has a great camera, a long-lasting battery, and a sleek design. Everyone should buy it."

Review B: "The new smartphone offers a 48-megapixel camera that can capture stunning details even in low-light conditions. Its battery life can last up to two days on a single charge, according to laboratory tests. The design is not only sleek but also ergonomic, making it comfortable to hold. Considering its price-to-performance ratio, it's a great option for users who prioritize both functionality and style."

Question: Which review is more persuasive and why?

A. Review A, because it is short and to the point.

B. Review B, because it provides specific details and logical reasoning.

C. Review A, because it uses positive language.

D. Review B, because it mentions the price.

Answer: B *(Rationale: Persuasive writing requires specific evidence and logical reasoning to support claims. Review B provides concrete details (48-megapixel camera, two-day battery life, ergonomic design) and links features to user needs, making it more persuasive. Review A is overly general and lacks supporting evidence. Option A is incorrect because brevity alone does not equal persuasiveness. Option C is incorrect because positive language without substance is ineffective. Option D is incorrect because mentioning the price is a minor detail, not the main reason for persuasiveness. This question trains students to evaluate AI-generated content based on key persuasive writing criteria.)*

Appendix 3. Sample MC for MECE Framework Application

Question: Which of the following sets of causes for urban traffic congestion best follows the MECE principle?

A. "Increasing number of cars, traffic accidents, bad weather"

B. "Infrastructure problems, high population density, lack of public transportation, and increasing number of private vehicles"

C. "Traffic lights, people not following traffic rules, road construction"

D. "Car ownership, traffic management, and accidents"

Answer: B *(Rationale: The MECE principle requires categories to be mutually exclusive (no overlap) and collectively exhaustive (covering all relevant causes). Option B meets this standard: infrastructure problems, high population density, lack of public transportation, and increasing private vehicles are distinct categories that comprehensively address major causes of traffic congestion. Option A has overlapping categories (traffic accidents and bad weather can both be temporary disruptions) and misses key causes. Option C is incomplete and overlapping (traffic lights and road construction are both infrastructure-related). Option D is incomplete and overlapping (accidents can be related to traffic management). This question tests students' understanding of structured idea organization, a critical skill for essay outlining.)*

Appendix 4. Sample MC for Evaluating AI Logic and Specificity

Scenario: You requested an AI to draft a response letter to a client who complained about a two-week delay in a construction project. The AI generated the following excerpt: *"We apologize for the delay. The weather was bad, and some materials arrived late, which is normal in this industry. We will try our best to finish soon."*

Question: From the perspective of professional crisis communication and logical completeness (using the 5W2H1E framework), what is the most critical flaw in the AI's output?

A) It uses overly formal language that distances the client.

B) It lacks specific details regarding the exact causes, concrete remedial actions, and a revised timeline.

C) It apologizes too much, making the company look weak.

D) It does not offer a financial discount to the client.

Answer: B *(Rationale: Professional communication requires specific facts and actionable solutions, not vague excuses.)*

Appendix 5. Sample True/False (T/F) Question for Root-Cause Analysis and Identifying Logical Overlaps (MECE)

Scenario: You asked an AI to generate a root-cause analysis for a recent project delay. The AI listed "Lack of funding," "Budget cuts," and "Insufficient manpower" as the three main causes.

Question: According to the MECE (Mutually Exclusive, Collectively Exhaustive) principle, the AI's response violates the "Mutually Exclusive" rule because "Lack of funding" and "Budget cuts" overlap in meaning. (True / False)

Answer: True. *(Rationale: "Lack of funding" and "Budget cuts" are functionally describing the same issue (financial constraints). Because they overlap, they are not "Mutually Exclusive." A proper AI prompt or human edit should consolidate these into a single "Financial Constraints" category and push the AI to identify other distinct causes to meet the MECE standard.)*

Appendix 6. Sample True/False (T/F) Question for Identifying AI Bias and Tone

Scenario: You asked an AI to write an internal memo evaluating the performance of two contractors. The AI wrote: *"Contractor A (a large multinational firm) is inherently more reliable and produced higher quality work than Contractor B (a local startup), as startups generally lack proper management."*

Question: The AI's evaluation demonstrates a subjective bias rather than an objective, evidence-based assessment. Therefore, a critical editor should remove the generalized assumption about startups and request specific performance metrics (e.g., defect rates, adherence to schedule) for both contractors before finalizing the memo. (True / False)

Answer: True *(Rationale: AI often relies on generalized stereotypes; students must learn to replace subjective bias with objective data.)*

Appendix 7. Pre- and Post-Test Performance Comparison

Evaluation Criterion	**Pre-Test Average Score (100)**	**Post-Test Average Score (100)**	**Improvement Rate**
Logical Coherence in Essays	65.0	82.0	26.15%
AIGC Evaluation Accuracy	28.0	79.0	182.14%
Framework Application (MECE/SWOT)	55.0	78.0	41.82%
Evidence Use in Writing	60.0	79.0	31.67%