How assessment rubrics reshape engagement networks in AI-assisted academic writing*

Nguyen-Ai-Thu Duong[1,*], Belle Dang[2], Luna Huynh[1] and Andy Nguyen[1]

[1]University of Oulu, Pentti Kaiteran katu 1 90570 Oulu, Finland
[2]Victoria University of Wellington, Kelburn Parade 6012 Wellington, New Zealand

Abstract

The growing use of generative artificial intelligence (AI) in academic writing foregrounds human-AI collaboration for hybrid intelligence, in which learning emerges through the coordinated regulation of human cognition and AI reasoning. This study investigates how assessment design relates to cognitive engagement in human–AI collaboration in writing. Forty undergraduate students completed two consecutive writing tasks with generative AI which were respectively assessed using a conventional rubric and an explicitly AI-aware rubric. Fine-grained writing-process data were coded using an ICAP-based scheme adapted for human–AI interaction and analyzed with Epistemic Network Analysis (ENA). Results indicate that overall distributions of engagement states remained largely stable across tasks; however, differences emerged in the co-occurrence structures among engagement states. These findings suggest that changes in assessment criteria are associated not with how frequently learners engage in particular cognitive states, but with how those states are structurally organized during writing. By adopting a process-oriented analytic lens, this study highlights the value of ENA for examining cognitive engagement and offers practical guidance for redesigning assessment rubrics as regulatory supports in hybrid intelligence writing.

Keywords

Human-AI collaboration, academic writing, ICAP, epistemic network analysis

1. Introduction

The rise of AI has accelerated the relevance of hybrid intelligence in education [10, 22]. Hybrid intelligence emphasizes the complementary integration of human cognitive capacities and AI rather than human replacement [12]. In academic writing, human–AI collaboration (HAIC) transforms how writing processes operate. Students' writing quality and learning are shaped not by whether AI is used, but by how learners interact with AI outputs, deciding what to adopt, revise, justify, or reject [18]. Accordingly, examining students' writing patterns in HAIC contexts enables educators to assess how learners engage with generative AI tools during learning processes, track skill development, and identify areas requiring additional support [11, 29, 31]. Despite these advances, the moment-to-moment unfolding of cognitive engagement in HAIC writing remains underexplored, motivating the systematic coding of collaborative behaviors as observable expressions of cognitive engagement.

Prior research shows that task complexity, alignment with learning objectives, and cognitive demands shape students' capacity for self-regulation [13]. At the group level, collaborative task and assessment design shapes how learners interact, co-regulate, and co-construct knowledge [41]. Specifically, studies indicate that learning-oriented assessment reframes tasks as learning activities, positions students as self or peer evaluators, and treats feedback as feed-forward that reconfigure learners' monitoring and strategy use [4, 30]. However, recent work argues that conventional assessment is increasingly inadequate for accurately gauging students' learning in AI-assisted environments [40], in part because generative AI can help students complete essays, proposals, and take-home exams, heightening concerns about cheating and academic integrity [9, 27]. These developments indicate that assessment methods

Joint Proceedings of LAK 2026 Workshops, co-located with the 16th International Conference on Learning Analytics and Knowledge (LAK 2026), Bergen, Norway, April 27 – May 1, 2026.
*Corresponding author.
Thu.Duong@oulu.fi (N. Duong); belle.dang@vwu.ac.nz (B. Dang); Luna.Huynh@oulu.fi (L. Huynh); Andy.Nguyen@oulu.fi (A. Nguyen)
0009-0008-2070-4001 (N. Duong)

in higher education must evolve to reflect the implications of AI, and that institutions should adapt to changing needs while ensuring equitable access to AI [33, 34]. Against this backdrop, a pressing question is whether AI-aware rubric redesign, explicitly acknowledging and evaluating AI use, can shift student behavior and elevate cognitive engagement toward meaningful, knowledge-building activity within academic tasks.

To fill this gap, we examine the influence of AI-aware assessment task design on students' cognitive engagement in HAIC tasks, employing a mixed-methods design combining ICAP-based qualitative thematic coding with Epistemic Network Analysis (ENA) [35, 38]. The study is guided by the following research questions:

1. RQ1. To what extent do overall engagement distributions and writing-process volumes differ between Task A (conventional rubric) and Task B (AI-aware rubric)?
2. RQ2. How does rubric design reconfigure the co-occurrence structure of engagement states during AI-assisted academic writing?

By pursuing these questions, the study offers insights of the structure of cognitive engagement, the epistemic relations among engagement levels, thereby testing whether moving from a conventional to an AI-aware rubric alters engagement in AI-assisted academic writing. Our findings motivate future studies to refine assessment criteria and develop timely instructional scaffolds that channel AI use toward constructive, knowledge-building participation while mitigating overreliance on AI.

2. Theoretical background

2.1. AI-assisted academic writing and ICAP framework

Academic writing is widely understood as a complex, multi-component, recursive process in which writers coordinate planning, text generation, and revision under task constraints [19]. This process imposes differential cognitive load during composition [5] and taxes multiple interacting systems [24]. Such studies have highlighted cognitive engagement and interaction as essential factors in academic writing process.

The adoption of advanced, AI-enabled writing assistants is substantively reshaping the academic writing process. Beyond posing challenges related to integrity, transparency, and uneven skill development, AI can accelerate multiple facets of the writing workflow (e.g., ideation, feedback, surface-level editing) [17, 23] and facilitate learners' use of planning, monitoring, and revision strategies [20, 26, 28]. Together, these observations suggest that learner-AI interaction during writing affects not only observable behaviors but also learners' internal cognitive processes.

Bringing these strands together, we adopt the ICAP framework (Passive, Active, Constructive, Interactive) to classify observable AI-assisted writing activities by their depth of cognitive engagement. This provides a principled way to operationalize engagement in HAIC writing and to examine whether rubric design functions as an instructional lever that shapes learners' transitions across engagement states over time. In ICAP, engagement is characterized along four modes, Passive, Active, Constructive, and Interactive, linking what learners do to the kind of cognitive processing those actions imply [7, 8]. Passive (P) refers to receiving information without producing any observable output beyond attention (e.g., reading or watching without acting). Active (A) involves manipulating presented information in meaning-preserving ways (e.g., highlighting, copying, repeating, or answering with minimal transformation). Constructive (C) entails generating new inferences or ideas that go beyond the given input (e.g., self-explanations, paraphrases that add relations, drawing novel connections, or posing "why/how" questions). Interactive (I) denotes contingent, back-and-forth exchanges in which contributions are built upon, critiqued, and integrated to co-construct understanding (e.g., proposing, challenging, justifying, and synthesizing) [7, 8].

In this experiment, the writing task was performed by a single author who was permitted to use GenAI as a peer-like collaborator. Accordingly, within this HAIC writing context, Interactive is operationalized as dialogic exchanges between the participant and the generative AI system. Specifically, the learner issues prompt and iteratively refines them so that the GenAI generates content in support of ongoing composition. This usage departs from the conventional ICAP definition of Interactive as peer-to-peer co-construction [6, 8].

Notably, AI-Active is introduced as a subset of Active in which the same form-focused, non-meaning-changing edits are triggered or executed via AI tools (e.g., Google/Word grammar or spelling suggestions, automatic spacing/formatting, autocorrect). Crucially, AI-Active does not change the propositional content or introduce reasoning; it simply polishes the surface with AI assistance. Distinguishing AI-Active from standard Active increases analytic sensitivity to AI mediation. If AI-assisted polishing is collapsed into generic Active, instances in which AI is present yet cognitively shallow are obscured; treating AI-Active as a separate code enables attribution to how AI is used (surface-level formatting and error correction vs. meaning-making) rather than merely whether AI appears. The distinction also preserves a clear boundary from Interactive, which in this study denotes learner-AI dialogue that advances meaning (e.g., prompt for generated content, edit prompt for new generated content). Without the AI-Active category, brief AI-triggered micro-edits could be misclassified as substantive "AI use," thereby inflating Interactive. From an instructional design standpoint, tracking AI-Active separately allows researchers to determine whether rubrics channel students toward Interactive → Constructive pathways (substantive reasoning and integration) rather than toward cosmetic, tool-driven edits.

Taken together, these distinctions motivate treating rubric design as an instructional lever for shaping how learners engage with AI, whether they remain at AI-assisted surface editing (AI-Active) or move toward deeper dialogic and meaning-making engagement (Interactive/Constructive). Furthermore, recent studies suggest that instructional design can improve learning outcomes by deliberately steering learners' cognitive engagement and interaction. In particular, evidence that human–AI teaming tends to yield larger gains on content-creation tasks than on decision tasks underscores the need to design writing workflows, and accompanying rubrics, that encourage learners to use AI for idea development and reasoning rather than for cosmetic edits alone [37]. This aligns with prior work showing that learning improves when instruction positions students as active contributors rather than passive recipients [15].

2.2. Assessment rubrics in AI-assisted academic writing

Rubrics are commonly defined as explicit scoring guides that articulate performance criteria and quality levels for a given task [3]. Rather than functioning merely as grading tools, analytic rubrics make assessment expectations transparent and operationalize learning outcomes into observable indicators, thereby serving both evaluative and instructional purposes. In the context of academic writing, rubrics act as mediational artifacts that connect task demands with students' cognitive and metacognitive regulation during planning, drafting, and revision [32].

Conventional rubrics for academic essays typically include criteria such as task fulfillment, organization and coherence, argumentation and use of evidence, lexical range and appropriateness, grammatical accuracy, and stylistic conventions, reflecting dimensions shown to structure both student performance and self-assessment in recent writing assessment research [1, 36]. However, the emergence of AI-assisted writing tools challenges the sufficiency of traditional criteria, as they do not explicitly address how text is produced, revised, or justified in human–AI collaborative contexts. Recent scholarship therefore calls for rethinking assessment that extend conventional dimensions with criteria capturing responsible, transparent, and cognitively meaningful engagement with AI [2, 33].

From a theoretical perspective, embedding these AI-aware criteria into assessment design is expected to reorient students' cognitive engagement: rather than optimizing for surface-level text quality alone, students are prompted to engage in higher-order activities such as evaluating AI output, justifying revisions, and reflecting on the division of labor between human and machine. In this sense, AI-aware rubrics function not only as assessment instruments but also as regulatory scaffolds that shape learners' strategic choices and depth of processing throughout the writing process, extending the principle of constructive alignment to human–AI collaborative learning environments.

3. Research method

3.1. Participants and procedure

The study involved 40 undergraduate English majors enrolled at a university in Vietnam. Participants were recruited voluntarily through an announcement circulated within the student community. Before registration, students were informed about the study, the procedures for data collection and use, and their right to withdraw at any time. All participants provided consent to be recorded. The experiment was conducted online via Zoom, a modality selected to accommodate the geographical dispersion of participants.

Each participant completed two consecutive 30-minute essay-writing tasks. In both tasks, students wrote a concise essay of 250–300 words of two different topics related to AI and were permitted to consult any external resources, including generative AI tools and web search. Task A was evaluated using a conventional rubric, whereas Task B employed an AI-aware rubric that explicitly accounted for AI support. Rubrics were embedded in the task prompts, and students were informed of the rubric changes prior to commencing Task B. In terms of writing topic, Task A examined AI's impact on employment and workforce policy, whereas Task B focused on AI-enabled digital transformation and its broader societal implications.

Throughout the session, screen recording captured participants' interactions with the writing tasks and their use of digital tools, providing a fine-grained record of the writing process. Forty students completed two consecutive writing tasks, yielding 80 recordings (one per student per task). Each file was assigned an anonymized participant ID and labeled by condition, Task A or Task B, to prepare for subsequent event coding of writing actions under the ICAP-based scheme.

To address the research questions, we coded 10,038 fine-grained writing actions produced by 40 students across two writing tasks guided by different rubrics: one conventional and one explicitly AI-aware. Using an ICAP-based scheme (Interactive, Constructive, Active, AI-Active, Passive), we then applied ENA to characterize epistemic relationships among engagement states within each task. This design yields evidence about the co-occurrence versus independence of engagement states, clarifying how rubrics design shapes the co-occurrence structure of cognitive engagement during HAIC writing.

3.2. Conventional assessment and AI-aware assessment design

The experiment employed two assessment rubrics. The conventional rubric for Task A evaluates five domains: content (completeness and relevance of the discussion of AI's labor-market impacts, 30%), analysis (quality of reasoning and evidential support, 30%), organization and structure (logical progression and effective transitions, 15%), quality of writing (clarity, formal tone, and control of grammar and mechanics, 15%), and adherence to word limit and source quality (250 − 300 words and judicious use of credible sources, 10%). The AI-aware rubric for Task B retains expectations for scholarly argumentation but explicitly assesses writing in AI-mediated contexts. It emphasizes originality and critical insight (independent thinking about AI's role and implications, 30%), argument structure and coherence (30%), use and integration of evidence (synthesis of traditional sources, digital data, and AI outputs, 10%), depth of analysis (balanced treatment of benefits, risks, and longer-term implications with reasoned recommendations, 20%), and academic writing quality and conventions (10%).

3.3. ICAP-based scheme coding

This study examines the epistemic relationships among cognitive engagement states during GenAI-supported collaborative writing. Accordingly, we analyzed the screen recordings to document and qualitatively code students' writing events. The qualitative analysis followed qualitative content analysis procedures [25] and an ICAP-based scheme adapted from the ICAP framework [6, 8] for the HAIC context.

In the open-coding phase, every observable action involved in producing the written product was assigned a brief descriptive code. Across the two tasks, the 40 students generated 10,057 coded actions.

Table 1
ICAP-based states used for labeling HAIC writing actions

ICAP-based state	Code	Definition	Action description
Passive	P	Learner attends to AI- or externally provided information without producing any observable contribution	Viewing instruction (or outline, rubric, prompt, generated content, essay content, search results, article). React to host notification of time.
AI-Active	A.AI	Learner performs form-focused, non-meaning-changing edits are triggered or executed via AI tools	Correct grammar and spelling using Google/Word suggestions. Edit essay forms use Word suggestions.
Active	A	Learner selects, rearranges, or reuses presented content without generating new ideas	Technical operation (e.g., open, close, search, go back, highlight, screenshot, submit). Copy or paste information (e.g., entire/portion of generated content, search outputs). Edit essay forms (adjust fonts, size, spacing). Check word count. Correct grammar and spelling. Search for information (word, concepts, articles) using non-AI tools.
Constructive	C	Learner generates new ideas or inferences that extend beyond the information provided by AI or external materials	Write content, edit content, outline content.
Interactive	I	Learner engages in contingent exchanges with AI, building on or critiquing AI outputs to co-construct understanding	Prompt, edit prompt. Paste content to formulate prompt.

These open codes were then consolidated and mapped to the corresponding cognitive-engagement levels in the ICAP-based taxonomy to form a layered code. Nineteen actions were excluded as technical artifacts (e.g., recording disconnects) that did not contribute to the writing process. The remaining 10,038 open codes were subsequently categorized under the ICAP-based scheme (see Table 1). Cross-validation between researchers confirmed the mutual exclusivity of the coding scheme. This procedure allows us to systematically characterize the composition and sequencing of cognitive engagement states across both tasks in a HAIC environment.

In our adaptation, Interactive denotes dialogic exchanges between learners and GenAI. This departs from the conventional ICAP usage, where Interactive refers to peer-to-peer co-construction [6, 8], and is adopted here to capture the bidirectional nature of learner-GenAI engagement in single-author HAIC tasks. Alongside Active, which in the original ICAP framework refers to learner-initiated, surface-level actions (e.g., formatting or self-identified grammar and spelling corrections), we introduce an additional category: AI-Active. AI-Active designates surface-level, form-focused edits facilitated by AI tools (e.g., grammar, spelling, or spacing suggestions). This distinction enables us to differentiate between formal adjustments made independently by the learner and those supported by AI assistance.

3.4. Data analysis

Given the substantial within-task variability in sequence length and the conceptual focus on learner–AI interaction, a network approach is appropriate to examine structure beyond counts. We therefore

Table 2
Descriptive statistics for two specific-task groups

	Task A		Task B	
	f	%	f	%
Total sequences (*)	40	100	40	100
Average sequence length (**)	122.67	(SD 48.69)	128.28	(SD 50.33)
Passive	164	33.46	1767	34.44
Active	1522	31.02	1502	29.27
Active with AI-assisted	78	1.59	79	1.54
Interactive	596	12.15	653	12.73
Constructive	1069	21.79	1130	22.02

Note. (*) Number of writing sequences, each representing one student's complete process. (**) Average number of actions per sequence.

used ENA [36, 43] to model the joint activation of ICAP-based states across 10,057 process-level coded action units, using a moving conversational window of four coded actions to capture local temporal coupling between adjacent cognitive states, consistent with prior ENA research on learning processes. ENA projections were derived using singular value decomposition (SVD), retaining two dimensions optimized to distinguish between Task A and Task B networks. Statistical significance of between-task differences was assessed using permutation-based tests on ENA centroid scores. To avoid inflated Type I error, inferential claims were based on centroid-level contrasts rather than individual edges.

4. Results and finding

4.1. To what extent do overall engagement distributions and writing-process volumes differ between Task A (conventional rubric) and Task B (AI-aware rubric)?

Table 2 presents descriptive statistics for the writing-process traces, including the total number of sequences, the average sequence length, and the contribution of each cognitive state. Overall activity volume was slightly higher in the second task: students produced 5,131 coded actions in Task B versus 4,907 in Task A, consistent with longer sequences per student (Task B: M = 128.28, SD = 50.33; Task A: M = 122.67, SD = 48.69). Dispersion was comparable (CV ≈ .39 in both tasks), indicating that while students tended to engage modestly more in Task B, the variability of their writing processes remained similar across tasks.

Across ICAP-based states, Passive and Active together dominated in both tasks (>60%). From Task A to Task B, Passive rose only slightly from 33.46% (n = 1,642) to 34.44% (n = 1,767; +0.98 percentage point (pp)), while Active declined from 31.02% (n = 1,522) to 29.27% (n = 1,502; −1.75 pp). Constructive was stable, nudging from 21.79% (n = 1,069) to 22.02% (n = 1,130; +0.23 pp). Human-AI Interactive also ticked up marginally from 12.15% (n = 596) to 12.73% (n = 653; +0.58 pp). By contrast, Active with AI-assisted remained rare and essentially unchanged (Task A: 1.59%, n = 78; Task B: 1.54%, n = 79; −0.05 pp), suggesting no shift toward superficial, tool-driven text polishing.

At the per-student level, ICAP compositions were highly stable across tasks. Paired comparisons showed only small, non-significant shifts in mean proportions: Passive increased by +1.05 percentage points (pp) on average (95% CI [−1.44, 3.55], p = .399, d_n = 0.14), Active decreased by −1.52 pp (95% CI [−4.17, 1.14], p = .255, d_n = −0.18), Interactive increased by +0.35 pp (95% CI [−1.99, 2.69], p = .764, d_n = 0.05), and Constructive changed by +0.07 pp (95% CI [−1.97, 2.12], p = .942, d_n = 0.01). AI-Active remained negligible, with no detectable paired change (Wilcoxon signed-rank: p = .899; rank-biserial r = −0.04). Collectively, these paired results indicate that, for most students, the relative time spent in Passive, Active, Interactive, and Constructive states in Task B closely mirrored Task A, with any individual shifts being small in magnitude and evenly distributed around zero.

Table 3
Edge weight of ICAP-based states in the Task A and Task B

Task A					
	I	C	A	A.AI	P
I		0.07	0.22	0.00	0.26
C	0.07		0.39	0.05	0.47
A	0.22	0.39		0.03	0.54
A.AI	0.00	0.05	0.03		0.03
P	0.26	0.47	0.54	0.03	
Task B					
	I	C	A	A.AI	P
I		0.04	0.17	0.00	0.30
C	0.04		0.43	0.05	0.46
A	0.17	0.43		0.04	0.50
A.AI	0.00	0.05	0.04		0.03
P	0.03	0.46	0.50	0.03	

Taken together, the engagement profile appears highly similar across tasks. The AI-aware rubric in Task B coincided with slightly longer sequences and a minor reweighting among categories (a small dip in Active alongside small gains in Passive, Constructive, and Interactive), but no notable movement toward AI-assisted form editing. Per-student comparisons reinforced this picture of stability: most learners allocated their time across Passive, Active, Interactive, and Constructive in much the same way across tasks, with any individual shifts small and evenly balanced in both directions. Nonetheless, comparable margins can still mask reorganizations in how states cluster and unfold over time. Accordingly, ENA is warranted to test whether the rubric change altered co-occurrence structure among ICAP states across Task A and Task B.

4.2. How does rubric design reconfigure the co-occurrence structure of engagement states during AI-assisted academic writing?

Table 3 and Figure 1 below present the ENA results for the two tasks. Table 3 reports the connection strengths among cognitive states, whereas Figure 1 visualizes the co-occurrence network of these states.

In Task A, the network is anchored by strong links from Passive to both Active and Constructive (P–A = 0.54; P–C = 0.47), with a moderate P–I tie (0.26). Connections among the higher-level states are weaker: A–C = 0.39 and A–I = 0.22, while I–C is minimal (0.07). Edges involving AI-Active are near-floor (P–A.AI = 0.03; A–A.AI = 0.03; A.AI–C = 0.05), indicating that AI-driven surface editing rarely co-occurred with other actions. In Task B, the overall shape remains similar. P–A (0.50) and P–C (0.46) continue to dominate; P–I grows modestly (0.30), A–C becomes slightly stronger (0.43), and A–I weakens (0.17). I–C is again very small (0.04). Edges tied to AI-Active stay negligible (P–A.AI = 0.03; A–A.AI = 0.04; A.AI–C = 0.05). Taken together, the Task B network shows a mild redistribution toward co-occurrences that connect Passive with Interactive and Active with Constructive, while ties among Interactive and Constructive remain weak.

Comparing tasks (Task B − Task A), the largest absolute changes are modest: P–I increases by +0.04 (0.30 vs. 0.26), A–C by +0.04 (0.43 vs. 0.39), P–A decreases −0.04 (0.50 vs. 0.54), and A–I decreases −0.05 (0.17 vs. 0.22), other differences are $\leq |0.03|$. Thus, the conventional "hub-and-spoke" around Passive (P–A, P–C) persists, with small shifts that favor P–I and A–C in Task B and slightly weaken A–I. Consistent with the descriptive counts, AI-Active remains peripheral in both tasks.

Turning to the ENA axes, the two group means separate significantly along MR1 (12.2% of variance explained): Task A (Mdn = 0.07, N = 40) > Task B (Mdn = 0.00, N = 40), U = 587.00, $p = .04$, $r = .27$, indicating a small-to-moderate shift in the primary contrast captured by MR1. By contrast, there is no difference along SVD2 (39.5% variance): Task A (Mdn = 0.08) vs. Task B (Mdn = 0.06), U = 804.00,

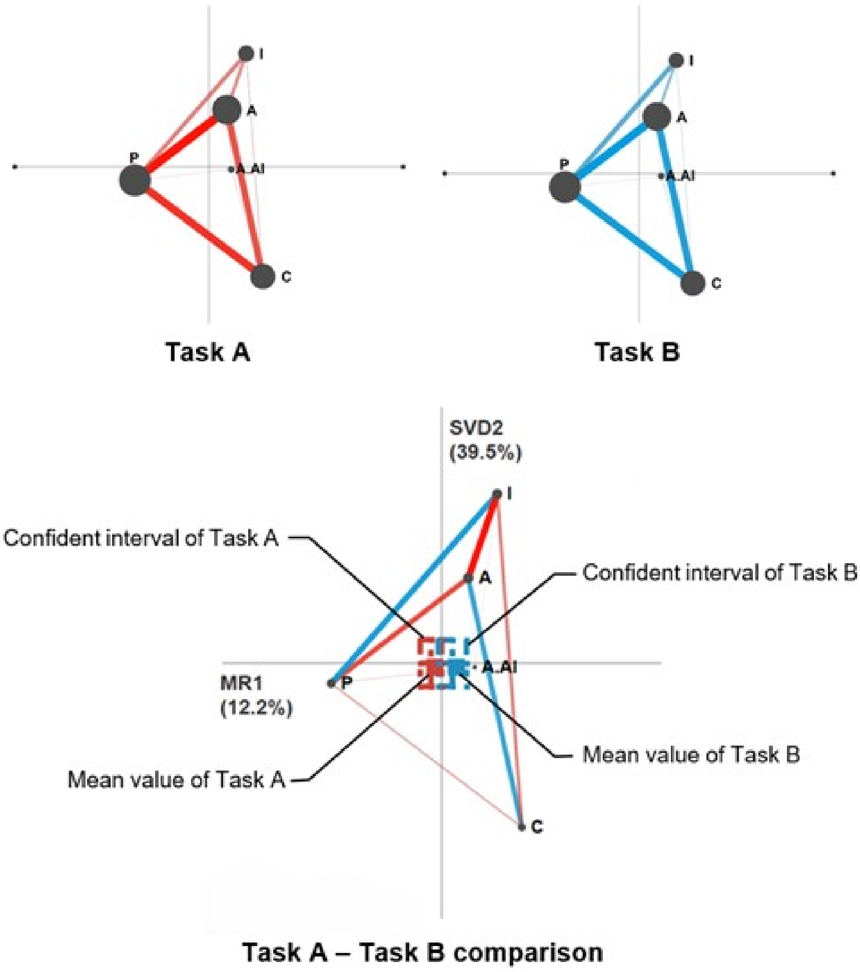

Figure 1: Epistemic networks of engagement states in the two tasks

p = .97, r = .00. In practical terms, the rubric change is associated with a subtle reorganization of co-occurrences summarized by MR1 (e.g., slightly more P–I and A–C and slightly less P–A and A–I), while the secondary contrast represented by SVD2 remains stable.

Across both tasks, engagement networks were organized around a persistent hub-and-spoke structure in which Passive engagement served as the primary entry point into Active and Constructive states. The introduction of an AI-aware rubric did not alter this overall topology but was associated with modest redistributions in co-occurrence patterns. Specifically, Task B showed slightly stronger links between Passive and Interactive engagement and between Active and Constructive engagement, alongside weaker ties between Passive and Active and between Active and Interactive states. These changes were captured by a small-to-moderate separation along the primary ENA dimension (MR1), while the secondary dimension remained stable. Together, these findings suggest that the rubric change influenced how engagement states were organized and combined during writing, without fundamentally reshaping the engagement ecology or elevating AI-mediated activity to a central role.

5. Discussion

This study contributes to emerging research on HAIC writing by examining how assessment design relates to the organization of learners' cognitive engagement, rather than to changes in engagement frequency or AI usage per se. Framed within hybrid intelligence, the findings align with recent scholarship that conceptualizes learning with AI as a coordinated human–AI system shaped by task design, assessment signals, and regulatory structures, rather than as a substitution or augmentation of human cognition [12, 21].

Consistent with contemporary learning analytics research, the results demonstrate that similar distributions of engagement states can conceal meaningful differences in process organization. This reinforces arguments that frequency-based indicators alone are insufficient for understanding complex learning processes, particularly in AI-rich environments [39]. Network- and sequence-oriented approaches, including ENA, have been increasingly advocated as means of capturing how learning activities are structured over time and how cognitive states are coordinated in situ [35, 38].

Importantly, the findings suggest that AI-aware assessment does not produce immediate or large-scale shifts in learners' engagement profiles. Rather, its influence appears modest, operating through subtle reconfigurations in how engagement states are combined during writing. This interpretation aligns with recent work in assessment research emphasizing that rubrics and criteria function as regulatory artifacts, signaling valued forms of engagement and shaping learners' strategic choices, rather than as direct levers for behavioral change [2]. In HAIC contexts, such regulatory effects may be especially incremental, as learners adapt established writing practices to accommodate AI support while retaining responsibility for meaning-making and judgment [14].

By foregrounding engagement organization rather than writing quality or AI adoption rates, this study responds to growing critiques of product-focused and tool-centric evaluations of generative AI in education [16]. The findings underscore the value of process-oriented analytics for examining how hybrid intelligence is enacted during learning activities, while avoiding assumptions that AI use necessarily entails deeper engagement or improved learning. At the same time, they point to the need for future research to triangulate process-level evidence with outcomes, learner perceptions, and longitudinal designs to better understand how assessment and AI jointly shape learning trajectories.

Taken together, this study positions AI-aware rubrics not as mechanisms that transform cognitive engagement outright, but as modest regulatory elements operating within a largely stable engagement ecology. For researchers and educators, this suggests that the educational impact of AI in writing may be better understood through gradual, process-level reorganization of engagement, rather than through immediate changes in surface behavior or performance, an orientation consistent with current directions in learning analytics and assessment research.

6. Conclusion and future research directions

This study examined how assessment design relates to the process of AI-assisted academic writing by analyzing learners' writing activity through descriptive measures and ICAP-based co-occurrence networks. Across tasks, the overall distribution of cognitive engagement states remained largely stable; however, Epistemic Network Analysis revealed differences in the co-occurrence structure among engagement states under conventional and AI-aware rubric conditions. These findings indicate that moving from a conventional to an AI-aware rubric does not primarily influence how frequently learners engage in particular cognitive states, but is associated with how those states are structurally related during writing. This interpretation underscores the value of making AI expectations explicit in instructional design, not to drive immediate behavioral change, but to shape the organization of cognitive engagement in AI-supported writing contexts.

More broadly, this study advances theory in human–AI collaborative writing by conceptualizing assessment criteria as regulatory contexts that shape learners' cognitive activity without presuming direct or causal effects on learning depth. By extending the ICAP framework to distinguish AI-related surface activity from other forms of engagement, the study provides a principled way to examine how AI use is situated within writing processes, rather than treating AI interaction as inherently productive or detrimental. This distinction enables researchers to more precisely investigate the role of AI in learners' cognitive engagement in increasingly AI-mediated educational environments.

The findings also demonstrate the value of process-oriented analytics, such as Epistemic Network Analysis, for examining cognitive engagement as a relational construct in AI-assisted writing. By focusing on co-occurrence structure rather than frequency alone, this approach reveals aspects of engagement organization that would remain invisible in aggregate summaries. These insights lay groundwork for future research that integrates co-occurrence modeling with temporal analyses, learning outcomes, and learner perspectives to better understand how assessment design and AI jointly shape writing development in hybrid intelligence contexts.

Declaration on Generative AI

During the preparation of this work, the authors used ChatGPT 5.2 in order to: Grammar and spelling check.

References

[1] Al-Nafjan, E. et al. 2025. Exploring self-assessment in EFL academic writing: a study of undergraduate Saudi students utilizing a rubric. *Language Testing in Asia.* 15, 1 (2025), 62. https://doi.org/10.1186/s40468-025-00397-2.

[2] Bearman, M. et al. 2023. Designing assessment in a digital world: an organising framework. *Assessment & Evaluation in Higher Education.* 48, 3 (Apr. 2023), 291–304. https://-doi.org/10.1080/02602938.2022.2069674.

[3] Brookhart, S.M. 2018. Appropriate Criteria: Key to Effective Rubrics. *Frontiers in Education.* Volume 3-2018, (2018). https://doi.org/10.3389/feduc.2018.00022.

[4] Carless, D. 2015. Exploring learning-oriented assessment processes. *Higher Education.* 69, 6 (2015), 963–976. https://doi.org/10.1007/s10734-014-9816-z.

[5] Carnota, A. et al. 2024. The impact of handwriting fluency on bursts of written language and text quality in undergraduates. *Reading and Writing.* (2024). https://doi.org/10.1007/s11145-024-10621-9.

[6] Chi, M.T.H. 2009. Active-Constructive-Interactive: A Conceptual Framework for Differentiating Learning Activities. *Topics in Cognitive Science.* 1, 1 (Jan. 2009), 73–105. https://doi.org/https://-doi.org/10.1111/j.1756-8765.2008.01005.x.

[7] Chi, M.T.H. 2021. Translating a Theory of Active Learning: An Attempt to Close the Research-Practice Gap in Education. *Topics in Cognitive Science.* 13, 3 (Jul. 2021), 441–463. https://doi.org/https://-doi.org/10.1111/tops.12539.

[8] Chi, M.T.H. and Wylie, R. 2014. The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes. *Educational Psychologist.* 49, 4 (Oct. 2014), 219–243. https://-doi.org/10.1080/00461520.2014.965823.

[9] Chiu, T.K.F. et al. 2023. Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence.* 4, (2023), 100118. https://doi.org/https://doi.org/10.1016/j.caeai.2022.100118.

[10] Cukurova, M. 2025. The interplay of learning, analytics and artificial intelligence in education: A vision for hybrid intelligence. *British Journal of Educational Technology.* 56, 2 (Mar. 2025), 469–488. https://doi.org/https://doi.org/10.1111/bjet.13514.

[11] Dang, B. et al. 2024. Decoding AI-Assisted Academic Writing Process: Insights from Lag Sequential and Epistemic Network Analyses. *Conference: Pacific Asia Conference on Information Systems (PACIS)At: Ho Chi Minh, Vietnam.* July (2024), 1–16.

[12] Dellermann, D. et al. 2019. Hybrid Intelligence. *Business & Information Systems Engineering.* 61, 5 (2019), 637–643. https://doi.org/10.1007/s12599-019-00595-2.

[13] Dignath, C. and Büttner, G. 2008. Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning.* 3, 3 (2008), 231–264. https://doi.org/10.1007/s11409-008-9029-x.

[14] Eaton, S.E. 2023. Postplagiarism: transdisciplinary ethics and integrity in the age of artificial intelligence and neurotechnology. *International Journal for Educational Integrity.* 19, 1 (2023), 23. https://doi.org/10.1007/s40979-023-00144-1.

[15] Freeman, S. et al. 2014. Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences.* 111, 23 (Jun. 2014), 8410–8415. https://doi.org/10.1073/pnas.1319030111.

[16] Gašević, D. et al. 2023. Empowering learners for the age of artificial intelligence. *Computers and Education: Artificial Intelligence.* 4, (2023), 100130. https://doi.org/https://-doi.org/10.1016/j.caeai.2023.100130.

[17] Giannakos, M. et al. The promise and challenges of generative AI in education. *Behaviour & Information Technology.* 1–27. https://doi.org/10.1080/0144929X.2024.2394886.

[18] Goldberg, P. et al. 2021. Attentive or Not? Toward a Machine Learning Approach to Assessing Students' Visible Engagement in Classroom Instruction. *Educational Psychology Review.* 33, 1 (2021), 27–49. https://doi.org/10.1007/s10648-019-09514-z.

[19] Hayes, John R 2012. Modeling and Remodeling Writing. *Written Communication.* 29, 3 (Jul. 2012), 369–388. https://doi.org/10.1177/0741088312451260.

[20] Järvelä, S. et al. 2013. Exploring Socially Shared Regulation in the Context of Collaboration. *J Cogn Educ Psych.* 12, 3 (2013), 267–2013. https://doi.org/10.1891/1945-8959.12.3.267.

[21] Kohnke, Lucas et al. 2023. ChatGPT for Language Teaching and Learning. *RELC Journal.* 54, 2 (Apr. 2023), 537–550. https://doi.org/10.1177/00336882231162868.

[22] Kong, X. et al. 2025. Examining human–AI collaboration in hybrid intelligence learning environments: insight from the Synergy Degree Model. *Humanities and Social Sciences Communications.* 12, 1 (2025), 821. https://doi.org/10.1057/s41599-025-05097-z.

[23] Lee, Y.-J. 2024. Can my writing be polished further? When ChatGPT meets human touch. *ELT Journal.* 78, 4 (Oct. 2024), 401–413. https://doi.org/10.1093/elt/ccae039.

[24] Li, J. et al. 2023. Investigating Latent Interactions between Students' Affective Cognition and Learning Performance: Meta-Analysis of Affective and Cognitive Factors. *Behavioral Sciences.*

[25] Mayring, P. 2020. Study on position control of piezoelectric actuators. *Forum: Qualitative Social Research / Forum Qualitative Sozialforschung.* 1, 2 (2020), Art. 20. https://doi.org/https://-doi.org/10.17169/fqs-1.2.1089.

[26] Molenaar, I. 2022. The concept of hybrid human-AI regulation: Exemplifying how to support young learners' self-regulated learning. *Computers and Education: Artificial Intelligence.* 3, (2022), 100070. https://doi.org/https://doi.org/10.1016/j.caeai.2022.100070.

[27] Moorhouse, B.L. et al. 2023. Generative AI tools and assessment: Guidelines of the world's top-ranking universities. *Computers and Education Open.* 5, (2023), 100151. https://doi.org/https://-doi.org/10.1016/j.caeo.2023.100151.

[28] Nguyen, A. et al. 2022. Exploring Socially Shared Regulation with an AI Deep Learning Approach Using Multimodal Data. *Proceedings of International Conference of the Learning Sciences, ICLS* .

[29] Nguyen, A. et al. 2024. Human-AI collaboration patterns in AI-assisted academic writing. *Studies in Higher Education.* 49, 5 (2024), 847–864. https://doi.org/10.1080/03075079.2024.2323593.

[30] Nicol, D.J. and Macfarlane-Dick, D. 2006. Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education.* 31, 2 (Apr. 2006), 199–218. https://doi.org/10.1080/03075070600572090.

[31] Oliveira, E.A. et al. 2020. Writing analytics across essay tasks with different cognitive load demands. 1 (2020), 60–70.

[32] Panadero, E. et al. 2017. Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review.* 22, (2017), 74–98. https://doi.org/https://-doi.org/10.1016/j.edurev.2017.08.004.

[33] Perkins, M. 2023. Academic Integrity considerations of AI Large Language Models in the post-pandemic era: ChatGPT and beyond. *Journal of University Teaching and Learning Practice.* 20, 2 (2023). https://doi.org/10.53761/1.20.02.07.

[34] Rajabi, P. et al. 2023. Exploring ChatGPT's impact on post-secondary education: A qualitative study. *Proceedings of the 25th Western Canadian Conference on Computing Education* (New York, NY, USA, 2023).

[35] Shaffer, D.W. et al. 2016. A Tutorial on Epistemic Network Analysis: Analyzing the Structure of Connections in Cognitive, Social, and Interaction Data. *Journal of Learning Analytics*. 3, 3 SE- (Dec. 2016), 9–45. https://doi.org/10.18608/jla.2016.33.3.

[36] Taylor, B. et al. 2024. Rubrics in higher education: an exploration of undergraduate students' understanding and perspectives. *Assessment & Evaluation in Higher Education*. 49, 6 (Aug. 2024), 799–809. https://doi.org/10.1080/02602938.2023.2299330.

[37] Vaccaro, M. et al. 2024. When combinations of humans and AI are useful: A systematic review and meta-analysis. *Nature Human Behaviour*. 8, 12 (2024), 2293–2303. https://doi.org/10.1038/s41562-024-02024-1.

[38] Williamson Shaffer, D. and Ruis, A. 2017. Epistemic Network Analysis: A Worked Example of Theory-Based Learning Analytics. *The Handbook of Learning Analytics*. C. Lang et al., eds. Society for Learning Analytics Research (SoLAR). 175–187.

[39] Winne, P.H. 2017. Cognition and metacognition within self-regulated learning. *Handbook of self-regulation of learning and performance*. Routledge. 36–48.

[40] Xia, Q. et al. 2024. A scoping review on how generative artificial intelligence transforms assessment in higher education. *International Journal of Educational Technology in Higher Education*. 21, 1 (2024), 40. https://doi.org/10.1186/s41239-024-00468-z.

[41] Zabolotna, K. et al. 2025. How does collaborative task design shape collaborative knowledge construction and group-level regulation of learning? A study of secondary school students' interactions in two varied tasks. *International Journal of Computer-Supported Collaborative Learning*. (2025). https://doi.org/10.1007/s11412-024-09442-3.