Tracing GenAI Literacy: Uncovering Student-AI Interaction Patterns in Academic Writing through Epistemic Network Analysis

Angxuan Chen[1], Jiyou Jia[1,*]

[1]Department of Educational Technology, Graduate School of Education, Peking University, Beijing, China

Abstract

As Generative AI (GenAI) becomes integral to education, fostering GenAI literacy is critical. However, current assessments largely rely on self-reported scales, lacking insights into how literacy manifests in actual learning processes. This study leverages Learning Analytics (LA) to bridge this gap. We collected interaction logs from 162 university students engaged in a GenAI-assisted abstract writing task. Using Epistemic Network Analysis (ENA), we modeled and compared the questioning strategies of students with varying GenAI literacy levels. Preliminary results reveal distinct interaction signatures: high-literacy students engage in iterative refinement and strategic questioning, while low-literacy students rely on direct generation commands. This work contributes to the workshop by demonstrating how process data can characterize GenAI literacy, paving the way for data-driven literacy assessment and real-time interventions.

Keywords

GenAI literacy, Learning Analytics, Epistemic Network Analysis, Interaction Patterns, Process Data

1. Introduction

The integration of Generative AI (GenAI) into educational settings has shifted the focus from mere access to "GenAI Literacy"—a multidimensional competence encompassing the understanding, application, evaluation, and ethical use of AI [1]. As GenAI tools like ChatGPT and DeepSeek become ubiquitous, they offer the potential to augment human cognition, but only if learners possess the requisite literacy to navigate them effectively.

Academic writing provides a particularly useful context in which to observe this literacy. Producing an abstract requires learners to identify central claims, select evidence, compress information, and communicate it in an appropriate scholarly register. When GenAI is introduced into this process, learners must additionally decide what to delegate, how to frame requests, whether an output is supported by the source material, and how to incorporate useful suggestions without surrendering authorship. GenAI literacy is therefore expressed through a sequence of judgments rather than through the technical quality of a single prompt. Two learners may obtain similarly fluent outputs while following very different processes: one may critically compare and revise the response, whereas another may accept generated text with little evaluation.

However, a significant gap exists in how GenAI literacy is currently measured. Existing research predominantly treats literacy as a static trait assessed via self-reported questionnaires (e.g., [2]). While useful, these instruments miss the "process" dimension: they cannot reveal how literacy manifests during actual human-AI collaboration. Does a student with high self-reported literacy actually prompt more effectively? Do low-literacy students struggle with hallucination management in real-time?

This distinction between perceived and enacted literacy is important for both research and instruction. Self-reports capture awareness and confidence, but do not show whether knowledge is activated during a task. Likewise, a fluent abstract does not reveal critical construction or uncritical acceptance of generated text. Combining trait measures with process evidence can expose such discrepancies and identify support needs.

First International Workshop on Advancing AI Literacy with Learning Analytics (AI-LIT) @ LAK26, April 2026

*Corresponding author.

jjy@pku.edu.cn (J. Jia)

Interaction logs can provide behavioral evidence for answering these questions because they preserve the order and context of learners' requests. Nevertheless, simple measures such as prompt length, number of turns, or the frequency of a particular command remain incomplete. They describe how often an action occurs but not how actions function together as a strategy. For example, an improvement command may represent thoughtful revision when connected to clarification and verification, but may represent superficial polishing when repeatedly issued without evaluation. Assessing literacy therefore requires attention to relations among behaviors and to the way these relations develop during the task. This relational perspective is well aligned with ENA, which models the co-occurrence structure of coded actions rather than treating each action as an independent event [3].

To address this challenge, we propose a Learning Analytics (LA) approach to characterize literacy through behavioral evidence. By analyzing interaction logs via Epistemic Network Analysis (ENA), this study aims to identify specific behavioral markers that distinguish high from low GenAI literacy users. This research moves beyond self-reported confidence to investigate how students actually structure their interaction with GenAI during complex academic tasks, providing a foundation for objective literacy assessment.

Specifically, the study addresses two research questions: (1) What structural patterns characterize the GenAI interactions of students with high and low literacy levels? (2) How can these patterns inform the design of process-based assessment and adaptive support? By connecting a validated literacy measure with coded interaction sequences, we examine whether theoretical differences in understanding, application, and evaluation are visible in authentic task behavior. The study contributes an empirical account of GenAI literacy as an enacted and temporally organized capability. It also illustrates how learning analytics can move beyond monitoring tool use toward identifying interaction patterns that may support timely, targeted scaffolding.

2. Methodology

2.1. Context and Participants

A total of 162 university students ($M_{\text{age}} = 20.1$) participated in this study. Prior to the task, students completed a validated GenAI Literacy Test [4] assessing technical understanding, interaction skills, and ethical awareness. Based on the average score, participants were classified into a High Literacy Group (HG, $n = 89$) and a Low Literacy Group (LG, $n = 73$).

The learning task required students to write an academic abstract for a research paper using a custom-built platform integrated with the DeepSeek Large Language Model [5]. To encourage meaningful interaction, the platform imposed a 30-word limit per prompt and disabled copy-pasting, forcing students to synthesize AI outputs rather than passively adopting them.

2.2. Data Collection and Coding Framework

The primary data source was the full interaction log (timestamped prompts and AI responses). We adapted a coding framework from Liu et al. [6] to categorize student prompts into specific communicative acts. The framework distinguishes between Command Behaviors and Questioning Behaviors:

Two researchers independently coded a subset of the data, achieving substantial inter-rater reliability (Cohen's $\kappa = 0.75$).

2.3. Epistemic Network Analysis (ENA)

To capture the structure of interactions, we utilized Epistemic Network Analysis (ENA) [3]. Unlike frequency counts, ENA models the co-occurrence of codes within a moving temporal window, visualizing how different behaviors (e.g., asking for improvement + asking for clarification) are linked in the students' cognitive process. This allows for a structural comparison of interaction strategies between the HG and LG groups.

Table 1
Coding Framework

Type	Description	Example
Abstract Generation Command (AGC)	Direct requests to generate text	Write an abstract for me
AI Improvement Command (AIC)	Requests to refine or polish text	Make this sentence more academic
Meta-Command (MC)	Inquiries about how to prompt or use the tool	How should I summarize this
Clarification Question (CQ)	Seeking explanation of concepts	What does 'epistemic' mean here?
Factual Question (FQ)	Retrieving specific facts from the text	What was the total sample size in the 2019 survey?
Challenging Question (CHQ)	Questioning the AI's logic or accuracy	Are you sure about that percentage?

3. Preliminary Findings

3.1. Interaction Patterns via ENA

The ENA models revealed structurally distinct interaction patterns between the two groups. The Low Literacy (LG) network is dominated by strong connections between Abstract Generation Commands (AGC) and Factual Questions (FQ) or Challenging Questions (CHQ). This pattern indicates a "transactional" approach. LG students tended to issue broad commands for content generation ("Write this") and paired them with basic fact retrieval. The link to CHQ often appeared as reactions to confusion or hallucinations rather than strategic critique. This structure suggests a reliance on GenAI as a search engine or a black-box generator rather than a collaborative agent.

The High Literacy (HG) network is characterized by strong connections between AI Improvement Commands (AIC), Clarification Questions (CQ), and Meta-Commands (MC). This pattern reflects a "collaborative" or "epistemic" approach. HG students frequently combined requests for text refinement with questions to clarify underlying concepts (CQ–AIC link). Furthermore, the integration of Meta-Commands suggests higher metacognitive regulation, where students actively inquire about optimal prompting strategies. This indicates an iterative cycle of Sense-Making $\rightarrow$ Refinement, treating the AI as a co-author and feedback provider.

3.2. Statistical Comparison

To verify these visual differences, we compared the projected ENA scores (positions in the high-dimensional space). A Mann-Whitney U test confirmed significant differences between the two groups on the primary dimension (X-axis) of the ENA space ($U = 683.50$, $p < .001$, $r = .35$). This statistical evidence confirms that GenAI literacy levels are associated with fundamentally different behavioral structures during the learning task.

4. Discussion and Implications

4.1. Behavioral Signatures of GenAI Literacy

Our findings provide empirical validation for theoretical constructs of GenAI literacy. The LG group's focus on Generation (AGC) corresponds to lower-level "Use & Apply" competencies, where the tool is used primarily to bypass cognitive effort. In contrast, the HG group's focus on Improvement (AIC) and Clarification (CQ) reflects higher-order "Evaluate & Create" competencies, involving critical refinement and deep engagement. This suggests that ENA can successfully "fingerprint" GenAI literacy through process data, offering a more granular view than survey-based methods.

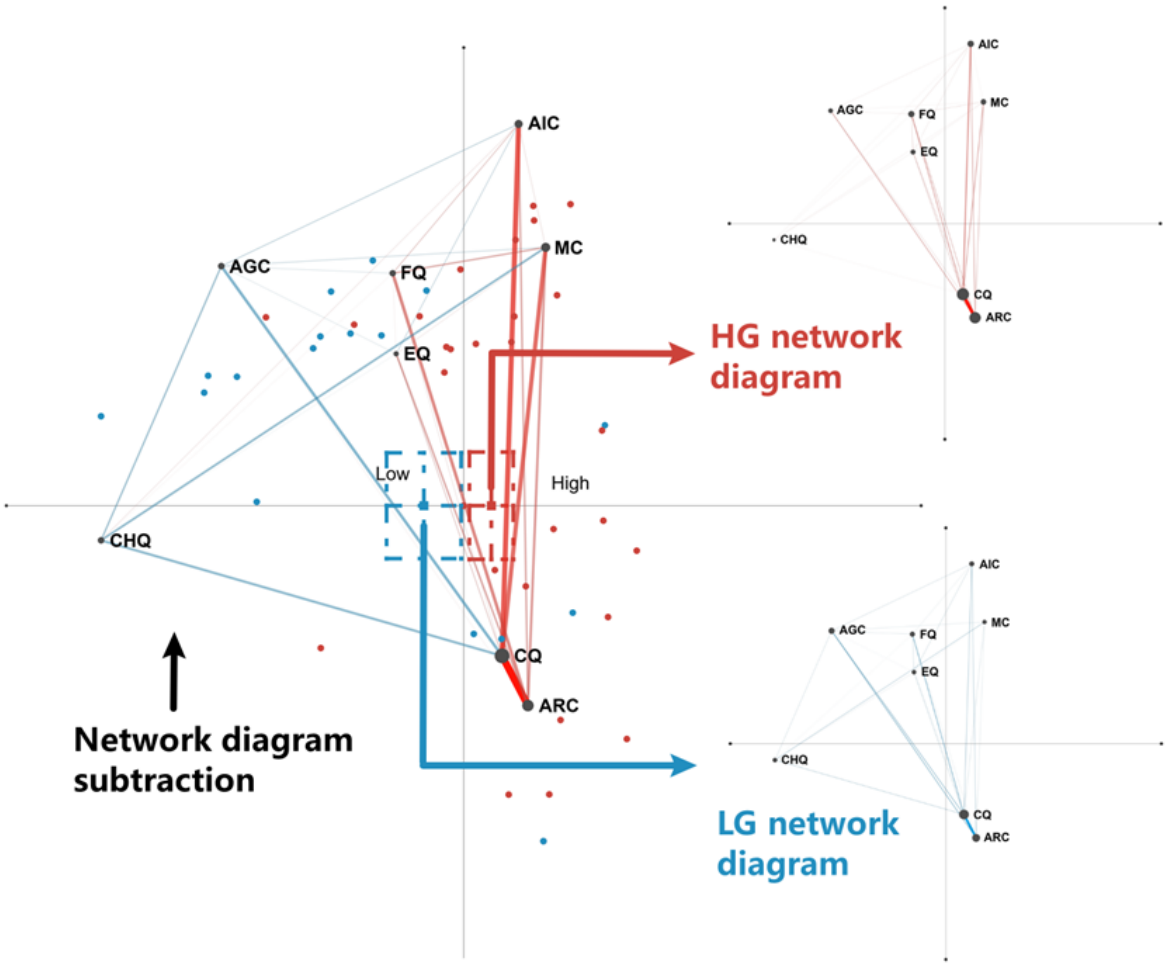

Figure 1: ENA results of LG, HG and network diagram subtraction.

Importantly, the distinction between the two groups lies not only in the frequency of individual prompt types, but also in how those behaviors are connected over time. A clarification question followed by an improvement command suggests understanding followed by revision, whereas a generation command linked to a factual question may reflect information collection without integration. ENA thus reveals strategic organization that isolated prompt counts miss.

These signatures are tendencies rather than fixed categories. Prompting may also be influenced by writing ability, domain familiarity, or platform constraints. Moreover, the meaning of challenging the AI depends on whether it leads to verification and revision. Future studies should combine network patterns with final-text quality and qualitative analyses of interaction sequences to distinguish productive critique from confusion.

4.2. Implications for Design and Assessment

These findings have significant implications for the design of GenAI-integrated learning environments. First, they suggest that literacy can be assessed *in situ*. Rather than relying on pre-tests, educational systems could analyze prompt patterns to estimate a learner's literacy level dynamically.

Second, identifying these patterns enables real-time scaffolding. For example, systems detecting the "Generation-Factual" loop typical of low literacy could intervene with prompts such as: "You are focusing on content generation. Try drafting your own summary first and asking the AI for critique to improve learning depth." Such data-driven interventions could help shift learners from transactional use to epistemic collaboration.

At the interface level, these interventions could be implemented as lightweight, context-sensitive prompts. After repeated generation commands, the system could invite learners to compare formulations, identify unsupported claims, or explain a revision. For students demonstrating clarification–improvement cycles, it could reduce scaffolding and prompt evaluation of argument coherence. This adaptive approach treats GenAI literacy as a capability that can be developed during the task, not merely classified before instruction.

For assessment, network indicators could complement surveys, rubrics, and final-product measures, but should not be used alone or for high-stakes decisions. The same prompt can serve different purposes, while concise expert interactions may resemble direct commands. Designers should calibrate indicators locally, explain how interaction data are interpreted, and allow learners to contest automated inferences. Privacy and data minimization are also essential because prompts may contain sensitive information. With these safeguards, process analytics can support formative feedback while respecting learner agency.

4.3. Conclusion

This study demonstrates that GenAI literacy is not just a static score but a dynamic behavioral capability. By applying Epistemic Network Analysis to interaction logs, we visualized how literacy shapes the human-AI collaboration process. High literacy facilitates strategic, refinement-oriented collaboration, while low literacy is characterized by transactional, generation-oriented dependence. These insights provide a necessary empirical foundation for developing automated literacy assessments and adaptive AI tutors.

Declaration on Generative AI

During the preparation of this work, the author used ChatGPT in order to: Grammar and spelling check. After using this tool/service, the author reviewed and edited the content as needed and takes full responsibility for the publication's content.

References

[1] T. K. Chiu, Y. Chen, K. W. Yau, C. S. Chai, H. Meng, I. King, S. Wong, Y. Yam, Developing and validating measures for AI literacy tests: From self-reported to objective measures, Computers and Education: Artificial Intelligence 7 (2024) 100282.

[2] O. Almatrafi, A. Johri, H. Lee, A systematic review of AI literacy conceptualization, constructs, and implementation and assessment efforts (2019–2023), Computers and Education Open (2024) 100173.

[3] D. W. Shaffer, W. Collier, A. R. Ruis, A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data, Journal of Learning Analytics 3 (2016) 9–45.

[4] Y. Jin, R. Martinez-Maldonado, D. Gašević, L. Yan, GLAT: The generative AI literacy assessment test, Computers and Education: Artificial Intelligence 9 (2025) 100436. doi:10.1016/j.caeai.2025.100436.

[5] DeepSeek-AI, D. Guo, D. Yang, H. Zhang, J. Song, R. Zhang, R. Xu, Q. Zhu, S. Ma, P. Wang, X. Bi, X. Zhang, X. Yu, Y. Wu, Z. F. Wu, Z. Gou, Z. Shao, Z. Li, Z. Gao, Z. Zhang, DeepSeek-R1: Incentivizing reasoning capability in LLMs via reinforcement learning, 2025. URL: https://arxiv.org/abs/2501.12948.

[6] J. Liu, S. Li, Q. Dong, Collaboration with generative artificial intelligence: An exploratory study based on learning analytics, Journal of Educational Computing Research 62 (2024) 1234–1266.