AI-Regulated Writing Framework: Integrating Self-Regulated Learning, AI Literacy, and Complex Adaptive System Dynamics

Tihana Banko*
tihana.banko@algebra.hr
https://orcid.org/0000-0001-5474-840X

https://doi.org/10.31192/np.24.3.10
UDK / UDC: 004.8:37.091.322=811.111
37.015.3:159.953
Prethodno priopćenje /
Preliminary communication
Primljeno / Received:
23. veljače 2026. / Feb 23, 2026
Prihvaćeno / Accepted:
26. lipnja 2026. / Jun 26, 2026



This article proposes the AI-Regulated Writing Framework, a micro-level conceptualization that integrates phases of self-regulated learning – forethought, performance, and self-reflection – along with competences of AI literacy and the dynamics of complex adaptive systems, in order to understand individual processes of AI-supported academic writing in English for Specific Purposes contexts. The model is proposed in response to limitations of previous research, which has predominantly focused on AI tools and ethical issues while overlooking the dynamic interactions between learners and AI. Accordingly, the pilot study conducted an initial examination of the effects of a brief, 15-minute scaffold that connects phases of self-regulated learning and AI literacy competences. The aim was to determine whether such support increases self-regulation, critical evaluation of AI tools, and adaptation driven by complex adaptive systems in AI-assisted writing. The results indicate substantial improvements in summary-writing performance (d=1.82), markers of self-regulated learning (d=1.92; e.g., an increase in monitoring of one's own writing process and use of AI from 25% to 83%), AI literacy (d=1.20; e.g., a 41% increase in awareness of AI limitations), and a 58% shift towards adaptive complex adaptive system patterns. The editing of AI-generated text shifted from predominantly surface-level changes

* Tihana Banko, PhD, Lecturer, Department of Interdisciplinary Sciences, Algebra Bernays University, Gradišćanska 24, HR-10000 Zagreb, Croatia.

(65% in Task 1) toward greater content-related (30%) and structural revisions (25%), which together accounted for 55% of Task 2 edits. These findings provide preliminary indicators consistent with the proposed framework and point to the need for tasks that foster self-regulated learning with AI support and systematic training in AI literacy in higher education contexts.

Keywords: *AI-assisted ESP writing; AI literacy; AI-Regulated Writing Framework; complex adaptive systems; self-regulated learning.*

Introduction

The rapid diffusion of generative artificial intelligence (AI) tools has begun to transform academic writing practices in higher education, particularly in English as a Foreign Language (EFL) and English for Specific Purposes (ESP) contexts. Students routinely use large language models to summarize sources, generate drafts, reformulate technical terminology and correct language. However, while most students use AI-tools for writing because of time efficiency, they are not sure whether the text is accurate or of high quality[1]. At the same time, teachers and institutions face considerable uncertainty about how AI should be positioned in higher education. The findings of a recent 2025 Chinese study reveal general concerns about academic integrity, fairness, equity, privacy, and data security, including specifically the invisible labor in dealing with ambiguous policies, uneven enforcement strategies, loopholes to avoid detection, disparities in access to state-of-the-art tools, and the cognitive impacts due to overreliance on GenAI tools[2]. This situation produces confusion for students, who routinely report mixed messages and inconsistent expectations about when and how GenAI use is allowed in their courses[3].

Existing research and debate on AI in academic writing have largely focused on tools and ethics: which AI systems students use, how these tools affect plagiarism detection, what counts as acceptable assistance, and how to update academic honesty policies[4]. Models have been suggested for embedding AI

[1] Cf. Martina STADNIK et al., AI and writing skills: Students' attitudes towards using AI to enhance their writing based on the example of Algebra University students, *Interdisciplinary Description of Complex Systems*, 23 (2025) 1, 46-71, 67.

[2] Cf. Jack TSAO, Trajectories of AI policy in higher education: Interpretations, discourses, and enactments of students and teachers, *Computers and Education: Artificial Intelligence*, 9 (2025) 100496, 1-15, 1.

[3] Cf. Brian W. STONE, Generative AI in Higher Education: Uncertain Students, Ambiguous Use Cases, and Mercenary Perspectives, *Society for the Teaching of Psychology*, 52 (2024) 1, 1-10, 6.

[4] Cf. Debby R. E. COTTON, Peter A. COTTON, J. Reuben SHIPWAY, Chatting and Cheating: Ensuring Academic Integrity in the Era of ChatGPT, *Innovations in Education and Teaching International* (2023), 1-12, 3; Mike PERKINS, Academic Integrity considerations of AI Large Language Models in the post-pandemic era: ChatGPT and beyond, *Journal of University Teaching & Learning Practice*, 20 (2023) 2, 1-24, 2-3, 14.

literacy into the higher education curriculum, such as the University of Florida five-category AI Literacy Model (Enabling AI, Know and Understand AI, Use and Apply AI, Evaluate and Create AI, AI Ethics)[5]. Such work is essential, but it tends to conceptualize AI as an external instrument or a risk factor, rather than as something that interacts dynamically with learners' self-regulation, competences and contexts. Much of the literature isolates either the technological layer (e.g. capabilities and limitations of AI tools) or the normative layer (e.g. rules, bans and ethical guidelines), with relatively little attention to the ongoing, moment-to-moment processes through which students plan, monitor and adapt their writing when AI is available. In other words, we know a great deal about what AI can do and what is permitted, but far less about how AI becomes integrated into individual learners' complex writing routines.

This article addresses that gap by proposing a micro-level, process-oriented framework that treats AI-assisted academic writing as a complex adaptive learning system at the level of the individual learner. Drawing on self-regulated learning (SRL) theory[6], AI literacy research[7] and micro-level perspectives from complexity theory[8], it introduces the AI-Regulated Writing Framework (ARWF), which models AI-assisted writing as Zimmerman's cycles of forethought, performance and self-reflection in which AI literacy competences and learner–AI feedback loops jointly shape emergent patterns of writing behaviour. The main contribution is threefold: (1) *theoretically*, ARWF integrates SRL, AI literacy and complex adaptive systems (CAS) reasoning into a coherent model of individual AI-assisted writing processes; (2) *methodologically*, it offers a set of observable indicators (e.g., goal setting, monitoring, critical editing, ethical reasoning) that can be used to study these processes; and (3) *empirically*, it provides an initial illustration of the framework through a small-scale pilot in ESP courses, demonstrating how a brief SRL+AI literacy scaffold can shift students' AI-use patterns and make their self-regulatory processes more explicit.

[5] Cf. Jane SOUTHWORTH et al., Developing a model for AI Across the curriculum: Transforming the higher education landscape via innovation in AI literacy, *Computers and Education: Artificial Intelligence*, 4 (2023) 100127, 1-10, 7.

[6] Cf. Barry J. ZIMMERMAN, Attaining self-regulation: A social cognitive perspective, In: Monique BOEKAERTS et al. (Eds.), *Handbook of self-regulation*, Academic Press, 2000, 13-39.

[7] Cf. Davy Tsz Kit NG et al. Conceptualizing AI literacy: An exploratory review, *Computers and Education: Artificial Intelligence*, 2 (2021) 100041, 1-11.

[8] Cf. Diane LARSEN-FREEMAN, Complexity theory: The lessons continue, In: Lourdes ORTEGA, Zhao Hong HAN (Eds.), *Complexity Theory and Language Development: In celebration of Diane Larsen-Freeman*, John Benjamins, 2017, 11-50.

1. Theoretical Foundations

1.1 Self-Regulated Learning and AI-Assisted Writing

Self-regulated learning (SRL) theory conceptualizes learners as proactive agents who actively manage their cognitive, motivational, and behavioural processes through iterative, goal-directed cycles responsive to feedback. Influential models include Zimmerman's cyclical phases of *forethought, performance, and self-reflection*[9]; Pintrich's motivationally rich framework spanning cognitive, behavioural, and contextual dimensions[10]; and Winne and Hadwin's COPES model, which foregrounds metacognitive monitoring and control[11]. This article adopts Zimmerman's model as its structural backbone given its parsimony and established application in writing research. Forethought involves task analysis, goal-setting and strategic planning; performance encompasses self-control, deployment of task strategies and self-observation; self-reflection includes evaluation of outcomes and adaptation for future efforts. In academic writing, forethought includes analysing assignment requirements and planning structure; performance involves drafting, revising, and seeking feedback; reflection entails evaluating text quality and planning improvements.

Generative AI tools introduce new dynamics into these SRL cycles. Structurally, the forethought phase encompasses learners' motivational beliefs, self-efficacy, and goal planning; the performance phase involves learning strategy use, task engagement, and real-time monitoring; and the self-reflection phase extends to include both reflective evaluation and technological adaptability. Unlike traditional learning contexts, AI's intervention has blurred the boundaries between the forethought and self-reflection phases. Learners may continuously shift between goal adjustment and self-evaluation during task execution, exhibiting a more flexible regulatory cycle[12]. Figure 1 illustrates how the three cyclical SRL phases map onto concrete AI-assisted writing processes, showing the iterative nature of student-AI interactions.

[9] Cf. Barry J. ZIMMERMAN, Becoming a self-regulated learner: An overview, *Theory Into Practice*, 41 (2002) 2, 64-70, 67.

[10] Cf. Paul R. PINTRICH, A Conceptual Framework for Assessing Motivation and Self-Regulated Learning in College Students, *Educational Psychology Review*, 16 (2004) 4, 385-407, 386-390.

[11] Cf. Philip H. WINNE, Allyson F. HADWIN, Studying as self-regulated learning, in: Douglas J. HACKER, John DUNLOSKY, Arthur C. GRAESSER (Eds.), *Metacognition in Educational Theory and Practice*, Lawrence Erlbaum, 1998, 277-304, 279-285.

[12] Cf. Fanfan HOU, Changyin ZHOU, How AI shapes self-regulated learning in foreign language education, *Discovering Computing*, 28 (2025) 344, 1-17, 14.

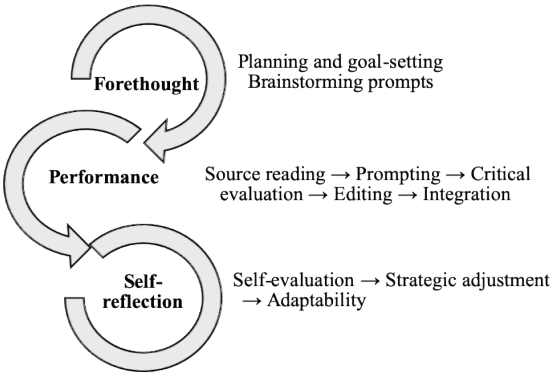


Figure 1. SRL phases mapped to AI-assisted writing processes.

1.2 AI Literacy: Competences for Responsible AI Use

AI literacy can be understood as comprising conceptual understanding of how AI systems work, practical skills for using and evaluating AI tools, and ethical judgement about appropriate and transparent use. In educational contexts, this encompasses basic conceptual knowledge of AI operations and limitations, awareness of bias, the ability to design effective prompts and critically evaluate outputs, and ethical competences related to transparency, fair use, privacy and bias.[13]

Drawing these insights together, we conceptualize four key competences for AI-assisted writing: (1) understanding the limitations of large language models, (2) effective prompt design tailored to specific writing tasks, (3) critical evaluation of AI-generated text against sources for accuracy and completeness, and (4) ethical integration that maintains transparency, protects authorial voice, and upholds academic integrity.

1.3 Complexity Theory: Micro-Level Dynamics in Individual Learning

Complexity theory views learning as an emergent process arising from interactions among multiple cognitive, metacognitive and contextual elements within the learner. Key micro-level complex adaptive system (CAS) concepts relevant for AI-assisted writing include feedback loops, non-linearity, attractor states and self-organisation[14].

[13] Cf. NG et al., *Conceptualizing AI literacy…*; Laura K. ALLEN, Panayiota KENDEOU, ED-AI Lit: An interdisciplinary framework for AI literacy in education, *Policy Insights from the Behavioral and Brain Sciences*, 11 (2024) 7, 3-10; Thomas K. F. CHIU, AI literacy and competency: definitions, frameworks, development and future research directions, *Interactive Learning Environments*, 33 (2025) 5, 3225-3229.

[14] Cf. Diane LARSEN-FREEMAN, Chaos /Complexity Science and Second Language Acquisition, *Applied Linguistics,* 18 (1997) 2, 141-165, 142-146.

It is important to note at the outset that this article employs CAS terminology — including attractors, feedback loops, non-linearity, and self-organisation — as a conceptual heuristic rather than as a formal analytical paradigm. That is, these concepts are used to describe and illuminate patterns in individual learner behaviour in ways that complement SRL and AI literacy perspectives, rather than to conduct dynamic systems modelling in the technical sense. Formal CAS analysis would require longitudinal data, multiple observation points, and computational modelling of system trajectories — none of which are feasible within the scope of this exploratory pilot. This heuristic application of CAS follows an established tradition in applied linguistics and educational research, where complexity concepts serve as productive interpretive lenses without necessitating full dynamical systems methodology.[15]

Feedback loops can be seen in iterative cycles of prompting, AI response, evaluation, editing and refined prompting that gradually led to local adaptation in a learner's writing behaviour. *Non-linearity* means that small changes in initial conditions can produce disproportionate effects in the learner's trajectory. For example, starting a task by reading the source text versus beginning with an AI-generated summary may lead to very different comprehension and writing outcomes. Over time, such cycles can give rise to relatively stable *attractor* patterns of behaviour, such as characteristic AI-use routines like treating AI mainly as a grammar checker, as a ghostwriter, or as a critical thinking partner. Finally, *self-organisation* refers to the tendency of complex systems to develop coherent order through local interactions, without central control. In AI-assisted writing, students can discover effective strategies through trial and feedback, as ongoing interactions between their self-regulation and AI outputs coalesce into more adaptive, stable practices.

More recent scholarship has extended CAS perspectives to technology-enhanced learning environments. Hiver and Al-Hoorie apply dynamic systems principles to L2 motivation and learner engagement, demonstrating how attractor-like stability and variability characterise individual developmental trajectories in classroom contexts[16]. Mercer's work on learner psychology as a complex dynamic system offers a micro-level account of how agency, identity, and self-regulation co-emerge through iterative interactions — a perspective directly relevant to AI-assisted writing[17].

[15] On the heuristic versus formal application of complexity concepts in applied linguistics and language education research, see Kees DE BOT, Wander LOWIE, Marjolijn VERSPOOR, A dynamic systems theory approach to second language acquisition, *Bilingualism: Language and Cognition*, 10 (2007) 1, 7-21, 8, fn. 2, 19.

[16] Cf. Phil HIVER, Ali H. AL-HOORIE, The Dynamic Ensemble for Second Language Motivation: Spinning a Transdisciplinary, Substantive-Methodological Synergy, *The Modern Language Journal*, 100 (2016) 4, 741-756, 742-745.

[17] Cf. Sarah MERCER, The Dynamic Nature of a Language Learner's Self-Concept, in: Sarah MERCER, Marion WILLIAMS (Eds.), *Multiple Perspectives on the Self in SLA*, Multilingual Matters, 2014, 139-158, 140-144.

2. Theoretical Gap and the AI-Regulated Writing Framework

While SRL, AI literacy, and micro-level complex adaptive systems (CAS) offer complementary lenses for AI-assisted writing, they have been treated as parallel rather than integrated frameworks.

Existing SRL research examines AI primarily in terms of how specific tools support the three SRL phases, without explicitly modelling how AI literacy competences shape regulation across those phases. For example, a recent qualitative systematic review synthesizes how chatbots, adaptive feedback systems, games and other applications afford goal setting, strategy use and reflection in higher education, organizing findings by forethought, performance and self-reflection[18]. However, while the review distinguishes human-centred from AI-centred SRL agency, it does not treat students' AI literacy (e.g., understanding limitations, critical evaluation, ethical use) as a separate factor influencing these regulatory processes. Beyond Zimmerman's model, neither Pintrich's motivationally-rich framework nor Winne and Hadwin's metacognitive COPES model has been systematically extended to AI-assisted writing contexts — leaving open questions about how motivational beliefs and real-time monitoring interact with generative AI tools.

On the other hand, most AI literacy studies concentrate on defining technical and ethical competences, without embedding them in cyclical self-regulatory processes. Conceptual frameworks typically distinguish dimensions such as knowing and understanding AI, using AI tools, evaluating AI outputs[19], and attending to ethical and societal implications[20]. These frameworks do not explicitly map AI competences onto SRL phases of forethought, performance and reflection.

Finally, CAS applications emphasize system-level emergence but have not operationalized individual learner-AI interactions as iterative self-organizing cycles of prompt-evaluate-revise-reflect. Larsen-Freeman applies complex dynamic systems theory to second language development, conceptualizing learning as emergent from agent interactions, feedback loops, non-linearity (wherein small changes produce large effects), and attractors representing stability. However, this work predates generative AI technologies and does not address technology-mediated writing processes.

To address this theoretical gap, we propose the AI-Regulated Writing Framework (ARWF), an innovative conceptual model that synthesizes SRL phases, AI literacy competences, and micro-level CAS dynamics into a coherent, operationalizable account of individual student processes during AI-assist-

[18] Cf. Min LAN, Xiaofeng ZHOU, A qualitative systematic review on AI-empowered self-regulated learning in higher education, *npj Science of Learning*, 10 (2025) 21, 1-16, 1.

[19] Cf. Ng et al., *Conceptualizing AI literacy…*, 4.

[20] Cf. Louie GIRAY, Kershnee SEVNARAYANC, Kgabo MAPHOTO, AI Literacy Framework for Academic Writing in the Age of Artificial Intelligence, *Internet Reference Services Quarterly*, 30 (2026) 1-11, 2.

ed academic writing. ARWF addresses this gap by integrating the monitoring emphasis of Winne and Hadwin and the motivational grounding of Pintrich into Zimmerman's cyclical structure, operationalized for AI-assisted writing. ARWF's key innovation lies in reconceptualizing AI not merely as a writing tool but as a co-regulator within SRL cycles, where iterative feedback loops between human metacognition and machine text generation create non-linear developmental trajectories toward stable, adaptive writing patterns.

ARWF uses Zimmerman's classic SRL model—cyclical phases of forethought, performance and self-reflection—as its backbone, preserving the cyclicality essential to self-regulation. This matches AI-assisted writing's iterative nature: students cycle through multiple prompt-revise-reflect loops rather than writing linearly. AI literacy permeates all phases as a meta-competence enabling informed decisions: understanding AI limitations guides forethought planning (e.g. *Should I use AI for outlining or drafting?),* critical evaluation skills shape performance (e.g. *Does this AI output match source technical terms?),* and ethical judgement informs reflection (e.g. *Did I over-rely on AI?).* Furthermore, ARWF conceptualizes each learner as their own complex adaptive system where iterative micro-processes self-organize through three CAS mechanisms. First, feedback loops drive adaptation, as seen in the sequence of prompting, AI output generation, editing, refined prompting, and improved output. Second, non-linearity amplifies small decisions: reading source text first versus AI summarization first produces dramatically different outcomes. Third, attractors emerge as stable AI-use patterns, for example positioning AI as a grammar helper in contrast to an undetected ghostwriter. Figure 2. illustrates ARWF integration, mapping SRL phases onto concrete AI-assisted writing processes with AI literacy threading continuously through cycles and CAS dynamics driving emergent adaptation.

3. The Pilot Study

3.1 Study Aims and Predictions

This pilot study tests whether a brief ARWF scaffold integrating SRL cycles and AI literacy is associated with observable changes self-regulation, critical AI evaluation, and CAS-driven adaptation in students' AI-assisted writing processes. The 15-minute scaffold should produce preliminary positive signals of heightened SRL behaviours, AI literacy markers, and CAS-emergent adaptation patterns across ARWF components, thereby justifying full-scale testing. Given that the scaffold explicitly directs students to set goals, monitor their process, and reflect on AI use, observed changes in these behaviours should be interpreted as indicators of scaffold-aligned engagement rather than direct evidence of underlying metacognitive development.

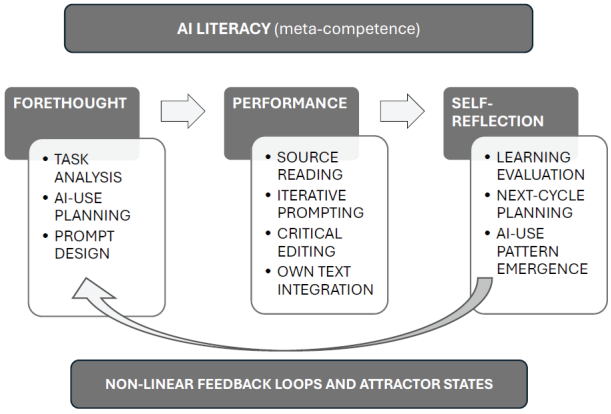


Figure 2. The AI-Regulated Writing Framework

At the micro level, each learner engages in SRL cycles of forethought, performance, and reflection tailored to a specific writing task. Forethought encompasses task analysis, goal setting such as summarising an article's main ideas in 180 words, and planning AI use such as employing AI for grammar checks rather than generating first drafts. Performance covers reading source texts, drafting, prompting AI, editing AI-generated text, and integrating it with one's own writing. Reflection assesses both the produced text and the learning process, including the helpfulness or harm of AI use.

AI literacy is expected to strengthen these cycles, because students with greater literacy are more likely to understand generative models' capabilities and limits, craft prompts for specific sub-tasks, and verify AI outputs against source texts and task requirements. Ethical awareness is likewise expected to guide these decisions by discouraging plagiarism and over-reliance on AI while promoting transparency. Within ARWF's complex adaptive system perspective, even a short SRL- and AI-literacy-oriented scaffold may trigger disproportionate shifts in how students regulate, evaluate, and adapt their AI-assisted writing practices.

3.2 Context, Participants, and Pilot Design

The pilot study took place at Algebra Bernays University in Zagreb, offering undergraduate programs in IT and related fields. 24 undergraduate students from four groups in the researcher's English for IT (ESP) course participated. Each group had two 90-minute writing sessions focused on summarizing IT-related texts. Informed consent was obtained from all participants. AI tools were already informally used by many students, but no prior structured SRL- or AI-literacy-specific teaching related to writing had occurred.

The study comprised two in-class writing tasks:

Task 1 (baseline): Students received a short IT-related article in English (~600 words) and wrote a 150–200-word summary using AI tools freely. They submitted (a) final summary, (b) main AI output used, and (c) reflections on AI use, changes made, and learning impact.

Task 2 (guided): Students received a comparable article/task plus a 15-minute ARWF scaffold (mini-lecture on SRL cycles, AI best practices and ethics). They completed pre-task SRL planning cards and post-task reflections on plan adherence and AI learning.

This minimal-intervention pilot design is methodologically appropriate for framework feasibility testing. Short scaffolds can effectively signal educational interventions in language learning contexts, as shown by single-session pilots reporting high feasibility and acceptability. The exploratory focus on SRL, AI literacy, and CAS markers dispensed with formal hypothesis testing. This approach follows best practices for small-sample pilots (n=24), which prioritise feasibility, fidelity, and preliminary signals over statistical power, as a range of 10–30 participants is suitable[21].

3.3 Data analysis

Data sources included: student summaries (Task 1 and Task 2), AI outputs submitted alongside summaries, reflection responses from both tasks, and pre-task SRL cards from Task 2. Analysis focused on three ARWF components using mixed methods.

Students' summaries were graded with a 4-criteria analytic rubric, with each criterion rated from 0 to 3: *content accuracy* (faithfulness to the source text), key idea coverage (inclusion of major sections and main points), *coherence and organisation* (logical structuring, paragraphing, and use of linking devices), and *language quality* (grammar, vocabulary, and appropriateness of paraphrasing). For each task, total rubric scores (0–12) were computed by summing the four analytic criteria and then used to calculate means and standard deviations. To examine change over time, individual gain scores were computed by subtracting each student's Task 1 total from their Task 2 total. The distribution of gain scores was then summarised with its mean and standard deviation, and a matched-pairs effect size (Cohen's *d*) was calculated by dividing the mean gain by the standard deviation of the gain scores. This provided an estimate of the magnitude of improvement in summary performance from Task 1 to Task 2.

[21] Cf. Patsawut SUKSERM, Determining the Appropriate Sample Size in EFL Pilot Studies, *Journal of Research Methodology*, 37 (2024) 3, 245-264, 256.

For a stratified subsample (n=12; 50% of participants), differences between the submitted AI outputs and the corresponding final summaries were inductively coded into three mutually exclusive categories: language-level edits (e.g., word choice, grammar), structural edits (e.g., organisation, sentence/paragraph boundaries), and content-level edits (e.g., adding or removing ideas, integrating or correcting source information). For each task, the proportions of these three edit types were calculated per student and then summarised at the group level to gauge the depth and focus of learners' engagement with AI-generated text.

Reflection texts and SRL cards were deductively coded using two binary coding schemes aligned with ARWF theory (see Table 1). Coding focused on four SRL elements and four AI-literacy elements, operationalized as eight binary codes (present=1/absent=0). Codes of 48 student reflections (24 Task 1 + 24 Task 2) were analysed. To assess inter-rater reliability, a second independent rater — a colleague with expertise in applied linguistics and no involvement in the study design or instruction — coded a randomly selected subsample of 25% of all reflections (n=12 out of 48, stratified across Task 1 and Task 2). The rater received the ARWF Binary Coding Scheme (Table 1) with definitions and examples but no additional guidance. Agreement between the two raters was calculated using Cohen's kappa (κ). Results indicated good reliability across all eight codes (κ = .78), confirming that the coding scheme is sufficiently transparent and replicable for use in future studies.

Table 1. ARWF Binary Coding Scheme

Element	Code	Definition	Reflection examples
SELF-REGULATED LEARNING			
Forethought	Explicit task-specific learning objective	States a clear intention for learning or performance	*my goal…, practice…, aimed to…*
Performance 1	Mid-task checking/ adjustment	Mentions checking understanding or adjusting approach during task	*followed plan…, noticed…, checked…*
Performance 2	Deliberate sequence/ steps	Describes purposeful sequence of actions or strategies	*first…then…, outline then draft…*
Reflection	Outcome evaluation/ learning gained	Evaluates results or expresses gains in understanding	*learned that…, realized…, improved…*
AI-LITERACY			
Limitations Aware	Acknowledges AI constraints	Recognizes flaws or limits of AI	*hallucinates, trash, generic, mistakes, errors…*
Critical Evaluation	Verifies against source	Cross-checks AI output with reliable information	*checked article, verify facts, corrected…*
Over-Reliance Concern	Worries about dependence	Expresses caution about overusing AI or losing skills	*depend too much, lazy, not improve own…*
Ethical Integration	Frames as tool (not replacement)	Positions AI as an aid or complement to own work	*editor, scaffold, advisor, polish…*

Frequencies that represent the proportion of students exhibiting each code were calculated with Cohen's *d* effect sizes for Task 1→Task 2 shifts. Drawing

on CAS as a heuristic lens, attractor-like patterns were approximated via total element gain scores across the eight ARWF binary codes. Students with a gain of ≥ 3 elements were classified as exhibiting a high-shift profile, provisionally consistent with movement toward a more adaptive AI-use pattern; those with a gain of ≤ 1 were classified as exhibiting a low-shift profile, provisionally consistent with persistence of a surface-level AI-use pattern. These are explicitly descriptive classifications indicating direction of change, not formal evidence of attractor dynamics, which would require longitudinal data.

4. Results

4.1 Changes in Summarizing Proficiency and Editing Focus

Students' summary performance was markedly higher in Task 2 than in Task 1. On Task 1, the mean total score (maximum=12) was 8.08 (SD=1.95), whereas on Task 2 the mean increased to 9.96 (SD=1.23). The average gain score (Task 2 – Task 1) was 1.88 points (SD=1.03), indicating that most students showed improvement rather than isolated individual gains. A matched-pairs effect size based on the gain scores yielded Cohen's d=1.82 (≥0.8[22]), which represents a very large pre- to post-task difference in rubric-based summary quality.

To ensure diversity without bias in this exploratory pilot, a subsample comprising 50% of participants (n=12) was stratified by baseline performance on Task 1. Participants were divided into tertiles based on summary scores (high, medium, low), with proportional sampling (n=4 per tertile) to capture variance in editing potential. Results from the stratified subsample of 12 students suggest clear differences in how students engaged with AI outputs across tasks. In Task 1, edits primarily consisted of language-level changes (M=65% of all edits), including vocabulary simplification, grammar corrections, and word order adjustments, whereas structural edits (M=20%) and content-level edits (M=15%) were relatively infrequent. By contrast, Task 2 showed a marked reduction in language-level edits to M=45% (a decrease of 20 points), accompanied by modest gains in structural edits (M=25%; +5 points) and a substantial increase in content-level edits—the largest shift observed (M=30%; +15 points, nearly doubling from baseline). Table 2 shows group-averaged proportions in editing patterns with the change value. Negative values indicate reduction in edit-type proportion; positive values indicate increase, which reflects a shift from surface-level to substantive engagement with AI outputs.

[22] Cf. Jacob COHEN, *Statistical Power Analysis for the Behavioral Sciences* (Second Edition), New York, Lawrence Erlbaum Associates, 1988, 26.

Table 2. Editing Patterns (Subsample n=12)

Type	Task 1 M %	Task 2 M%	Change (% point)
Language-level	65	45	-20
Structural	20	25	+5
Content-level	15	30	+15

4.2 SRL and AI-literacy indicators in reflections

The analysis of reflective data demonstrates a substantial increase in both SRL and AI-literacy indicators following the scaffolded intervention. As shown in Table 3, all SRL components improved dramatically—from baseline frequencies near one-third to post-task levels exceeding 80%—with the largest relative gains in reflection (+59%) and monitoring (+58%). The overall SRL increase of 48%, supported by a large effect size (Cohen's d=1.92), indicates that students displayed substantially more behaviours consistent with metacognitive engagement following the scaffold. It should be noted, however, that since the scaffold explicitly prompted these behaviours, the observed frequencies may partly reflect students' compliance with task instructions rather than spontaneous metacognitive development. Similar though slightly smaller effects appear in AI-literacy (Table 4), with total gains of 30% and a large effect size (Cohen's d =1.20). The greatest improvements occurred in recognizing AI limitations (+41%) and framing AI as critical support rather than replacement (+46%), implying improved judgment and responsible integration of AI tools.

Finally, CAS analysis (Table 5) shows that 58% of participants exhibited a high-shift profile toward *critical support* attractor pattern, suggesting a directional change from passive AI use toward more reflective, self-regulated engagement. For example, in the first task one student used AI almost only as a grammar checker, accepting the output with minimal cosmetic changes. In the second task, the same student first read the source article, then requested alternative formulations for key ideas and checked each sentence against the source and their own notes. This shift illustrates a move from a 'grammar checker' pattern toward a more stable attractor in which AI functions as a critical partner.

Table 3. SRL Elements Frequency (n=24)

Element	Task 1 n (%)	Task 2 n (%)	Shift (+/-%)
Goal setting	9 (38%)	22 (92%)	+54%
Monitoring	6 (25%)	20 (83%)	+58%
Strategy Use	10 (42%)	21 (88%)	+46%
Reflection	8 (33%)	22 (92%)	+59%
Total	1.4 (35%)	3.3 (83%)	+48% (Cohen's d = 1.92)

Table 4. AI-Literacy Elements Frequency (n=24)

Element	Task 1 n (%)	Task 2 n (%)	Shift (+/-%)
Limitations Aware	4 (17%)	14 (58%)	+41%
Fact checking	2 (8%)	12 (50%)	+42%
Over-Reliance Concern	6 (25%)	10 (42%)	+17%
AI Support (not replacement)	5 (21%)	16 (67%)	+46%
Total	0.6 (15%)	1.8 (45%)	+30% (Cohen's d = 1.20)

Table 5. CAS-Inspired Behavioural Shift Profiles (n=24)

Profile	n	%
High Shift (≥+3)	14	58%
Medium (+2)	5	21%
Low Shift (≤+1)	5	21%

5. Discussion

The pilot study provides preliminary empirical support for the AI-Regulated Writing Framework, suggesting that a brief 15-minute scaffold integrating SRL cycles and AI literacy competences was associated with observable shifts in undergraduate ESP students' AI-assisted writing processes. Large effect sizes in summary proficiency (Cohen's d = 1.82), SRL markers (d = 1.92), and AI-literacy indicators (d = 1.20) are consistent with enhanced self-regulation, alongside a 58% shift toward adaptive *critical support* attractors in CAS analysis. These findings are consistent with ARWF's core proposition: AI acts as a co-regulator within iterative SRL phases, where AI literacy enables feedback loops that may support non-linear adaptation at the individual learner level.

Quantitative improvements in rubric scores—from surface-level language edits (65% in Task 1) to substantive content and structural revisions (combined 45% in Task 2)—reflect the performance phase dynamics of the ARWF. Analyses of editing patterns reveal that several students progressed from merely *cleaning up* AI-generated text to reorganizing paragraphs, reshaping topic sentences, and adding or deleting ideas to align more closely with the source article. Collectively, these patterns demonstrate a shift from surface-level polishing in Task 1 to deeper, more substantive engagement with AI-generated content in Task 2, consistent with observed gains in rubric-based performance. This transition from passive AI acceptance to active integration is consistent with the CAS heuristic of non-linearity, whereby small changes co-occurring with the scaffold (e.g., explicit goal setting through pre-task cards) amplified into enhanced engagement, such as source-verification prompting. Such dynamics mirror Larsen-Freeman's conceptualization of micro-level adaptation

in language development, wherein local rules (e.g., iterative checks) produce disproportionately significant outcomes.

Reflective data further validates ARWF integration. Substantial SRL gains suggest the scaffold prompted students to make forethought and self-reflection visible in their writing, though it remains an open empirical question whether these gains reflect lasting metacognitive change or task-specific compliance with scaffold instructions. AI-literacy surges—particularly limitations awareness (+41%) and fact checking (+42%)—positioned AI as a scaffold rather than replacement, fostering ethical co-regulation. High-shift profiles (58%) suggest attractor-like patterns consistent with an *AI as critical partner* orientation, aligning with recent AI-SRL synergies where literacy training clusters learners into adaptive profiles[23].

These patterns are consistent with ARWF's theoretical novelty: unlike models that map AI affordances separately across SRL phases or literacy dimensions, it operationalizes human-AI interactions as self-organizing cycles testable via observable indicators (e.g., editing depth, reflection codes).

Results extend AI-education research by grounding ethical/tool debates in process-oriented SRL-CAS. Similar pilots in ESP/EFL contexts report associations between scaffolds and foundational writing/metacognition (e.g., ChatGPT gains in proficiency/autonomy[24]). However, ARWF uniquely operationalizes CAS-inspired attractor profiles as observable behavioural indicators, offering a pragmatic methodological starting point for IT-ESP contexts where technical accuracy demands critical evaluation.

6. Limitations

A central methodological limitation concerns the potential confound between scaffold instructions and outcome measurement. The Task 2 scaffold explicitly directed students to set goals, plan AI use, monitor their writing process, and reflect on outcomes — precisely the behaviours subsequently coded in their reflections as SRL and AI-literacy indicators. Consequently, observed increases (e.g., monitoring from 25% to 83%) may partly or wholly reflect students' alignment with task instructions rather than genuine metacognitive development. Future studies should address this confound by incorporating measurement instruments independent of scaffold prompts, such as think-aloud protocols during writing, AI prompt logs, or revision traces captured via keystroke logging, which would provide behavioural evidence less susceptible to demand characteristics.

[23] Cf. Hou, Zhou, *How AI shapes…*, 14.

[24] Cf. Amr M. MOHAMED et al., ChatGPT's impact on ESP writing proficiency and learner autonomy, *Technology in Language Teaching & Learning*, 7 (2025) 3, 1-25, 14.

As an exploratory pilot, findings are preliminary and non-generalizable. The small sample (n=24) limits statistical power, precluding inferential tests; effect sizes, while large, may inflate without controls[25]. Absence of a comparison group confounds intervention effects from practice. Short-term design (two sessions) captures immediate shifts but not stability of attractors. Self-reports risk social desirability bias, and single-instructor implementation at one institution constrain transferability to diverse EFL/ESP settings.

Conclusion

The framework and pilot study yield several implications for ESP instruction and research. First, tasks should prioritize human-led co-regulation by embedding SRL processes to guide AI use, such as requiring pre-task planning and post-task reflection, ensuring students maintain agency rather than deferring to AI outputs. Second, AI literacy should be cultivated as a core competence through brief, integrated activities addressing AI capabilities/limitations, prompt engineering, and ethical considerations, complementing domain-specific instruction to promote reflective, responsible tool integration.

Future research should employ larger samples and controlled experimental designs, including comparisons between courses receiving structured SRL and AI-literacy scaffolding and those without such interventions, alongside advanced analyses of text revisions and learning outcomes. Longitudinal studies are recommended to examine the stability of attractor states across academic semesters. The framework should be extended to diverse ESP domains (e.g., media, business) and varying proficiency levels (A2–C1). Finally, psychometric validation of ARWF indicators is essential to support their application in broader educational technology contexts.

The exploratory pilot in ESP courses provides an initial, context-specific illustration of how small changes in local classroom rules can shift AI-use patterns and make learners' self-regulatory processes more visible. The hope is that this perspective will support more nuanced, educationally grounded debates and studies about AI-assisted writing in higher education.

[25] Cf. Erik WESTLUND, Elizabeth A. STUART, The Nonuse, Misuse, and Proper Use of Pilot Studies in Experimental Evaluation Research, *American Journal of Evaluation,* 38 (2017) 2, 246-261, 256.

Tihana Banko*

Okvir regulacije pisanja umjetnom inteligencijom: integracija samoreguliranog učenja, pismenosti u području umjetne inteligencije i dinamike složenih adaptivnih sustava

Sažetak

Ovaj članak donosi prijedlog Okvira regulacije pisanja umjetnom inteligencijom, mikro-razinske konceptualizacije koja integrira faze samoreguliranog učenja – predviđanje, izvođenje i samorefleksiju – te kompetencije pismenosti u području umjetne inteligencije (UI), kao i dinamiku složenih adaptivnih sustava radi razumijevanja individualnih procesa akademskog pisanja uz podršku UI u kontekstima engleskog jezika struke. Model je predložen zbog nedostataka dosadašnjih istraživanja, koja su pretežno usmjerena na alate UI i etička pitanja, zanemarujući dinamične interakcije između učenika i UI. U skladu s time, pilot istraživanje je preliminarno ispitivalo učinak kratke, 15-minutne podrške koja povezuje faze samoreguliranog učenja i pismenost u području UI. Cilj je bio utvrditi povećava li takva podrška samoregulaciju, kritičku evaluaciju alata UI i prilagodbu vođenu složenim adaptivnim sustavima u pisanju uz potporu UI. Rezultati upućuju na značajna poboljšanja: vještine sažimanja (d=1,82), pokazatelja samoreguliranog učenja (d=1,92; npr. porast učestalosti praćenja vlastitog procesa pisanja i korištenja umjetne inteligencije s 25 % na 83 %), pismenosti u području UI (d=1,20; npr. porast svijesti o ograničenjima +41 %) te pomaka od 58 % prema adaptivnim obrascima složenih adaptivnih sustava. Proces uređivanja teksta generiranog alatima UI prelazi s pretežno površinskih zahvata (65 % u Zadatku 1) prema sadržajnim (30 %) i strukturnim promjenama (25 %), koje zajedno čine 55 % zahvata u Zadatku 2. Ovi nalazi pružaju preliminarne pokazatelje konzistentne s predloženim okvirom te upućuju na potrebu za zadacima koji potiču samoregulirano učenje uz potporu UI i sustavnom izobrazbom u području pismenosti UI u kontekstima visokog obrazovanja.

Ključne riječi: okvir regulacije pisanja umjetnom inteligencijom; pisanje na engleskom jeziku struke; pismenost u području umjetne inteligencije; samoregulirano učenje; složeni adaptivni sustavi.

* Dr. sc. Tihana Banko, predavačica, Sveučilište Algebra Bernays, Katedra za interdisciplinarne znanosti, Gradišćanska 24, HR-10000 Zagreb; e-mail: tihana.banko@algebra.hr.