

AI-Assisted Gap-Based Research Question Formulation in Higher Education

Authors:

1st Dr. Gulnoza Odilova, 2nd Yassine Ait Lachguer

Head of Department Theory and Practice of Translation

International English Language Teacher, Vosiq International School

Kimyo International University in Tashkent

Author note:

g.odilova@kiut.uz

aitlachguer1745@gmail.com

<https://doi.org/10.5281/zenodo.23357214>

AI-Assisted Gap-Based Question Formulation in Academic Research

Abstract

GenAI, undoubtedly, can polish an academic question, yet its formulation doesn't automatically prove the researcher's ability to defend the gap thoroughly. A conceptual article on gap-based research-question formulation has a critical foundation, treating the operation as a sequence of evidence-judgments: reading the literature, choosing and defending a gap, distilling a research problem and defending the question. Gap justification, under Gen AI intervention, should relate to structured inquiry, process-sensitive appraisal literature, and a specific framework that guides the formulation of a reliably relevant and feasible research question. The proposed framework of research questions connects these operations to evidence of quality, theoretical grounds presented as the basis of research questions, and GenAI intervention. The claim is not only that the process should be assessed against the product (question formulation vs. question defensibility), but it is also what to inspect, at the level of the researcher's competence evidence, from literature review to a defensible question. The framework is, therefore, conceptual and still requires evidence, but the framework provides a basis

for an assessment in which the question is not the only evidence of research competence.

Keywords: generative artificial intelligence research-question formulation, research gaps higher education, assessment and epistemic reasoning, gap-based inquiry, scholarly reasoning

1- Introduction

Forming a defensible question also requires a researcher to read the literature, judge what is still unsettled, explain why the unsettled matters, and state a question answerable in a feasible study. Research-methods scholarship treats question work as an intellectual exercise with theoretical assumptions and choices in it rather than as the mechanical search for an empty cell in the literature (Alvesson & Sandberg, 2013; Sandberg & Alvesson, 2011), and higher-education publications show that gap claims can be implicit and their justifications can be weak (Wald et al., 2024).

GenAI makes this process newly visible. A large language model can suggest candidate gaps, summarize the literature, generate questions, critique scope and propose methods. Kim et al. (2025) apprised the cognitive level of designated questions by teachers; Mikaelyan et al. (2026) evaluated a structured GenAI-supported undergraduate research process involving literature checking and human feedback. The claims that AI-supported question formulation has been neglected would therefore be incorrect. The question is what an AI-assisted question allows an educator to infer about the researcher who submitted it. A polished question could be the product of the student's own distillation, of joint iteration, or of an exchange in which the AI provided the reported gap as well as the question. Identical-looking products can, therefore, differ in their distribution of intellectual work. Having said that, final-question formulation and gap-identification ought to be evaluated distinctly but as related outcomes (Bearman et al., 2024; Messeri & Crockett, 2024; Oliveira et al., 2025).

This article is about gap-based research-question work, about converting a proposed unresolved scholarly issue into an evidence-bearing, important, researchable question. It does not say that AI-supported question work or process-sensitive assessment is new. Its narrower contribution is to fix the evidence that should decide whether the gap-to-question move is defensible. That matter separates, in this setting, the educational purpose of forming a good question from the purpose of developing a question that students can defend in front of the problem it represents.

2- Conceptual Approach and Scope

This focused conceptual synthesis thus connects the line of work on how to build questions and where the gaps are (Alvesson & Sandberg 2013; Sandberg & Alvesson 2011) on how people justify the gaps in higher education research (Wald et al., 2024) on how people make evaluative judgements (Bearman et al. 2024) on how AI can assist in the formation of questions (Kim et al. 2025; Mikaelyan et al., 2026) and on Oliveira et al (2025) DRIVE framework of the learning that is tracked in interactions with GenAI. It attempts to determine what the GenAI-in-education literature measures when the skill is gap-based question formation. It is not a report of every paper in the literature. This is a synthesis, not a systematic review, not a model.

3- Materials and methods

This article takes a critical concept approach to this AI-assisted gap-based research-question construction in higher education. Based on the literature on research-question construction, problematization, inquiry-based learning and generative AI, it questions the link between quality of a question and the quality of the reasoning behind it (Alvesson & Sandberg, 2013; Bearman et al., 2024; Messeri & Crockett, 2024; Sandberg & Alvesson, 2011). Literature interpretation, gap identification, gap evaluation, problem articulation and question construction are five related areas of attention. It identifies three analytical angles, the question itself, the rationale for it and the process by which it is produced. Using literature critically, it develops a conceptual

framework of how to differentiate between plausible research questions and justified gap identification and justified scholarly reasoning. The framework is conceptual and can be tested empirically in future research.

4- Gap-Based Research-Question Formulation as a Multi-Stage Process

4-1 A gap is not yet a research problem

Gap-spotting names something no study examined and identifies the question the gap requires. Sandberg and Alvesson (2011) use the term for common work and note that it can produce incremental questions where a researcher extends an argument without ever challenging it. The warning applies to the GenAI output because a model can generate a plausible missing population, context, variable or method. Few studies examined this population cannot establish that the missing part matters.

A gap is not the same as a research problem. A gap is a theoretical, data or practical inconsistency that does not fit in with literature. Research problem indicates why a gap is important to knowledge, theory, method or practice. Wald and colleagues (2024) noted in 124 journals, the gaps were mentioned solely implicitly without showing why the research problem mattered. They found out that finding a gap is not the same as justifying a study.

GenAI could blur them by creating the fluent gap statement and the matching question in one response, and a pair can appear established even if the claim is not checked against the literature or its significance is not argued, which is an inference risk not every AI-produced gap is wrong. The defensible process checks the alleged gap against the sources and argues why it is worth pursuing (Messerli & Crockett, 2024; Wald et al., 2024).

5- From literature to researchable a question

Gap-based question formulation is reflected in the idea of understanding the literature review, spotting the gaps or inconsistencies, investigating the gaps, stating the

problem and distilling the question, which is the motive for the whole research. The chain of thought is an inquiry-cycle model which involves reasoning, relevance and joining the dots if necessary.

Literature interpretation reports what sources inherently establish and how their findings are related. Gap identification names a possible omission, contradiction, limitation or unresolved relationship. Gap evaluation asks whether the claim is correct, consequential and not already well addressed. Problem articulation explains the intellectual significance of the issue. Question construction and validation then specify what is to be investigated and check whether the question follows from the evidence and is answerable with available methods. These operations cannot be collapsed into one score of “research-question quality,” because each can fail independently.

For example, a student may notice that two studies report different effects of AI feedback and fail to notice that they use different measures or populations. The difference exists, and its cause and importance are unclear. The student cannot yet ask why AI feedback works differently, but can ask what differences exist, and whether they are explained by the measures, task or population. Question construction, therefore, involves interpretation as much as wording formation (Pedaste et al., 2015).

6- GenAI as a Redistribution of Intellectual Work

The role of GenAI should therefore not be treated as one kind of support. It is determined by the way in which the thinking work is shared with the researcher and the system. There are three roles that are pivotal; generation: when the system is used to develop a possible gap or a question; scaffolding: when the system is used to help the student to organize, clarify or enhance an idea that is still being developed; critique: when the system is used to verify assumptions, point out spots or suggest alternative ways of looking at something. All these roles have levels of thinking and academic skill.

Such differentiation is important because a reliably feasible final question does not necessarily indicate that the researcher had independently identified, assessed and justified the gap. As a matter of fact, GenAI can, undoubtedly, enhance question quality and structured inquiry, but this makes it susceptible whether the researcher can reason and judge independently (Kim et al., 2025; Mikaelyan et al., 2026; Bearman et al., 2024). Assessment should therefore be measured not only on the quality of final question, but also on the premise and reasoning through which the question was created, entailing the students' understanding to elaborate on it.

The literature already has evidence for each of these claims. The studies do not all measure the same thing. Kim et al. (2025) investigated AI coaching and what it does to the quality of questions educators pose. In the same reasoning, it is found that AI feedback improves that quality, but they did not show the researcher to explain or defend the gap of research they are inquiring about. Mikaelyan et al. (2026) investigated an eight-step process in which undergraduates explored, checked sources and received human feedback, showing the process feasibility and usefulness, but they did not show that it made students acquire transferrable skills in identifying gaps. The DRIVE framework of Oliveira et al. (2025) is the equivalent concept for GenAI interaction as evidence of reasoning and visible expertise. It supports the argument for investigation process evidence rather than end product evidence only.

7- The problem of inferring competence from the product

A refined question is a product; question-formulation competence is a distinct skill which requires analytic and academic reasoning to elaborate and defend the relevance of a research question. A valid question may demonstrate access to an effective human-AI process exclusively and nothing else. It does not, on its own, establish that the student recognized a real gap and explained why the question belongs to the literature. It is not an argument against collaboration. It limits the inference I can draw from the artefact alone. Messeri and Crockett (2024) warn that AI-supported

science can create the appearance of understanding when fluent or productive output is taken to represent understanding of the reasoning.

Assessment comes in the form of triadic evidence. Product evidence that targets the question clarity, focus, relevance, reliability, and feasibility. Rationale evidence is whether the student can explain and elaborate on the questions. Process evidence is manifested in student's agency and ownership with the question; checking resources, accepting, amending, rejecting, and checking AI assumptions. These variations of evidence are interchangeable, linear, but not causal in the sense that process logic does not prove understanding, and a rationale can be rehearsed or produced by AI Gen unless the student can defend it against established sources and challenges.

When the purpose is research reasoning, product-only assessment may reward effective AI use without establishing whether the student can carry out the operations. A prohibition on AI is not the only alternative to product-only assessment. A defensible alternative assesses the collaborative product against source-grounded justification and evidence of the student's reasoning (Bearman et al., 2024; Oliveira et al., 2025).

8- Results

8-1 Evaluating AI-Mediated Gap-Based Question Formulation

The proposed framework interrelated five concomitant areas: literature-interpretation, gap identification, gap evaluation, problem articulation and question construction. All aligned with three kinds of evidence, product, rationale and process. A well-formulated question can be authentically relevant but still misrepresents the source, overstate the significance or fail to adhere to the literature review guidelines. A source map and a stated rationale, thus, make the chain inspectable and investigate the defensibility of the question. The framework evolves around an epistemic work-gap-based inquiry, not around a developmental sequence or a validated scale, and its categories and proposed uses remain to be reliably tested.

8-2 Propositions for Research and Practice

Four claims follow. The first is that final-question quality and gap-formulation competence should be reported and judged as related but analytically different outcomes. The second is that generation, scaffolding and critique should be reported as functions of use rather than as functions of one intervention, because the functions may shift intellectual work. The third is that source tracing and explicit justification should be assessed where possible, as this would support claims about how a student constructs and supports a question, and the claim should be tested. The fourth is that transfer of gap-formulation reasoning to a new source and topic should be assessed, because passing a familiar supported task does not establish that a student will generate the same reasoning without support on a fresh topic.

These GenAI propositions challenge the premises of coaching, structured inquiry workflow and process-sensitive assessment (Kim et al., 2025; Mikaelyan et al., 2026; Oliveira et al., 2025). Quality of question, usefulness, accuracy of source, justificatory adequacy and transfer should be disentangled in future work.

9- Discussion

9-1 When a Good Research Question Conceals Weak Research Reasoning

The central constraint with using AI to assist in the creation of research questions based on gaps is not that the AI cannot generate a question. It is the fact that the question is founded on scholarly thinking. A good sounding question may be based on a literature-based gap or a poorly considered claim of importance. Identifying a gap does not necessarily mean that it is a worthwhile pursuit. It is the process of turning the gap into a research question that needs critical thinking (Alvesson & Sandberg 2013; Sandberg & Alvesson 2011).

So, we should separate the quality of the question from the reasoning behind it and from the process used to get there. The smooth output of AI can make the process smoother, and this may mislead people to think they understand the topic, while, in

reality, their agency and thinking have been replaced (Messerli & Crockett 2024). In education, we need to ensure that students and researchers can back up the gap that they found explaining why it matters, defend their question and change it when they are given feedback. The value of AI in research lies not in the questions it produces but in the way it can enhance students' own scholarly judgment. We still need studies to establish if such thinking skills transfer to new problems and remain strong when students are no longer receiving AI support.

10- Limitations and Scope

The article is, therefore, a focused conceptual synthesis rather than a systematic review or validation. The comparison against selected related work does not show that nothing else exists to cover the same concerns. A recent framework already provides reasons to examine GenAI interaction as a form of learning, and the Socratic Challenger presents a structured undergraduate research-inquiry workflow. The present contribution is narrower: it ties gap-based question generation to product, rationale and process evidence. Its contribution should therefore be read as a proposed assessment lens for a defined activity rather than as an entirely new theory or validated instrument. Comparison with future work should test the categories against existing assessment instruments and their usefulness in a real research methods setting.

11- Conclusion

Irrefutably, GenAI can produce a plausible research question, but feasibility doesn't linearly signify that the targeted gap is accurate, significant, or understood by the researcher. Existing research already addresses AI-supported question development, structured inquiry, and process-sensitive assessment. This article's narrower contribution is to specify what educators should examine when evaluating the reasoning that connects literature to a gap-based question. By linking five formulation operations to product, rationale, and process evidence, the framework offers a testable way to distinguish a polished question from a defensible one. Its value will depend on subsequent empirical comparison and validation.

References

Alvesson, M., & Sandberg, J. (2013). *Constructing research questions: Doing interesting research*. SAGE Publications.

Bearman, M., Tai, J., Dawson, P., Boud, D., & Ajjawi, R. (2024). Developing evaluative judgement for a time of generative artificial intelligence. *Assessment & Evaluation in Higher Education*, 49(6), 893–905.

<https://doi.org/10.1080/02602938.2024.2335321>

Kim, P., Wang, W., & Bonk, C. J. (2025). Generative AI as a coach to help students enhance proficiency in question formulation. *Journal of Educational Computing Research*, 63(3), 565–586. <https://doi.org/10.1177/07356331251314222>

Messeri, L., & Crockett, M. J. (2024). Artificial intelligence and illusions of understanding in scientific research. *Nature*, 627(8002), 49–58.

<https://doi.org/10.1038/s41586-024-07146-0>

Mikaelyan, A., McKenney, E. A., Mathieson, O. L., & Toprani, D. (2026). The Socratic Challenger: A structured GenAI-assisted workflow for undergraduate

research inquiry. *Frontiers in Education*, 11, Article 1913451.

<https://doi.org/10.3389/feduc.2026.1913451>

Oliveira, M., Zednik, C., Bombaerts, G., Sadowski, B., & Conijn, R. (2025).

Assessing students' DRIVE: A framework to evaluate learning through interactions with generative AI. *Computers and Education: Artificial Intelligence*, 9, Article 100497. <https://doi.org/10.1016/j.caeai.2025.100497>

Pedaste, M., Mäeots, M., Siiman, L. A., de Jong, T., van Riesen, S. A. N., Kamp, E. T., Manoli, C. C., Zacharia, Z. C., & Tsourlidaki, E. (2015). Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educational Research Review*, 14, 47–61.

<https://doi.org/10.1016/j.edurev.2015.02.003>

Sandberg, J., & Alvesson, M. (2011). Ways of constructing research questions: Gap-spotting or problematization? *Organization*, 18(1), 23–44.

<https://doi.org/10.1177/1350508410372151>

Wald, N., Harland, T., & Daskon, C. (2024). The gap statement and justification in higher education research: An analysis of published articles. *European Journal of Higher Education*, 14(2), 308–323. <https://doi.org/10.1080/21568235.2023.2189132>