Journal of Modern Research in English Language Studies
Print ISSN: 2676-5357
Online ISSN: 2676-5985

Integrating Self-Regulated Learning in Flipped Learning to Enhance Argumentative Writing with Digital Tools

Leyli Nouraei Yeganeh[1*], Majid Nemati[2]

[1*](Corresponding author) Department of Foreign Languages and Literature, University of Tehran; Tehran, Iran. *leili.yeganeh@ut.ac.ir*
[2]Department of Foreign Languages and Literature, University of Tehran; Tehran, Iran. *nematim@ut.ac.ir*

Article info

Article type:
Research article

Received:
2025/03/14

Accepted:
2025/06/24

Abstract

This quasi-experimental study investigated the integration of flipped learning (FL) with self-regulated learning (SRL) strategies to enhance argumentative writing (AW) skills among 240 intermediate English-proficient medical students at an Iranian university during Fall 2023. The study aimed to determine whether embedding SRL strategies into FL environments would lead to greater improvements in AW proficiency compared to FL alone. Participants were randomly assigned to an experimental group (FL with SRL) and a control group (FL without SRL). In the experimental group, digital tools such as Google Docs, EdPuzzle, and Padlet were aligned with SRL phases to facilitate goal-setting, pre-class preparation, and reflective critique. Results indicated that while both groups improved, students who engaged with SRL-enhanced FL instruction demonstrated greater gains in AW and self-regulatory skills. These findings suggest that integrating SRL strategies into FL models can foster more effective writing development and learner autonomy in English for specific purposes (ESP) contexts, offering a scalable instructional approach for medical education. The instructional design and digital tool alignment proposed in this study can be adapted to support writing development and learner autonomy across various global ESP programs and diverse educational contexts.

***Keywords*:** argumentative writing, digital tools, flipped Learning, medical education, self-regulated learning

Cite this article: Nouraei Yeganeh, L., & Nemati, M. (2026). Integrating self-regulated learning in flipped learning to enhance argumentative writing with digital tools. *Journal of Modern Research in English Language Studies, 13* (1), 171-195.

DOI: 10.30479/jmrels.2025.21792.2495

1. Introduction

Flipped Learning (FL) has emerged as a transformative educational approach, particularly in language instruction. This shift promotes active learning opportunities and supports deeper engagement with course materials, especially in language education contexts (Aidoo et al., 2022). This model has gained increasing relevance in English for Specific Purposes (ESP) contexts, including medical education, due to its potential to enhance learning outcomes, boost student motivation, and promote active engagement (Li et al., 2023). FL enables students to prepare independently by providing flexible access to pre-class materials through digital platforms. Consequently, in-class sessions focus on practical applications, such as collaborative writing or peer critique (Padovano et al., 2024). Despite these advantages, effectively implementing FL presents challenges. Students often struggle with unstructured pre-class activities, insufficient feedback, and a lack of self-discipline during independent learning (Cevikbas & Kaiser, 2023). These challenges are particularly problematic for mastering complex skills such as Argumentative Writing (AW), where structured support and continuous feedback prove crucial.

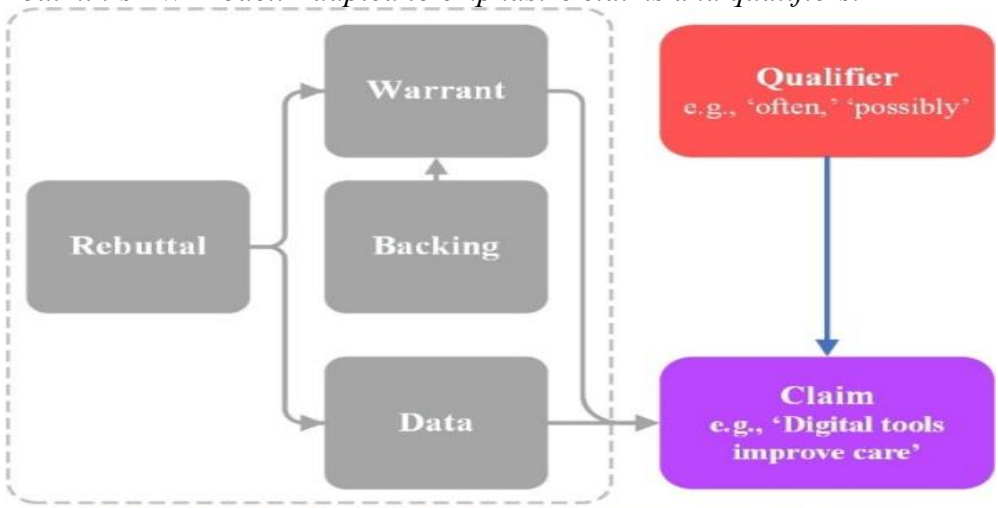
Self-Regulated Learning (SRL), as defined by Zimmerman (2008), offers a promising framework to tackle these challenges. SRL includes the cognitive, metacognitive, motivational, behavioral, and emotional processes that enable learners to set goals, monitor their progress, and reflect on outcomes (Sanz-Angulo et al., 2025; Zare et al., 2024). Incorporating SRL strategies into FL environments may assist students in navigating the independent demands of flipped classrooms more effectively. However, research exploring the integration of FL and SRL specifically to enhance AW skills remains limited (Padovano et al., 2024; Weng et al., 2024). Precise, evidence-based writing is essential in ESP settings, especially in medical education. Medical students must express complex arguments in English as a foreign language (EFL) for research abstracts, clinical analyses, and policy recommendations (Deng et al., 2024; Lotfi Gaskaree et al., 2025). Nevertheless, students with intermediate English proficiency often view AW as tedious and challenging, struggling with both structural and rhetorical demands (Teng & Wang, 2023). Traditional instruction often lacks sufficient iterative practice and feedback, emphasizing the need for innovative pedagogical approaches (Mekheimer, 2025).

Digital tools offer promising solutions by supporting structured writing practice. Collaborative platforms like Google Docs facilitate drafting and feedback, interactive tools like EdPuzzle enable pre-class comprehension checks, and reflective platforms like Padlet encourage post-task critical thinking (Jin, 2024). While FL models have enhanced grammar and vocabulary acquisition (Sanz-Angulo et al., 2025), their potential to improve

higher-order skills like AW, particularly when aligned with SRL strategies, requires further empirical validation (Sarwar et al., 2024). Importantly, the instructional challenges and solutions explored in this study are not unique to the Iranian context; they reflect broader trends in global ESP and medical education settings, where fostering autonomous, proficient writers remains a critical priority. This study addresses these gaps by investigating whether integrating SRL strategies into FL instruction enhances AW skills among medical students more effectively than FL alone. Specifically, the study builds on Toulmin's model of argumentation to analyze improvements in claims and qualifiers, two critical components of effective academic writing. Additionally, it examines the impact of the FL-SRL model on students' self-reported learning skills, including motivation, planning, assessment, and self-directedness (see Figure 1).

Figure 1

Toulmin's AW Model. Adapted to emphasize claims and qualifiers.



Note: Focus on claims and qualifiers; other elements excluded. 20% of pilot essays omitted warrants.

The study prioritizes claims and qualifiers because they represent foundational elements in constructing coherent, persuasive arguments, especially for learners at an intermediate English proficiency level. Given the cognitive load required to formulate complete warrants and rebuttals, focusing first on strengthening claims and qualifiers allows for more manageable, measurable progress in AW skills essential for medical ESP contexts. This targeted focus provides a necessary scaffold before introducing more complex Toulmin elements in future instructional phases. This study uniquely operationalizes SRL phases by systematically aligning them with targeted digital tools. It integrates Google Docs for explicit goal-setting and

collaborative drafting during the forethought and performance phases. EdPuzzle supports self-monitoring during pre-class preparation, and Padlet structures a reflective critique post-task. This intentional pairing of SRL strategies with specific digital technologies distinguishes the current study, offering a practical, replicable framework for enhancing higher-order AW skills in ESP contexts.

This study conceptualizes AW development through the dual lens of Toulmin's (2003) argumentation model and Zimmerman's (2008) SRL framework. While Toulmin's model provides a systematic approach to analyzing the structural components of arguments, particularly claims and qualifiers, Zimmerman's SRL model outlines the cognitive, motivational, and metacognitive processes learners use to plan, execute, and evaluate their learning. We posit that aligning SRL phases (forethought, performance, and reflection) with the instructional focus on Toulmin elements can scaffold learners' development of higher-order writing skills. Specifically, goal-setting and strategic planning (forethought) can support claim formulation, while reflective phases can reinforce the nuanced use of qualifiers. This theoretical integration guides the intervention design and the study's analytical focus, offering a cohesive framework for enhancing AW in ESP contexts.

This study aims to evaluate the effectiveness of integrating SRL strategies into FL instruction to enhance AW skills and promote learner autonomy in medical ESP education. The following research questions guide the investigation:

1. Does integrating SRL strategies into FL instruction enhance AW skills more effectively than FL alone?
2. How do claims and qualifiers in AW improve under the FL-SRL approach?
3. How does integrating digital tools with SRL strategies influence students' self-reported learning skills in a FL environment?

By strategically leveraging digital platforms within the FL-SRL framework, this study aims to provide practical guidance for ESP instructors and contribute to a broader understanding of technology-mediated writing development in higher education.

2. Literature Review

2.1. Flipped Learning in Language Education

FL significantly influenced language education by shifting foundational content delivery, such as grammar instruction and writing conventions, to pre-class activities using digital resources (Aidoo et al., 2022). This model freed up in-class time for active learning, including collaborative writing tasks, peer review sessions, and interactive discussions, thus fostering deeper engagement and practical skill development (Li et al., 2023). In ESP

contexts, particularly in medical education, FL provided flexible access to content, allowing students to engage with materials asynchronously while enabling instructors to design targeted, application-based activities like drafting clinical reports or policy recommendations (Padovano et al., 2024).

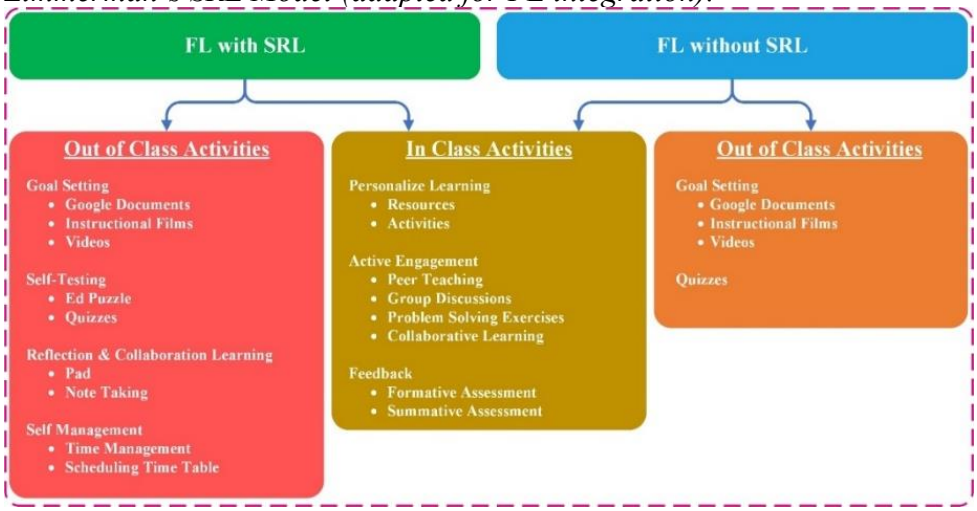
While several studies reported gains in vocabulary acquisition and student motivation (Sanz-Angulo et al., 2025; Wong & Liem, 2022), these advancements were often linked to foundational language skills rather than higher-order academic writing. However, challenges persisted: research by Cevikbas and Kaiser (2023) indicated that students frequently struggled with pre-class preparation, particularly when tasks lacked structure or adequate feedback. Notably, many studies emphasized descriptive outcomes while neglecting long-term retention, variations in student engagement levels, or the contextual dependency of FL's effectiveness. These findings suggested that while FL can enhance basic language skills, its efficacy in developing complex skills like AW remains contingent upon additional support mechanisms.

2.2. Self-Regulated Learning and Its Role in FL

SRL, as conceptualized by Zimmerman (2008), involved a cyclical process of forethought (goal-setting and planning), performance (self-monitoring and strategy use), and self-reflection (evaluation and adaptation). SRL frameworks empowered learners to take greater ownership of their progress, which is critical in the autonomous learning environments that FL created (Zare et al., 2024). Empirical evidence indicated that integrating SRL strategies into FL contexts can significantly enhance academic outcomes. For instance, Zariholhosseini et al. (2025) found that students who employed SRL strategies during FL interventions achieved higher writing quality and greater task engagement than their peers relying on traditional instructor feedback. Digital tools amplified these benefits: Google Docs facilitated real-time goal-setting and collaboration, EdPuzzle supported self-assessment through embedded quizzes, and Padlet encouraged reflective practice and peer feedback (Jin, 2024; Lee et al., 2024). While research documented SRL's role in enhancing general language performance (Öztürk & Çakıroğlu, 2021), its specific application to AW development within FL environments remained underexplored. Particularly in ESP contexts, where precision and critical reasoning were crucial, understanding how SRL strategies interact with FL to foster AW proficiency was essential (Weng et al., 2024). However, much of the current literature lacked critical methodological rigor, often relying on self-reported improvements without triangulating these with performance-based assessments or instructor observations. Additionally, conflicting findings existed regarding SRL's scalability, especially in resource-limited environments where students encountered uneven digital literacy or limited access to feedback mechanisms (see Figure 2).

Figure 2

Zimmerman's SRL Model (adapted for FL integration).



2.3. Argumentative Writing in ESP and Medical Education

AW required higher-order skills such as constructing coherent claims, integrating evidence, and addressing counterarguments—skills that challenged many ESP learners, particularly those in medical education programs (Teng & Wang, 2023). In clinical and academic settings, students must make nuanced arguments to advocate for treatments or debate ethical dilemmas (Deng et al., 2024). Toulmin's model (2003) provided a structured framework for analyzing argument quality, identifying essential components such as claims, grounds, warrants, qualifiers, rebuttals, and backing. Research by Rapanta et al. (2025) indicated that second-language (L2) learners often overgeneralized arguments (e.g., omitting qualifiers) and struggled with evidentiary support, resulting in simplistic or absolutist writing. Yang (2022) similarly found that L2 learners frequently produced vague claims and failed to justify their arguments through evidence and reasoning.

Recent studies further confirmed these challenges in medical education: Yeganeh et al. (2025) demonstrated that medical students' difficulties in structuring claims and supporting evidence significantly hindered their ability to produce persuasive clinical case analyses. Similarly, Lacina et al. (2025) emphasized the need for explicit instruction in argument structure to improve medical students' academic writing performance, particularly when responding to ethical dilemmas and treatment debates. Prior interventions focused on general writing skills rather than explicitly targeting AW components, leaving a gap in strategies designed to strengthen claims and qualifiers essential for developing critical thinking and argument sophistication. Moreover, while FL and SRL each supported language development individually, their combined application to foster higher-order

AW skills remained largely underexplored. Existing studies rarely aligned FL's flexible instructional model with SRL's structured goal-setting, monitoring, and reflection processes to systematically scaffold AW development. Moreover, limited research existed on how learners transferred these academic argumentation strategies across genres or disciplines, highlighting the need for more longitudinal studies that assessed enduring learning. This study addresses this gap by investigating how integrating SRL strategies within an FL environment can enhance critical components of academic argumentation among ESP learners.

2.4. Digital Tools in FL and AW Development

Digital technologies offered specific affordances for supporting AW development within FL models. Google Docs facilitated collaborative drafting and iterative feedback, with Maghsoudi and Mansouri (2025) reporting a 15% improvement in writing coherence when peer editing was integrated into digital platforms. Similarly, EdPuzzle enhances pre-class preparation by embedding formative assessment questions in videos, improving conceptual understanding and recall (Sanz-Angulo et al., 2025). Padlet supports post-task reflection and critical analysis, with Lee et al. (2024) noting that 78% of users revised arguments after engaging with peer feedback. Moreover, Hwang (2025) emphasized that aligning digital tools with SRL phases—goal-setting, monitoring, and reflection—enhanced learner autonomy and writing self-efficacy. In a multimodal learning environment, tools like smartphones and laptops facilitated access to content and collaborative activities, increasing engagement and completion rates (Cui & Zhang, 2025). Despite these reported benefits, existing literature often lacks direct comparisons of tool effectiveness across SRL phases or detailed analysis of learner interaction patterns with these tools. Additionally, some studies have generalized findings from basic writing or grammar exercises to more advanced tasks like AW, which may have overstated the impact of digital tools on complex writing development (Deng et al., 2024).

2.5. Research Gaps

While FL promoted engagement and SRL fostered autonomous learning, their combined application to enhance AW skills among medical students has remained an underexplored area. Previous research has largely overlooked the need for structured interventions that support argumentative structure and critical reasoning in ESP writing (Mekheimer, 2025). Moreover, interventions that leveraged SRL principles—explicit goal-setting before writing, strategic monitoring during drafting, and structured reflection post-feedback—are scarce (Huang & Chen, 2022). This study addresses these gaps by investigating the effects of embedding SRL strategies into FL environments

to improve AW. Specifically, it focuses on enhancing the use of claims and qualifiers—two critical but often neglected components of L2 students' argumentative essays (Chen, 2025). The current study bridges two under-connected frameworks: Toulmin's model and Zimmerman's SRL by targeting these elements through a theoretically grounded design. Furthermore, it contributes a replicable intervention model to advance writing instruction in the medical ESP context. Additionally, by employing accessible digital tools, the study offers a scalable model that aligns with calls for technology-mediated ESP pedagogy (Ngo & Hastie, 2025).

3. Method

This study employed a quasi-experimental pre-test/post-test design over 16 weeks in Fall 2023 at the University of Kashan in Iran. It investigated the effectiveness of integrating FL with SRL strategies to enhance AW skills among medical students. The researchers randomly assigned the participants to an experimental group (FL with SRL) or a control group (FL without SRL), dividing both groups into two classes of approximately 60 students each to ensure instructional feasibility and fidelity. The intervention followed Zimmerman's (2008) SRL model, utilizing digital tools to support the forethought, performance, and reflection phases of learning.

3.1. Participants

The participants comprised 240 medical students (aged 18–22) enrolled in a mandatory ESP course. Students demonstrated intermediate English proficiency (B2 CEFR level), confirmed by a placement test. The sample was 55% female (n = 132) and 45% male (n = 108), reflecting the program's demographic profile. An experienced ESP instructor taught the experimental and control groups, ensuring consistency in teaching style, materials, and classroom management. Pre-intervention surveys revealed that 95% of students were unfamiliar with FL models, and 85% reported difficulties with AW tasks, highlighting the need for targeted support. Pre-test AW scores confirmed baseline equivalence between groups (experimental: $M = 8.73$, $SD = 2.26$; control: $M = 8.50$, $SD = 2.30$; $t(238) = 0.79$, $p = 0.42$). To maintain homogeneity, students with advanced or beginner proficiency levels were excluded (n = 48). Before participating, all students signed informed written consent forms following the institutional ethical guidelines approved by the university's research ethics committee. After confirming eligibility, researchers randomly assigned the participants to the experimental or control group in a 1:1 ratio using a computer-generated random number list. A research assistant not involved in instruction or assessment conducted the randomization independently, ensuring allocation concealment and minimizing selection bias. The resulting groups were balanced in terms of

gender and English proficiency level. The same instructor taught the experimental and control groups using identical core content, class schedules, and assessment rubrics to mitigate potential instructor effects as confounding variables. The only difference between groups was the delivery method (SRL integration versus standard FL). Furthermore, an independent observer recorded and randomly reviewed all classroom sessions to ensure instructional fidelity and adherence to the assigned intervention protocols.

3.2. Materials and Instruments

3.2.1. Homogeneity Test

A 60-minute placement test evaluated reading, writing, and grammar skills, scoring 100 points. Two independent raters established intermediate proficiency (50–75) according to CEFR standards and achieved high inter-rater reliability (Cohen's kappa = 0.85). The test showed strong internal consistency (Cronbach's $\alpha = 0.89$).

3.2.2. Argumentative Writing Tests

Students completed pre- and post-test essays responding to the prompt: "Has digital technology enhanced the gap in healthcare service quality worldwide?" and did the writing tasks digitally through Google Docs during in-class sessions under instructor supervision to ensure standardization and prevent academic dishonesty. Essays were scored across four dimensions—claim, qualifier, evidence, and coherence—using a 0–5 rubric. Researchers piloted the rubric with 15 essays to ensure clarity. Unaware of group assignments, two raters evaluated all essays, with discrepancies resolved through discussion (inter-rater reliability: Cohen's kappa = 0.87). To maintain inter-rater reliability, both raters participated in a two-stage training process. First, they jointly reviewed and discussed the scoring rubric and sample essays to align their interpretations of each criterion. Then, each rater independently scored a set of 20 pilot essays, after which scores were compared and discrepancies were discussed to reach consensus. Only after achieving at least 85% agreement on the pilot set did the raters proceed to evaluate the study essays independently. Periodic cross-checks during scoring helped ensure consistency throughout the assessment process.

3.2.3. Self-Regulated Questionnaire (SRQ)

The SRQ (Turan et al., 2009) comprises 41 items divided into four subscales: motivation, planning, assessment, and self-directness. The SRQ was administered digitally through Google Forms at the beginning and end of the course, ensuring all responses remained anonymous. A 5-point Likert scale was employed, with higher scores showing stronger SRL skills. The SRQ exhibited high reliability (Cronbach's $\alpha = 0.85$ in this study) and has been

previously validated for use in ESP and medical education contexts (Fakhri et al., 2023), further supporting its suitability for this population.

3.2.4. Digital Tools

Each digital tool was strategically aligned with a specific SRL phase:

- **Forethought (Goal-Setting):** The study used Google Docs to facilitate weekly goal-setting activities (e.g., "Include two qualifiers in my claim."). It tracked student edits, and instructors provided written feedback within 48 hours.
- **Performance (Monitoring and Strategy Use):** EdPuzzle provided brief pre-class videos integrated with quizzes that enabled students to assess their comprehension before class discussions.
- **Reflection (Evaluation and Adjustment):** Padlet hosted biweekly reflective critiques in which students shared their strategies, challenges, and planned revisions based on feedback from their peers and teachers.

Additionally, Quizizz offered gamified in-class quizzes to reinforce argumentative concepts. Usage logs indicated high engagement across platforms, with most students accessing materials through mobile devices. These tools were selected over alternatives due to their effectiveness, accessibility, and alignment with SRL processes. The researchers selected Google Docs for its collaborative real-time editing and commenting features, supporting goal-setting, drafting, and peer feedback more effectively than static platforms like Microsoft Word or Moodle forums. They favored EdPuzzle over generic video platforms because it integrated formative assessments directly into videos, facilitating active monitoring and retrieval practice—features that passive platforms like YouTube do not inherently support. Padlet was chosen for the reflection phase because its open, visual board format encouraged more interactive and peer-supported reflective practices than linear tools such as blogs or discussion boards. These choices ensured that digital engagement actively supported each SRL phase while remaining accessible for all students on mobile and desktop devices.

3.2.5. Syllabus Text

The course designers adapted the content from Eemeren and Grootendorst's (2003) argumentation model, covering key topics such as the structure of arguments, common logical fallacies, and the evaluation of sound reasoning. The design team selected these components to support students' development of critical thinking and persuasive writing skills. Instruction followed a flipped learning model, where students engaged with pre-recorded video lectures before class and participated in collaborative, application-based activities during in-class sessions (see Table 1).

Table 1
ESP AW Course Topics.

Weeks	Topics	Activities
1	Pre-test	Essay writing
2	Differences of Opinion	Video + discussion
3	Argumentation and Discussion	Video + role-play
4-5	Standpoints and Argumentation	Video + drafting
6	Unexpressed Standpoints and Unexpressed Premises	Video + analysis
7	The Structure of Argumentation	Video + outlining
8-9	Fallacies (1 & 2)	Video + identification tasks
10	The Soundness of Argumentation	Video + peer critique
11-13	Written Argumentation	Video + essay revision
14-15	Oral Argumentation	Video + debate practice
16	Post-test	Essay writing

3.3. Procedure

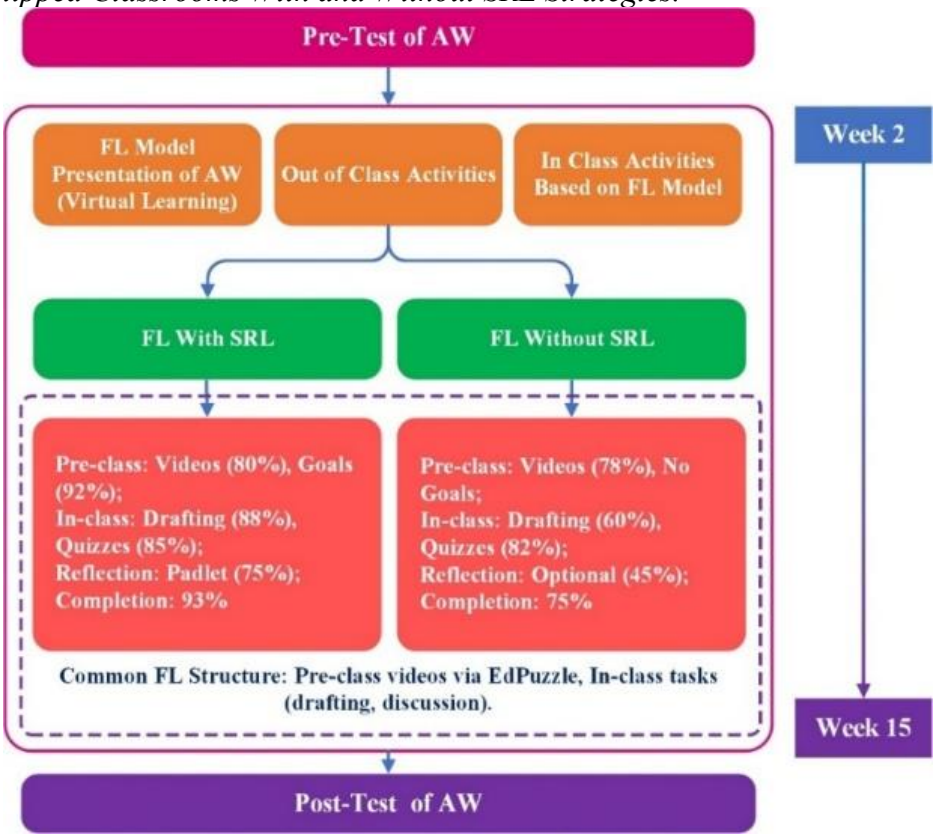
Throughout the 16-week intervention, both groups participated in a flipped ESP writing course. In the experimental group, SRL strategies were specifically integrated:

- **Forethought phase:** Students established weekly goals on Google Docs before drafting sessions. One goal was to ensure each paragraph includes one supporting evidence. Teacher feedback prompted necessary goal refinement.
- **Performance phase:** Students monitored their progress by tracking study times and collaborating on essay drafts in Google Docs. EdPuzzle quizzes before class helped them assess their understanding of key AW components.
- **Reflection phase:** Students posted reflective entries on Padlet discussing their learning strategies and revisions, responding to structured prompts provided by the instructor.

The control group followed the same flipped syllabus without explicit SRL support. Although both groups had access to digital tools, the study systematically guided their usage only in the experimental group. To minimize bias, both groups completed pre- and post-tests independently of their regular coursework. Furthermore, the instructor participated in two SRL-focused professional development workshops before the intervention began, which covered goal-setting, self-monitoring, and reflective facilitation strategies to ensure the consistent delivery of SRL strategies. The study recorded all sessions to monitor instructional fidelity. In addition to video recordings, a trained observer conducted biweekly fidelity checks using a standardized SRL implementation checklist to guarantee the consistent integration of SRL elements across all sessions (see Figure 3).

Figure 3

Flipped Classrooms With and Without SRL Strategies.



3.4. Data Analysis

The researchers conducted all analyses using SPSS 22.0. They performed an ANCOVA using pre-test scores as a covariate to compare AW post-test scores between groups while controlling for initial proficiency differences. They also evaluated improvements in AW scores within groups using paired-sample t-tests. For the Toulmin elements (claims and qualifiers), gains within the experimental group were further analyzed with paired-samples t-tests. The SRQ subscale and total scores were compared using Mann-Whitney U tests, appropriate for ordinal data from Likert scales. Effect sizes (η^2 for ANCOVA, Cohen's d for t-tests, and r for Mann-Whitney U) were calculated to determine the practical significance of the findings. The study applied a two-tailed significance level of 0.05 for all tests. Before conducting ANCOVA, the study assessed the assumption of homogeneity of regression slopes. No significant interaction was found between the covariate (pre-test scores) and the independent variable (group), confirming that the assumption

was met ($p > 0.05$). This validated the use of ANCOVA to compare adjusted post-test scores across groups.

4. Results and Discussion

4.1. Results

4.1.1 Argumentative Writing Outcomes

Pre-test AW scores confirmed baseline equivalence between the experimental ($M = 8.73$, $SD = 2.26$) and control ($M = 8.50$, $SD = 2.30$) groups ($t(238) = 0.79$, $p = 0.42$; Levene's $F = 0.07$, $p = 0.78$). Following the 16-week intervention, both groups demonstrated significant improvement. The experimental group improved from 8.73 to 16.80 ($t = -25.71$, $p < 0.001$, $d = 3.36$), while the control group improved from 8.50 to 14.78 ($t = -15.55$, $p < 0.001$, $d = 2.48$). ANCOVA, controlling for pre-test scores, indicated a significant difference favoring the experimental group ($F(1, 237) = 38.67$, $p < 0.001$, $\eta^2 = 0.14$), with an adjusted mean difference of 1.92 points (95% CI [1.37, 2.65]). According to Cohen's (1988) guidelines for interpreting effect sizes, an η^2 value of 0.01 indicates a small effect, 0.06 a medium effect, and 0.14 a large effect. Therefore, the observed effect size of $\eta^2 = 0.14$ represents a large effect, suggesting that integrating SRL into FL substantially impacted students' AW gains (see Table 2).

Table 2

AW Scores and ANCOVA Results

Group	Pre-test Mean (SD)	Post-test Mean (SD)	Adjusted Post-test Mean	F	p	η^2
Experimental	8.73 (2.26)	16.80 (2.48)	16.75	38.67	<0.001	0.14
Control	8.50 (2.30)	14.78 (2.53)	14.83			

4.1.2 Toulmin Element Improvements

Toulmin's analysis of the AW components revealed greater gains in using qualifiers compared to claims among students in the experimental group. Qualifier scores increased from 1.46 ($SD = 0.50$) to 3.19 ($SD = 0.66$) ($t = -25.32$, $p < 0.001$, $d = 2.92$), while claim scores rose from 1.52 ($SD = 0.50$) to 2.75 ($SD = 0.72$) ($t = -17.21$, $p < 0.001$, $d = 1.98$). However, the analysis indicated that approximately 15% of essays from the experimental group still lacked appropriate qualifiers post-intervention. This suggests that while the SRL-enhanced FL approach significantly improved the overall use of qualifiers, a substantial subset of students encountered challenges in applying nuanced argumentative language. One plausible explanation is that these students had lower underlying English language proficiency, which limited their ability to use hedging or modal expressions accurately and confidently.

Additionally, insufficient prior exposure to academic English conventions—particularly hedging devices and modality structures commonly used in scientific arguments—may have contributed to persistent gaps. It is also possible that the metacognitive gains fostered during the reflection phase were not completely transferred into students' writing behavior, indicating a need for more explicit bridging strategies between the reflection and performance phases. These students may have benefited from additional targeted mini-lessons on modality and nuance or more explicit peer modeling activities during the performance phase. Such interventions could include sentence starters for qualified claims, scaffolded examples of hedging language in context, and formative peer assessment rubrics focusing specifically on qualifiers. Future instructional designs could incorporate scaffolded exercises to support learners requiring extended practice using qualifiers (see Table 3).

Table 3

Toulmin Elements (Experimental Group)

Element	Pre-test Mean (SD)	Post-test Mean (SD)	Gain	t	p	d
Claim	1.52 (0.50)	2.75 (0.72)	1.23	-17.21	<0.001	1.98
Qualifier	1.46 (0.50)	3.19 (0.66)	1.73	-25.32	<0.001	2.92

4.1.3 Self-Regulated Learning Outcomes

The experimental group showed significant improvements in self-reported SRL skills across all dimensions compared to the control group. Total SRQ scores increased from 104.5 to 166.7 ($r = 0.49$, $p < 0.001$). Gains were also evident across subscales: motivation, planning, assessment, and self-directness (see Table 4).

Table 4

SRQ Scores (Experimental vs. Control)

Dimension	Exp Pre (SD)	Exp Post (SD)	Ctrl Pre (SD)	Ctrl Post (SD)	p	r
Motivation	18.2 (4.1)	28.5 (3.8)	17.9 (4.0)	20.3 (4.2)	<0.001	0.45
Planning	20.1 (4.5)	32.4 (4.0)	19.8 (4.3)	22.5 (4.4)	<0.001	0.50
Assessment	48.3 (9.2)	78.6 (8.5)	47.9 (9.0)	54.2 (9.3)	<0.001	0.48
Self-Directness	17.9 (3.9)	27.2 (3.6)	17.6 (3.8)	19.1 (4.0)	<0.001	0.42
Total	104.5 (18.7)	166.7 (17.8)	103.2 (18.5)	116.1 (19.0)	<0.001	0.49

4.1.4 Digital Tool Impact

Digital tools contributed differently to AW improvement. Google Docs supported the highest mean score increase (1.5 points), followed by EdPuzzle (1.2 points), Quizizz (0.9 points), and Padlet (0.8 points). These patterns were statistically significant across comparisons ($p < 0.05$). Among the tools

utilized, Google Docs produced the highest learning gains, attributed to its affordances for real-time collaborative drafting and continuous peer-instructor feedback. This tool supported deeper cognitive engagement by enabling learners to iteratively apply argumentative structures while aligning with SRL's performance and reflection phases. In contrast, EdPuzzle, while effective in enhancing comprehension during the forethought phase, involved more passive learning through video-based input, which limited opportunities for active construction or social negotiation of meaning. These differences underscore that tools promoting active knowledge transformation, rather than mere content reception, yield stronger outcomes in complex skill development such as AW (see Table 5).

Table 5

Digital Tool Impact on AW Scores

Tool	Experimental Mean Increase (SD)	Control Mean Increase (SD)	p-value
Google Docs	1.5 (0.3)	0.8 (0.4)	<0.05
EdPuzzle	1.2 (0.2)	1.0 (0.3)	<0.05
Padlet	0.8 (0.2)	0.3 (0.1)	<0.01
Quizizz	0.9 (0.3)	0.7 (0.2)	<0.05

Figure 4 compares the ANCOVA-adjusted AW post-test scores and improvements in Toulmin elements between groups. The experimental group attained a higher adjusted post-test score (M = 16.75) than the control group (M = 14.83), controlling for pre-test scores. Within the experimental group, Toulmin analysis indicated greater improvements in qualifiers (1.46 to 3.19) than in claims (1.52 to 2.75), suggesting enhanced critical thinking and precision. Error bars reflect variability (±1 SD). Figure 5 illustrates changes in self-reported SRL skills across four dimensions (Motivation, Planning, Assessment, and Self-Directness). The experimental group showed significant gains across all dimensions ($p < .001$), while the control group's improvements were more moderate. Figure 6 presents the effect sizes linked to AW improvements. The between-group ANCOVA effect size was moderate ($\eta^2 = 0.14$), whereas within-group pre-to-post effect sizes were large for both the experimental ($d = 3.36$) and control ($d = 2.48$) groups. Figure 7 indicates higher engagement rates among students in the experimental group using digital tools (e.g., 88% used Google Docs) compared to the control group (e.g., 60%). This greater engagement likely contributed to the observed performance gains, highlighting the importance of structured SRL support.

Figure 4

AW Post-Test Scores and Toulmin Element Improvements.

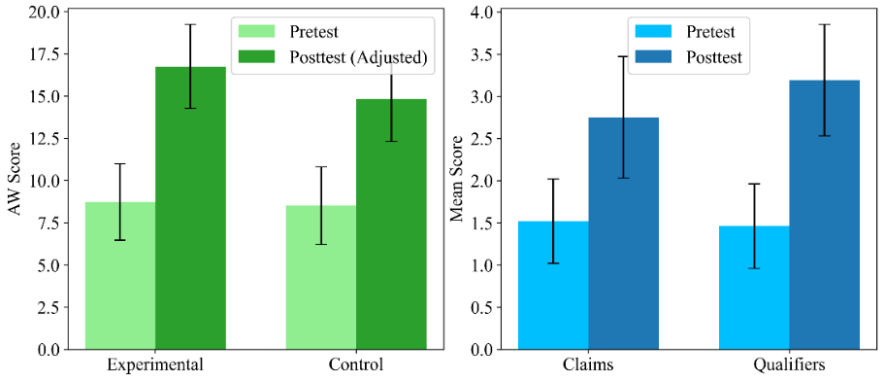


Figure 5

Changes in Self-Reported SRL Skills by Group.

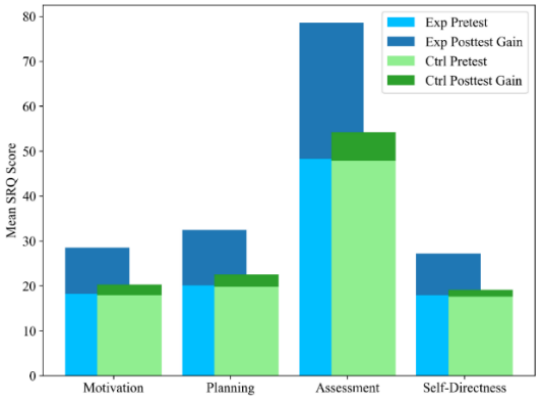


Figure 6

Effect Sizes of AW Improvement Across Groups.

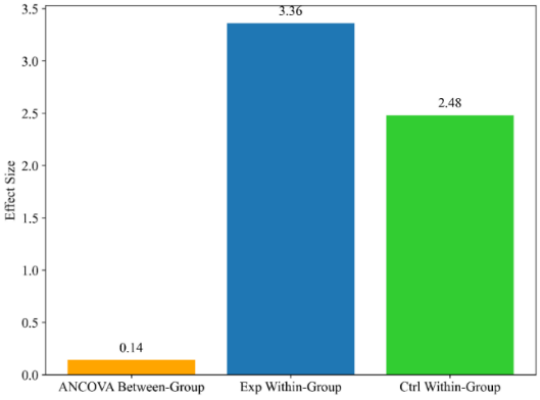
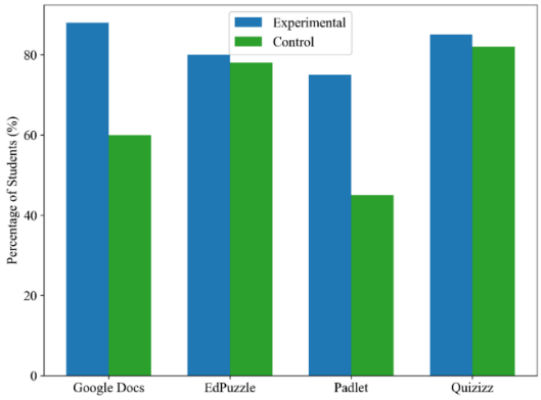


Figure 7

Digital Tool Engagement Rates in Experimental and Control Groups.



4.2. Discussion

The findings of this study suggest that integrating SRL strategies into FL environments significantly enhances students' AW skills compared to FL alone. The large within-group effect size observed for the experimental group ($d = 3.36$) and the moderate between-group effect size ($\eta^2 = 0.14$) reinforce the impact of structured self-regulation strategies on developing complex writing competencies. In particular, the greater improvement in qualifiers compared to claims highlights the role of SRL in fostering nuanced, critical argumentation—an essential skill for professional communication in medical education (Rapanta et al., 2025; Toulmin, 2003). Importantly, the digital tools used in the intervention played a pivotal role, but their effectiveness appeared closely tied to how they were integrated with the SRL framework. Rather than the tools producing gains, their deliberate alignment with SRL phases—goal-setting, performance monitoring, and reflection—maximized their educational impact (Hwang, 2025; Zimmerman, 2008). Google Docs, for instance, supported collaborative goal-setting and iterative drafting, facilitating active engagement during the performance phase. EdPuzzle reinforced knowledge acquisition during the forethought phase through retrieval-based quizzes, while Padlet enabled structured post-task reflections, strengthening metacognitive awareness. This targeted integration ensured that digital environments actively scaffolded students' cognitive and metacognitive processes, consistent with effective SRL practices (Jin, 2024; Turan et al., 2009).

By extending Zimmerman's (2008) SRL theory into a socially mediated digital learning context, this study demonstrates that the regulation of cognition and motivation can be collaboratively constructed through tool-supported peer interaction, rather than being solely internalized individually. It reframes SRL as a dynamic interplay between individual strategy use and

technologically mediated social engagement, thus broadening its theoretical scope for hybrid and online learning environments. Likewise, this study contributes to the development of FL theory by embedding metacognitive strategy training within flipped instruction. While traditional FL models emphasize passive content delivery followed by in-class application, the current design shows how strategically integrating SRL processes through digital prompts, self-assessment tools, and reflective journaling can deepen students' cognitive engagement and transform the FL model from procedural instruction to a fully metacognitive learning cycle. As illustrated in Table 5, the digital tools differentially impacted AW gains. Google Docs produced the highest mean score increase (1.5 points) compared to the control group (0.8 points), likely due to its emphasis on collaborative drafting and immediate feedback cycles that support deeper cognitive engagement. EdPuzzle contributed a 1.2-point increase by scaffolding pre-class understanding of argumentative structures, although its relatively passive, pre-class nature may have limited its impact compared to in-class collaboration (Sanz-Angulo et al., 2025). Padlet played a moderate role (0.8-point gain), fostering reflective practices after writing tasks, although the asynchronous format may have somewhat diluted immediacy and learning transfer (Hwang, 2025). Quizizz, which focused more on content recall and the identification of argument elements, yielded a smaller but significant contribution (0.9-point gain).

The greater impact of Google Docs can be attributed to its capabilities for interactive revision, peer feedback, and sustained engagement during SRL's performance and reflection phases. It enabled learners to iteratively apply feedback, refine their argument structure, and internalize quality criteria for effective writing. In contrast, EdPuzzle's value lies in activating prior knowledge and guiding comprehension during the forethought phase, but its lower interactivity limits engagement with higher-order skills. These findings suggest that tools supporting sustained knowledge construction and collaborative metacognitive dialogue better align with the deeper cognitive demands of AW than tools designed for passive input or factual recall. Thus, this study identifies which tools are effective and clarifies why, based on how tool-specific features interact with SRL phases and cognitive load.

Beyond writing outcomes, the experimental group also demonstrated significant gains in SRL skills, particularly in the planning and assessment dimensions. These improvements reinforce the idea that embedding SRL prompts into digital environments enhances academic skill development and fosters greater learner autonomy (Lai & Hwang, 2016). Students' increased motivation, organizational skills, and self-directed learning behaviors reflect the effectiveness of strategically integrating SRL within flipped classrooms. While the overall results support the efficacy of SRL-enhanced flipped instruction, alternative explanations must also be considered. Although the

groups were statistically equivalent at baseline in AW and SRL skills, individual differences in prior writing experience, intrinsic motivation, or familiarity with digital tools could have influenced the outcomes (Broadbent et al., 2023; Lai & Hwang, 2016). Students with stronger initial self-regulatory tendencies may have benefited more from the SRL scaffolds, amplifying the intervention's effects. Future research should include more detailed pre-assessments of writing experience, motivation, and digital literacy to isolate the instructional effects more precisely (Zimmerman, 2008).

Moreover, the varying degrees of effectiveness among digital tools can be explained by the different cognitive and metacognitive processes each tool supports. Through collaborative editing and goal revision, Google Docs actively engaged students in iterative drafting aligned with performance and reflection phases. EdPuzzle strengthened pre-class conceptual understanding but lacked the collaborative depth needed for complex argument construction. Padlet encouraged reflective thinking; however, being asynchronous, it may not have fully captured immediate critical engagement. Quizizz, while beneficial for reinforcing factual knowledge, primarily supported lower-order skills rather than higher-order argumentation. These patterns reaffirm that tools fostering active construction and social interaction—key elements of SRL—produce the greatest impact on complex academic skills (Turan et al., 2009). The study thus informs instructional design by identifying how the sequencing and affordances of each tool can be mapped onto SRL processes to optimize cognitive engagement and learning outcomes. From a pedagogical perspective, the findings suggest that instructors can effectively enhance AW development by embedding explicit SRL prompts into digital learning activities. Weekly goal-setting through platforms like Google Docs, formative assessments via EdPuzzle, and structured post-task reflections using Padlet offer practical, low-cost strategies for improving ESP writing instruction. By leveraging these freely available digital tools within a structured SRL framework, instructors can promote academic skill development and greater learner autonomy.

5. Conclusion and Implications

This study provides strong evidence that integrating SRL strategies within FL environments significantly enhances students' AW skills and promotes the development of essential self-regulatory behaviors. Medical students who participated in SRL-enhanced FL instruction demonstrated greater improvements in constructing nuanced arguments, using evidence effectively, and applying critical qualifiers compared to their peers in a standard FL setting. In addition, they reported higher levels of motivation, planning, assessment, and self-directed learning. These findings suggest that explicitly guiding students through goal-setting, performance monitoring, and reflective practices fosters deeper cognitive engagement and enhances learner

autonomy, particularly in ESP contexts such as medical education (Turan et al., 2009; Zimmerman, 2008). Theoretically, this study extends Zimmerman's (2008) SRL model into digital learning environments by demonstrating how specific tools—Google Docs for collaborative goal-setting and drafting, EdPuzzle for knowledge monitoring, and Padlet for structured reflection—can scaffold distinct phases of the SRL cycle. Furthermore, the findings contribute to Toulmin's (2003) model of argumentation by illustrating that SRL strategies can significantly improve the precision and complexity of student arguments, addressing the higher-order demands of academic and professional writing more effectively than traditional FL models focused primarily on content delivery.

Importantly, the study advances Zimmerman's model by illustrating that embedding SRL strategies into technology-mediated writing instruction supports individual cognitive and metacognitive regulation, while facilitating social regulation through collaborative platforms. By explicitly aligning digital engagement with SRL phases, the study expands the traditional conception of SRL from a primarily self-driven process to a socially supported, digitally enhanced learning model, especially relevant in contemporary hybrid and online educational contexts. This integration highlights the evolving nature of self-regulation in 21st-century academic environments, where technology, collaboration, and autonomous learning intersect. Practically, this study offers a low-cost, scalable instructional framework for ESP instructors aiming to enhance students' writing and self-regulation skills. Instructors are encouraged to incorporate weekly goal-setting activities within collaborative platforms like Google Docs to help students articulate explicit, task-specific objectives before writing. During the performance phase, tools like EdPuzzle can deliver interactive pre-class videos embedded with formative quizzes, allowing students to monitor their understanding of argumentative structures before in-class application. For the reflection phase, platforms such as Padlet can prompt students to evaluate their writing process, discuss challenges, and propose improvement strategies, guided by reflective questions such as, "Which feedback helped you strengthen your argument today?" Embedding SRL prompts directly into digital learning activities ensures that students actively regulate their writing development, enhancing academic performance and autonomy (Hwang, 2025; Jin, 2024). These examples illustrate that the SRL-enhanced FL model can be flexibly adapted across disciplines and technological environments, broadening its applicability beyond medical ESP education.

Despite these promising outcomes, several limitations must be acknowledged. The 16-week intervention period may have been too short to fully capture the long-term retention of writing and self-regulation skills, as the study observed minor skill decay in the limited post-test follow-up. Future

research should extend the intervention duration to 12–18 months to better assess the stability of gains over time. Studies should also expand the analysis to include additional Toulmin elements, offering a more comprehensive picture of AW development. Furthermore, applying the FL-SRL model across a broader range of ESP fields, such as engineering, law, and public health, and across varied levels of technological access, would strengthen the model's generalizability. Incorporating qualitative methods, such as student interviews, reflective journals, and focus groups, could also provide richer insights into the motivational and cognitive processes that drive successful engagement in SRL-enhanced FL environments. In conclusion, this study provides valuable theoretical and practical insights into digital-era writing instruction. It shows that a structured alignment of digital tools with SRL principles within an FL framework can significantly enhance higher-order academic writing skills and promote greater learner independence. Beyond the specific context of medical ESP education, these findings highlight the potential of SRL-enhanced FL models to improve complex writing skills across various disciplines where critical thinking, evidence integration, and structured argumentation are essential. Technology-mediated SRL strategies can be adapted to support writing development in fields such as engineering, law, public policy, and teacher education, offering scalable and flexible solutions to enhance academic writing outcomes in both high- and low-resource settings. However, the scalability and transferability of this model must be approached with caution. Thus, while the model offers a flexible framework, its effectiveness should be validated through further research in diverse educational and cultural environments. Beyond classroom practice, the findings have implications for curriculum designers and educational policymakers, who may consider embedding SRL-enhanced FL approaches into broader academic writing curricula, particularly in ESP and other writing-intensive programs. Technology developers can also leverage these insights to build tools that facilitate SRL phases and provide instructors with real-time feedback and progress analytics. Additionally, ensuring equitable implementation requires attention to infrastructural disparities, prompting the need for adaptable models and inclusive design that accommodate diverse technological, cultural, and educational contexts. This study contributes to the growing body of research demonstrating that thoughtfully structured digital learning environments can significantly improve academic writing skills and foster greater learner independence in ESP education when aligned with SRL principles.

References

Aidoo, B., Macdonald, M. A., Vesterinen, V.-M., Pétursdóttir, S., & Gísladóttir, B. (2022). Transforming teaching with ICT using the flipped classroom approach: Dealing with COVID-19 pandemic. *Education Sciences*, *12*(6), 421. https://doi.org/10.3390/educsci12060421

Broadbent, J., Panadero, E., Lodge, J., & Fuller-Tyszkiewicz, M. (2023). The self-regulation for learning online (SRL-O) questionnaire. *Metacognition and Learning*, *18*(1), 135–163. https://doi.org/10.1007/s11409-022-09319-6

Cevikbas, M., & Kaiser, G. (2023). Can flipped classroom pedagogy offer promising perspectives for mathematics education on pandemic-related issues? A systematic literature review. *ZDM–Mathematics Education*, *55*(1), 177–191. https://doi.org/10.1007/s11858-022-01388-w

Chen, X. (2025). *Implementing e-assessment for learning in primary EFL writing: A case study in China.* Springer. https://doi.org/10.1007/978-981-97-9491-1_2

Cui, Y., & Zhang, H. (2025). Can student accurately identify artificial intelligence generated content? An exploration of AIGC credibility from user perspective in education. *Education and Information Technologies*, 1-26. https://doi.org/10.1007/s10639-025-13448-1

Deng, R., Jiang, M., Yu, X., Lu, Y., & Liu, S. (2024). Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies. *Computers & Education, 212,* Article e105224. https://doi.org/10.1016/j.compedu.2024.105224

Eemeren, F. H. V., & Grootendorst, R. (2003). *A systematic theory of argumentation: The pragma-dialectical approach*. Cambridge University Press. https://doi.org/10.1017/CBO9780511616389

Fakhri, N., Amini, M., Moosavi, M., Taherifard, E., & Saber, M. (2023). Validity and reliability of the Persian version of modified motivated strategies for learning questionnaire: A methodological study among medical students. *BMC Medical Education*, *23*(1), 553. https://doi.org/10.1186/s12909-023-04547-z

Lacina, J., Roberts, S. K., & Crawford, P. A. (2025). Celebrating pathways to joyful and meaningful writing with young children. *Early Childhood Education Journal*, *53*(1), 1–6. https://doi.org/10.1007/s10643-025-01914-x

Huang, J., & Chen, G. (2022). Individualized feedback to raters in language assessment: Impacts on rater effects. *Assessing Writing*, *52*, 100623. https://doi.org/10.1016/j.asw.2022.100623

Hwang, S. (2025). Unpacking the impact of writing feedback perception on self-regulated writing ability: The role of writing self-efficacy and self-

regulated learning strategies. *Behavioral Sciences*, *15*(2), 100. https://doi.org/10.3390/bs15020100

Jin, S. (2024). Tapping into social media: Transforming EFL learners' writing skills and alleviating anxiety through YouTube. *Education and Information Technologies*, 29(9), 10707–10728. https://doi.org/10.1007/s10639-023-12252-z

Lai, C.-L., & Hwang, G.-J. (2016). A self-regulated flipped classroom approach to improving students' learning performance in a mathematics course. *Computers & Education*, *100*, 126–140. https://doi.org/10.1016/j.compedu.2016.05.006

Lee, H.-J., Lee, E.-H., & Kwon, H.-J. (2024). Online support for international students' engagement in learning: A case study on Padlet usage at a university in Korea. *Social Sciences*, *13*(5), 232. https://doi.org/10.3390/socsci13050232

Li, R., Lund, A., & Nordsteien, A. (2023). The link between flipped and active learning: A scoping review. *Teaching in Higher Education*, *28*(8), 1993–2027. https://doi.org/10.1080/13562517.2021.1943655

Lotfi Gaskaree, B., Fallah, N., & Nili-Ahmadabadi, M. (2025). Alignment of Iranian EAP programs with critical principles: An exploration of educator perspectives. *Journal of Modern Research in English Language Studies*, *12*(2), 173–197. https://doi.org/10.30479/jmrels.2024.20467.2395

Maghsoudi, M., & Mansouri Nejad, A. (2025). Digital flow: Insights from English teacher educators. *Journal of Modern Research in English Language Studies*, *12*(2), 143–171. https://doi.org/10.30479/jmrels.2024.20660.2411

Mekheimer, M. A. (2025). EFL idiom instruction: Teacher practices and challenges. *Social Sciences & Humanities Open*, *11*, 101416. https://doi.org/10.1016/j.ssaho.2025.101416

Ngo, T. N., & Hastie, D. (2025). Artificial intelligence for academic purposes (AIAP): Integrating AI literacy into an EAP module. *English for Specific Purposes*, 77, 20–38. https://doi.org/10.1016/j.esp.2024.09.001

Öztürk, M., & Çakıroğlu, Ü. (2021). Flipped learning design in EFL classrooms: Implementing self-regulated learning strategies to develop language skills. *Smart Learning Environments*, *8*(1), 2. https://doi.org/10.1186/s40561-021-00146-x

Padovano, A., Longo, F., Manca, L., & Grugni, R. (2024). Improving safety management in railway stations through a simulation-based digital twin approach. *Computers & Industrial Engineering*, *187*, Article e109839. https://doi.org/10.1016/j.cie.2023.109839

Rapanta, C., Macagno, F., & Jenset, G. (2025). A close look at children's and adolescents' arguments: Combining a developmental, educational, and philosophical perspective. *European Journal of Psychology of Education*, *40*(1), 3. https://doi.org/10.1007/s10212-024-00914-6

Sanz-Angulo, P., Galindo-Melero, J., De-Diego-Poncela, S., & Martín, Ó. (2025). Promoting soft skills in higher engineering education: Assessment of the impact of a teaching methodology based on flipped learning, cooperative work, and gamification. *Education and Information Technologies*, 1-44. https://doi.org/10.1007/s10639-025-13322-0

Sarwar, M. N., Maqbool, M. A., Ullah, S., Rana, A. S., Khan, S. U., Ibrahim, A. A., Alam, K., Zafar, S., Ullah, Z., & Nazar, M. F. (2024). Fostering conceptual understanding of photocatalysis for sustainable development: A social constructivism flipped-classroom model. *Sustainability*, *16*(23), 10324. https://doi.org/10.3390/su162310324

Teng, M. F., & Wang, C. (2023). Assessing academic writing self-efficacy belief and writing performance in a foreign language context. *Foreign Language Annals*, *56*(1), 144–169. https://doi.org/10.1111/flan.12638

Toulmin, S. E. (2003). *The uses of argument* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511840005

Turan, S., Demirel, Ö., & Sayek, İ. (2009). Metacognitive awareness and self-regulated learning skills of medical students in different medical curricula. *Medical Teacher*, *31*(10), 477–483. https://doi.org/10.3109/01421590903193521

Weng, X., Xia, Q., Ahmad, Z., & Chiu, T. K. (2024). Personality traits for self-regulated learning with generative artificial intelligence: The case of ChatGPT. *Computers and Education: Artificial Intelligence*, 7, 100315. https://doi.org/10.1016/j.caeai.2024.100315

Wong, Z. Y., & Liem, G. A. D. (2022). Student engagement: Current state of the construct, conceptual refinement, and future research directions. *Educational Psychology Review*, *34*(1), 107–138. https://doi.org/10.1007/s10648-021-09628-3

Yang, R. (2022). An empirical study of claims and qualifiers in ESL students' argumentative writing based on the Toulmin model. *Asian-Pacific Journal of Second and Foreign Language Education*, 7(1), 6. https://doi.org/10.1186/s40862-022-00133-w

Yeganeh, L. N., Fenty, N. S., Chen, Y., Simpson, A., & Hatami, M. (2025). The future of education: A multi-layered Metaverse classroom model for immersive and inclusive learning. *Future Internet*, *17*(2), 63. https://doi.org/10.3390/fi17020063

Zare, J., Delavar, K. A., Derakhshan, A., & Pawlak, M. (2024). The relationship between self-regulated learning strategy use and task

engagement. *International Journal of Applied Linguistics*, *34*(3), 842–861. https://doi.org/10.1111/ijal.12535

Zariholhosseini, E., Shafiee, S., & Tabatabaei, O. (2025). EFL teachers' creativity and L2 learners' academic attainment: The moderating role of teaching experience. *Journal of Modern Research in English Language Studies*, *12*(2), 1–22. https://doi.org/10.30479/jmrels.2024.20247.2364

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal*, *45*(1), 166–183. https://doi.org/10.3102/0002831207312909