WILEY

ORIGINAL ARTICLE

Structured or Semi-Structured? The Use of Reflection Journals in Postgraduates' Generative Artificial Intelligence Literacy Development in an L2 Academic Writing Context

Danyang Zhang[1] | Lanyu Wen[1] | Junjie Gavin Wu[2]

[1]College of International Studies, Shenzhen University, Shenzhen, China | [2]Faculty of Applied Sciences, Macao Polytechnic University, Macao SAR, China

Correspondence: Lanyu Wen (2400221005@mails.szu.edu.cn)

ABSTRACT

Generative artificial intelligence (GenAI) not only has the potential to aid L2 academic writing but also poses unique challenges concerning impacts and ethics. Reflection journals, which promote critical thinking and metacognitive awareness, have the capacity to guide GenAI-assisted writing, yet remain underexplored. This study examines and compares the effects of structured reflection journals (SRJs) and semi-structured reflection journals (SSRJs) on enhancing postgraduates' GenAI literacy within an L2 academic writing context. Based on the frameworks proposed by UNESCO and Digital Promise, the study developed a GenAI literacy framework that includes four dimensions: (a) *Operational Competence in GenAI Tools*, (b) *Ethics and Security in GenAI Usage*, (c) *Critical Evaluation of GenAI Outputs and Autonomy* and (d) *Reflection in GenAI Application*. The study assessed the GenAI literacy levels of 39 participants before and after four GenAI-assisted writing tasks. Reflection journals were completed after each task (20 in the SRJ group and 19 in the SSRJ group). Results revealed significant improvements in GenAI literacy across three dimensions (Dimensions 1 to 3), with no significant difference between the SRJ and SSRJ groups. However, thematic analysis of reflective content showed that SRJs, with their more comprehensive guiding questions, encouraged deeper engagement with GenAI-related rules and more thorough evaluations of GenAI-generated content compared to SSRJs. By comparing different types of reflection journals as effective scaffolds in L2 writing pedagogy, this study encourages the integration of reflective practices into GenAI-assisted L2 academic writing classrooms, aiming to enhance students' critical evaluation skills and ethical awareness.

1 | Introduction

In L2 academic writing, the inherent complexity of academic writing requires not only linguistic proficiency but also intellectual rigour to address challenges like avoiding plagiarism and understanding intellectual property and data privacy (Ou et al. 2024; Yang et al. 2024). The integration of GenAI tools into academic writing further exacerbates these challenges: learners must now balance GenAI's efficiency gains with ethical accountability (Yang et al. 2024). This underscores the critical role of classroom teaching in cultivating GenAI literacy, which refers to individuals' abilities to 'effectively navigate and integrate AI technologies into their writing tasks' (Warschauer et al. 2023, 2).

Given the critical need to cultivate GenAI literacy in the classroom, reflective practices emerge as a vital pedagogical strategy to address these complex demands. According to Dewey (1997),

reflection is 'active, persistent, and careful consideration of any belief or supposed form of knowledge in the light of the grounds that support it, and the further conclusions to which it tends' (p. 6). Reflection journals, which embody the notion of reflection as a rigorous and thoughtful process, have become a key tool in higher education for fostering students' critical examination of their beliefs, values and experiences (Minott 2008). By engaging in reflective writing, learners are compelled to think deeply about their learning experiences, analyse their strengths and weaknesses and make connections between theory and practice (Brookfield 2017). Therefore, in the GenAI-assisted academic writing context, reflection journals can help learners understand how GenAI can be integrated into their academic writing processes, such as using GenAI for idea generation, research assistance and proofreading. This can also help foster reflection on the ethical and social implications of relying on such technologies.

The fact that reflection journals can either be structured or semi-structured offers learners different levels of guidance and freedom. Structured reflection journals (SRJs) may provide specific prompts and frameworks (Alt and Raichel 2020; Callens and Elen 2015) that direct learners' attention to the application of GenAI in their academic writing. On the other hand, semi-structured journals (SSRJs) allow for more flexibility, enabling learners to explore their unique perspectives and experiences regarding GenAI in academic writing (Alt and Raichel 2020; Callens and Elen 2015). Understanding the differential impacts of these two types of journals could provide valuable insights for educators, particularly in designing more effective instructional strategies to enhance L2 learners' GenAI literacy. This will also help language learners better understand, evaluate and use GenAI-generated content for their academic research and writing.

Against this backdrop, this study aims to explore the effects of SRJs and SSRJs in enhancing postgraduates' GenAI literacy within the context of L2 academic writing. Our findings are expected to contribute to guiding the development of more effective teaching strategies and helping postgraduate L2 learners enhance their GenAI literacy in a way that is both productive and ethical.

2 | Literature Review

2.1 | AI Literacy Frameworks and Practices

2.1.1 | AI Literacy: A Framework to Understand, Evaluate and Use Emerging Technology by Digital Promise

In 2024, Digital Promise, a US-based educational non-profit organisation, released 'AI Literacy: A Framework to Understand, Evaluate, and Use Emerging Technology' (Lee et al. 2024). The framework aims to provide information for educational leaders and instructors from different backgrounds to help them design learning activities that enhance students' AI literacy.

Under this framework, AI literacy refers to the knowledge and skills that enable individuals to critically understand, evaluate and utilise AI systems and tools to participate safely and effectively in an increasingly digital world. It includes four major elements: (a) *AI Literacy Practices*, (b) *Core Values* (i.e., fundamental principles that support students to use AI tools safely and effectively, including being centred on human judgement and justice), (c) *Modes of Engagement* (i.e., using AI tools in an interconnected manner to demonstrate AI literacy) and (d) *Types of Use* (i.e., using AI tools to serve different purposes). The first element, *AI Literacy Practices*, includes: (a) algorithmic thinking, abstraction and decomposition, (b) data analysis and inference, (c) data privacy and security, (d) digital communication and expression, (e) ethics and impact and (f) information and mis/disinformation. These aspects integrate computational thinking, data literacy, digital citizenship and media literacy.

2.1.2 | AI Competency Framework for Students by UNESCO

In 2024, UNESCO proposed an AI competency framework both for teachers and students. The student framework's objective is to foster the development of core AI competencies in students, enabling them to become responsible and creative citizens in the AI era (UNESCO 2024). The framework comprises four aspects: (a) *Human-centred Mindset*, (b) *Ethics of AI*, (c) *AI Techniques and Applications* and (d) *AI System Design*. Each aspect is further categorised into three progression levels: *Understand*, *Apply* and *Create*. Specifically, the *Human-centred Mindset* aspect focuses on students' values, beliefs and critical thinking skills, ensuring they can assess the suitability and ethical implications of AI. *The Ethics of AI* aspect equips students with ethical value judgements and social skills to navigate the AI landscape. The *AI Techniques and Applications* aspect provides students with conceptual knowledge and operational skills related to AI tools. The *AI System Design* aspect hones students' systemic design thinking and comprehensive engineering skills for AI system development (UNESCO 2024).

2.1.3 | Alignments Between the Aforementioned Frameworks

Given that 'Generative artificial intelligence (GenAI) is a highly potent sub-category of artificial intelligence (AI)' (Schryen et al. 2025, 12), the generic AI-literacy frameworks offered by UNESCO and Digital Promise can serve as blueprints for deriving a GenAI-literacy model for L2 academic field. Besides, the two frameworks present detailed and instructive structures for evaluating students' AI literacy, focusing on assessing both the comprehension of diverse AI applications and the ability to apply such knowledge in practical contexts. Therefore, this study selects these two frameworks as the theoretical foundations.

The two AI frameworks share several alignments. First, both the *Types of Use* and *Modes of Engagement* in the Digital Promise Framework and the *AI techniques and applications* and *AI system design* from the UNESCO framework emphasise AI foundations and application skills that should be grasped. Second, the UNESCO framework includes the *Ethics of AI*. Similarly, the *Types of Use* and *Modes of Engagement* from the Digital Promise Framework also mention the responsible and
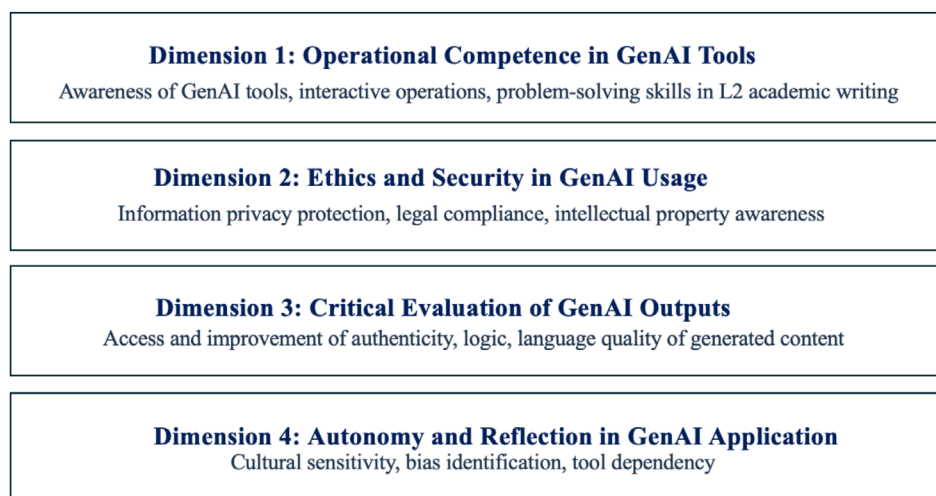
Dimension 1: Operational Competence in GenAI Tools

Awareness of GenAI tools, interactive operations, problem-solving skills in L2 academic writing

Dimension 2: Ethics and Security in GenAI Usage

Information privacy protection, legal compliance, intellectual property awareness

Dimension 3: Critical Evaluation of GenAI Outputs

Access and improvement of authenticity, logic, language quality of generated content

Dimension 4: Autonomy and Reflection in GenAI Application
Cultural sensitivity, bias identification, tool dependency

FIGURE 1 | The GenAI literacy framework for the present study.

safe use of AI. Thirdly, both frameworks are oriented towards human-centredness. The *Human-centred mindset* aspect in the UNESCO framework highlights human agency, accountability and citizenship, whereas the *Core Value* and *Modes of Engagement* from the Digital Promise framework emphasise human judgement and justice. These values underpin human autonomy and reflection in both frameworks. Besides, both frameworks also recognise the need for students to evaluate the use of AI. The content in the progression levels of the UNESCO framework—*Understand*, *Apply* and *Create*—aligns with Digital Promise's interconnected *Modes of Engagement* and *AI Literacy Practices*, touching on actionable practices for critically evaluating AI outputs.

Based on the AI literacy models provided by Digital Promise and UNESCO, their alignments, as well as considering the characteristics of postgraduates' L2 academic writing, this study proposes a framework for students' GenAI literacy that includes four dimensions: (a) *Operational Competence in GenAI Tools*, (b) *Ethics and Security in GenAI Usage*, (c) *Critical Evaluation of GenAI Outputs* and (d) *Autonomy and Reflection in GenAI Application*. This is specifically applicable in the L2 academic writing context. Figure 1 shows the proposed GenAI literacy framework.

2.1.4 | Practices to Promote Students' GenAI Literacy Development

Scholars have attempted to investigate students' GenAI literacy across various education levels. For instance, studies show that incorporating hands-on activities in K-12 education, such as digital storytelling (Ng et al. 2022) and education robots (Williams et al. 2024), can effectively promote AI literacy among students. These approaches engage students in active learning and help them better understand AI concepts. Similarly, Miao et al. (2021) indicated that leveraging AI tools and resources, such as educational games and simulations, makes learning about AI more engaging and interactive at the K-12 level. For senior secondary school students, Kong et al. (2024) showed that project-based learning can improve students' ability to use

AI concepts for problem-solving, enhance their metacognitive strategies, and help them understand AI ethics. Within this new technological environment, Zhang et al. (2025) focused on university foreign language majors and developed a reliable scale to guide students in searching, understanding and evaluating information resources.

Tzirides et al. (2024) study focused on postgraduates, exploring the impact of using GenAI tools and cyber-social teaching methods on their GenAI literacy. The study found that students felt more comfortable using GenAI tools and had a better understanding and evaluation of GenAI Applications in education after completing the courses on advanced digital technologies for education. The research also revealed students' trial-and-error experiences with GenAI image generation and their understanding of GenAI's advantages and disadvantages compared to human evaluations.

The design of GenAI literacy curricula should be age-appropriate and relevant to students' interests and needs. Su and Yang (2023) proposed a framework for AI education in early childhood education, emphasising the importance of play-based learning and integrating AI concepts into existing curricula. Besides, recent advancements in AI literacy research emphasise the need for structured frameworks to assess competencies in specific contexts. Liu et al. (2025) developed and validated the Generative AI Literacy in Digital Multimodal Composing Scale (GAIDMCS). This four-dimensional instrument measures affective, behavioural, cognitive and ethical literacy among university EFL learners in China. This scale highlighted the importance of integrating prompt engineering, ethical considerations and multimodal AI tools into AI literacy curricula. Their findings suggested that explicit training in evaluating AI outputs and understanding ethical risks enhances students' ability to critically engage with GenAI in digital composing tasks.

Although advancements in GenAI literacy have been documented in foundational educational contexts, a critical gap remains in the development of classroom interventions and curricula tailored to GenAI integration, particularly within L2

academic writing environments. This deficiency underscores the pressing need for targeted instructional strategies that address the dual challenges faced by postgraduate learners: navigating linguistic complexities inherent in L2 writing while simultaneously incorporating GenAI tools into their compositional processes. Examining how these interventions function will shed light on the opportunities and challenges of integrating GenAI into higher education L2 writing practices and offer actionable insights for pedagogy and curriculum design.

2.2 | L2 Academic Writing With Reflective Practice

Reflective writing practice is supported by a robust theoretical foundation and extensive practical exploration. Theoretically, constructivism posits that learning involves students actively constructing knowledge based on prior experiences (Piaget 1953). Similarly, metacognitive theory emphasises learners' awareness and regulation of their cognitive processes (Flavell 1979), whereas Dewey's (1997) theory of reflective thinking positions reflection as a critical bridge between experience and deep learning. As Allan and Driscoll (2014) noted, reflection journals serve as tools for fostering metacognitive thinking and documenting knowledge construction, offering learners a structured space to articulate their learning processes.

For L2 learners, who face language barriers and cultural differences, reflection journals provide a practical means to address L2 writing challenges. Recent studies, such as those by Li and Hebert (2024) and Zhang and Hyland (2023), highlight the role of reflection journals as scaffolds and catalysts in peer feedback contexts. Meanwhile, Zhang et al. (2022) examined how reflective writing enhances self-regulated writing strategies. Specifically, the design of reflection journals significantly influences L2 acquisition outcomes. SRJs, for instance, strike a balance between autonomy and guided prompts, enabling learners to articulate their thoughts within an organised framework (Choy et al. 2019). SSRJs, in contrast, encourage greater emotional and intellectual engagement by prompting learners to reflect on autonomous decisions, thereby fostering self-regulation and goal-oriented learning (Widad and Abdellah 2024). Unstructured reflection journals (URJs), which impose no fixed format, allow learners to explore experiences independently, often revealing subconscious knowledge challenges (Callens and Elen 2015). This freedom also gives instructors insights into learners' unguarded struggles and interests, enabling timely pedagogical adjustments (Alt and Raichel 2020).

Empirical studies have begun exploring these formats. For example, Alt and Raichel (2020) conducted a mixed-methods study on 97 undergraduate education students. They aimed to investigate how SSRJs and URJs impact students' metacognitive awareness, which encompasses knowledge and regulation of cognition. Their qualitative analysis showed that SSRJs can contribute to learners' deeper, long-term reflection, whereas URJs yielded mainly descriptive responses. Quantitatively, the SSRJ group scored higher in metacognitive regulation and knowledge of the post-test, indicating that this type of journal is more effective for students' perceptions of how they regulate their cognition. The study by Callens and Elen (2015) explored the impact of different reflection approaches and methodologies on pre-service teachers' critical reflection. The researchers investigated SRJ and URJ approaches, as well as digital storytelling and learning journal approaches. In total, 164 student teachers participated in two experiments. Results suggested that SRJs are most effective in supporting critical reflection, whereas URJs are less effective. The study also revealed that learner characteristics, like willingness to reflect, can be important factors. However, these studies have overlooked postgraduate students, a population facing unique pressures such as L2 barriers and high-stakes academic demands (Bitchener and Basturkmen 2006).

Although reflection has long been acknowledged as beneficial for L2 learning, we have a limited understanding of how different types of reflection journals, particularly SRJs and SSRJs, impact the GenAI literacy development of postgraduate L2 learners. Postgraduates are in a unique position, being at the crossroads of advanced degrees. Compared with undergraduates, postgraduates' L2 theses are expected to be publishable research contributions, thereby heightening both the ethical stakes and the methodological rigour when using GenAI-assisted academic writing; unlike doctoral students, they still need to consolidate advanced academic writing skills. Moreover, investigating the effects of instructional interventions, such as different types of reflection journals, on EFL learners' GenAI literacy development within an L2 academic writing classroom remains underexplored.

In response to these gaps, the present study focuses on the role of SRJs and SSRJs in enhancing postgraduates' GenAI literacy in an L2 academic writing context. The study has two main objectives. First, it aims to establish a theoretical framework for GenAI literacy in L2 academic writing, combining the various aspects of GenAI literacy that have been previously studied in isolation. Second, it endeavours to provide practical guidelines for further reflective practice. These guidelines are intended to help educators create more targeted GenAI-integrated teaching modules and for postgraduates to optimise their GenAI-empowered L2 academic writing and reflection processes.

The two overarching research questions and the sub-questions of the present study are:

1. Does an eight-week L2 academic writing reflective practice via either SRJs or SSRJs significantly improve postgraduates' GenAI literacy? If so, is there any statistically significant difference between the SRJ and SSRJ groups?

2. How do postgraduates from the SRJ and SSRJ groups reflect on their GenAI-assisted L2 writing practice? What are the differences in reflections between the two groups?

3 | Methodology

3.1 | Participants

The study included 39 participants. They were all first- or second-year postgraduate students of a university in southern China. Of

TABLE 1 | Independent sample *t*-tests of the two groups' pre-test.

	Group (mean ± SD)			
	The SRJ group (*n* = 20)	The SSRJ group (*n* = 19)	*t*	*p*
Pre-test_Operational Competence in GenAI Tools	3.46 ± 0.75	3.04 ± 0.86	1.618	0.114
Pre-test_Ethics and Security in GenAI Usage	3.09 ± 0.68	2.91 ± 0.93	0.675	0.504
Pre-test_Critical Evaluation of GenAI Outputs	3.74 ± 0.62	4.02 ± 0.47	−1.585	0.121
Pre-test_Autonomy and Reflection in GenAI Application	3.66 ± 0.57	3.86 ± 0.74	−0.946	0.350
Pre-test_Overall Score	112.20 ± 15.88	111.84 ± 15.61	0.071	0.944

these, 33 were from mainland China, whereas the rest were from other countries such as Pakistan and Russia. There were 31 female participants and eight male participants. Of the participants, 23 were learning English as a foreign language, six were studying Japanese, six were studying Spanish, and four were studying French. Based on the postgraduate entrance examinations, the participants had similar English language proficiency (around IELTS = 6.5). All participants had prior experience of using GenAI tools for L2 writing, indicating a certain level of proficiency in their GenAI use. Detailed anonymised participant information can be found in Appendix S1. Participants are identified by numbers (e.g., Participant 1) in the following sections.

3.2 | Research Procedure

In this study, a GenAI literacy pre-test was conducted, and participants were categorised into low, medium and high GenAI literacy groups based on their mean scores. Participants from each level (low, medium and high GenAI literacy) were randomly assigned to either the SRJ or SSRJ group, ensuring that both groups were comparable in initial proficiency and each included participants from all three levels. After group allocation, we utilised independent sample *t*-tests to examine the differences in pre-test scores between the groups across the overall and four literacy dimensions.

As shown in Table 1, the *p*-values for the pre-test scores of the four dimensions were all greater than 0.05, demonstrating no significant differences between the two groups at the pre-test stage and confirming the appropriateness of group allocation.

During the eight-week writing course, participants were required to complete four tasks assisted by GenAI platforms. They chose which platform to use themselves. The first task focused on writing an abstract and introduction, the second on writing a literature review and research methodology, the third on writing results, discussion and conclusion sections and the fourth on references and ethics. They used GenAI platforms to help them in generating, evaluating and revising the content. Then, the SRJs and SSRJs were distributed to the participants of the two groups, allowing them to record their thoughts and experiences during GenAI-assisted L2 academic writing. After the eight-week course, a post-test was conducted on all participants to measure the impact on their GenAI literacy. Figure 2 depicts the research procedure.

3.3 | Data Collection and Analysis

3.3.1 | Pre- and Post-GenAI Literacy Tests

Based on the four core dimensions of the GenAI literacy framework (Figure 1), a questionnaire was designed to assess participants' GenAI literacy levels before and after four times of the L2 writing tasks.

The test first gathered participants' demographic information, language proficiency and prior experiences with GenAI. These items aimed to account for the potential influence of participants' pre-existing differences on the post-test results. This was followed by items relating to the four GenAI literacy dimensions. These items included a validity check question ('This is a test question. Please select 'Agree''). The full test is provided in Appendix S2.

The test originally set 35 items for the Genai literacy test. Based on CR values; three items were eliminated ($p > 0.05$), and 32 items were set for the final literacy test. The results of overall and dimensional Cronbach's α (overall: 0.948; Dimension 1: 0.94; Dimension 2: 0.938; Dimension 3: 0.924; Dimension 4: 0.845), KMO value (0.841) and Bartlett's test ($p < 0.01$) indicate high appropriateness of Exploratory Factor Analysis (EFA). Although this study did not adhere to all best practices in scale development (e.g., insufficient sample size for Confirmatory Factor Analysis), the rigorous item screening process (e.g., Cronbach's alpha > 0.3) and theoretically coherent factor solution (67.41% variance explained) support the provisional validity of the measure.

Regarding data analysis, after checking the assumptions, two-way mixed MANOVA (and ANOVA) tests were conducted using SPSS to analyse the effects of Time and groups on overall and dimensional GenAI literacy scores.

3.3.2 | Reflection Journals

This study collected 80 SRJs (comprising a total of 73,962 words) and 76 SSRJs (43,406 words). To ensure that the SRJs and SSRJs comprehensively covered and allowed reflection upon the four literacy dimensions and to help participants enhance their GenAI literacy in L2 writing, this study established priori categories and subcategories for the SRJs and SSRJs. The SRJs, through pre-set questions, explicitly guided students to reflect on the four dimensions of GenAI literacy by addressing specific
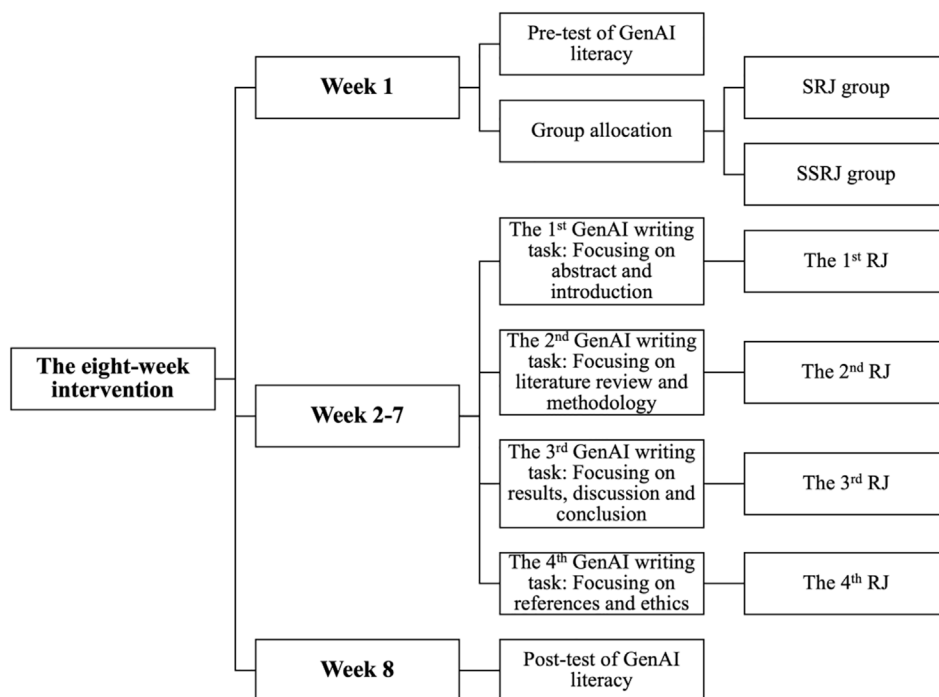
FIGURE 2 | The research procedure.

categories and their corresponding subcategories. In contrast, the SSRJs offered greater flexibility, with only category-level questions, allowing students to freely express their thoughts and experiences within the framework. The full SRJs and SSRJs used in this study are provided in Appendix S3. To explore how or why SRJs and SSRJs may (or may not) impact GenAI literacy, as addressed in Research Question 1, qualitative analysis of the reflection journal content from both groups was conducted.

Braun and Clarke (2006) proposed a six-step thematic analysis approach to identify themes in qualitative data, including familiarising, coding, theme identification, review, definition and reporting. Boyatzis (1998), on the other hand, suggested a deductive approach based on pre-set theoretical frameworks. This study combined both methods, as the SRJs and SSRJs were designed based on the GenAI literacy framework, with certain categories inherently embedded in the reflection journals, making Boyatzis's (1998) deductive approach suitable for initial coding. At the same time, Braun and Clarke's (2006) method was used to further refine categories, identify themes and generate a report. The content of the two groups' journals was then compared. The analysis process is shown in Figure 3.

4 | Results

4.1 | Results of the GenAI Literacy Test

4.1.1 | Postgraduates' GenAI Literacy Development (RQ 1)

Two-way mixed MANOVA (and ANOVA) tests were conducted to examine the main effects of Time (within-subjects factor: pre-test vs. post-test) and Group (between-subjects factor), as well as their interaction effect, on both the overall
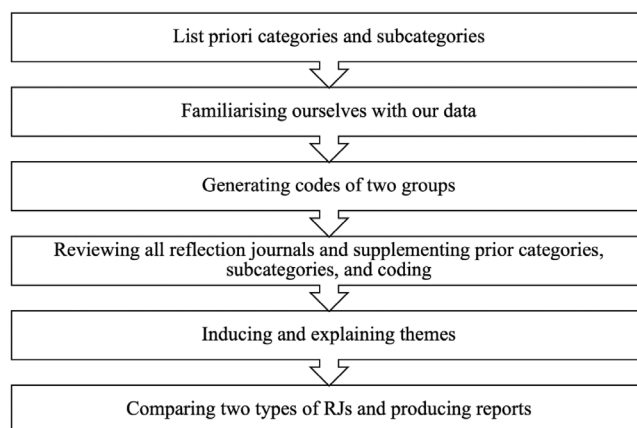
FIGURE 3 | The reflection journal analysis process.

GenAI literacy score and its four dimensional scores. The means and standard deviations for overall and Dimension 1–4 literacy scores are presented in Table 2 below. Before conducting the two-way mixed MANOVA (and ANOVA), the assumptions of normality were satisfied. Normality was assessed using the Kolmogorov–Smirnov test, with most data conforming to normal distribution. The exceptions were data relating to the *Operational Competence in GenAI Tools* and *Critical Evaluation of GenAI Outputs* dimensions, which met approximate normality criteria ($|skewness| < 3$, $|kurtosis| < 10$). These results satisfied the assumptions of the two-way mixed MANOVA (and ANOVA), allowing further analysis of group and time differences in GenAI literacy.

MANOVA revealed a significant main effect of time (Pillai's Trace $= 0.305$, $F(4,71) = 7.8$, $p < 0.001$). However, the main effect of group (Pillai's Trace $= 0.108$, $F(4,71) = 2.145$, $p = 0.084$)

TABLE 2 | Descriptive statistics for participants' GenAI literacy scores in the pre- and post-tests.

	Pre-test				Post-test			
	SRJ group (*n* = 20)		SSRJ group (*n* = 19)		SRJ group (*n* = 20)		SSRJ group (*n* = 19)	
	M	SD	*M*	SD	*M*	SD	*M*	SD
Overall	112.20	15.90	111.84	15.61	128.60	17.28	127.00	15.90
Dimension 1	3.46	0.75	3.04	0.86	4.11	0.61	3.98	0.47
Dimension 2	3.09	0.68	2.91	0.93	3.94	0.53	3.74	1.08
Dimension 3	3.74	0.62	4.02	0.47	4.08	0.61	4.12	0.45
Dimension 4	3.66	0.57	3.86	0.74	3.90	0.70	3.97	0.44

and the time x group interaction (Pillai's Trace = 0.044, F (4,71) = 0.825, p = 0.513) were not significant across four dimensions. To examine specific dimension patterns, subsequent ANOVA tests were conducted for each dimension. The results indicated significant main effects for Time on overall scores (F (1,37) = 31.939, $p < 0.001$, partial η^2 = 0.463), Dimension 1 (F (1,37) = 34.816, $p < 0.001$, partial η^2 = 0.485), Dimension 2 (F (1,37) = 29.249, $p < 0.001$, partial η^2 = 0.442) and Dimension 3 (F (1,37) = 8.402, p = 0.006, partial η^2 = 0.185), but there was no significant main effect for Time on Dimension 4 (F (1,37) = 2.660, p = 0.111, partial η^2 = 0.067). These results suggested that Time significantly influenced the overall and Dimension 1–3 scores, meaning that the reflection journal practice (regardless of group) significantly influenced participants' GenAI literacy scores in terms of overall score and Dimension 1–3. Dimension 4, *Autonomy and Reflection in GenAI Application*, however, referring to long-term technology dependency and cultural sensitivity, might call for longer observations and interventions.

Besides, the results presented no significant main effects for Group on overall score (F (1,37) = 0.05, p = 0.824, partial η^2 = 0.001), Dimension 1 (F (1,37) = 2.52, p = 0.121, partial η^2 = 0.064), Dimension 2 (F (1,37) = 0.786, p = 0.381, partial η^2 = 0.021), Dimension 3 (F (1,37) = 1.059, p = 0.310, partial η^2 = 0.028) and Dimension 4 (F (1,37) = 0.643, p = 0.428, partial η^2 = 0.017); and no significant interaction effects between Time and Group were found for overall scores (F (1,37) = 0.049, p = 0.825, partial η^2 = 0.001), Dimension 1 (F (1,37) = 1.105, p = 0.300, partial η^2 = 0.029), Dimension 2 (F (1,37) = 0.010, p = 0.922, partial η^2 = 0.000), Dimension 3 (F (1,37) = 2.382, p = 0.131, partial η^2 = 0.060) and Dimension 4 (F (1,37) = 0.082, p = 0.231, partial η^2 = 0.010). These revealed that there were no statistically significant differences between the groups, and the effect of Time did not vary significantly depending on the group assignment.

4.2 | Results of Participants' Reflections and Experiences (RQ2)

Qualitative analysis of SRJs and SSRJs explores the impacts of reflective practice in GenAI literacy development and nuanced differences between the two approaches. Following coding, the reflection journals were revisited to ensure no codes,

subcategories or categories were omitted. The priori subcategories and categories were confirmed as final, and themes were inductively derived and explained (Figure 4).

4.2.1 | Dimension 1 Operational Competence in GenAI Tools: Human Perceptions and GenAI Proficiency

The reflection journals primarily revealed two main themes: (1) *perceptions of GenAI tools and basic operation* and (2) *digital communication and expression*. Regarding *perceptions of GenAI tools and basic operation*, participants in both groups generally agreed that GenAI tools are highly accessible and user-friendly. For instance, Participant 10 in the SRJ group mentioned that the GenAI tool she used was easy to obtain and operate. She reported that she could download and install it anytime and anywhere from major app stores. She could easily follow the instructions, which allowed for smooth use, and that the availability of function guides enabled her to understand more complex features. In the SSRJ group, Participant 28 reported that the selected GenAI tool ran smoothly and provided answers in a structured format with clear logic, which were relatively comprehensive. These responses indicate that students praised the accessibility and usability of GenAI tools, with participants recognising their significant advantages in improving the efficiency and quality of their L2 writing.

However, the SRJ group expanded on this theme by mentioning strategies to address challenges encountered when using GenAI for L2 writing, whereas the SSRJ group did not. For example, Participant 4 mentioned that she was unsure which input instructions or prompts would be most effective in generating the content she wanted, so she asked her peers in the same group about their experiences and sought their advice.

This difference suggests that participants in the SRJ group, when faced with the limitations of GenAI tools, were guided by the reflective questions to actively utilise external resources, such as online materials or peer experiences, to resolve issues. In contrast, their SSRJ counterparts did not mention such strategies, possibly indicating a lack of awareness of external resources or insufficient questions to guide deeper reflection.

Regarding *digital communication and expression*, differences emerged between the two groups on this theme. Specifically, the
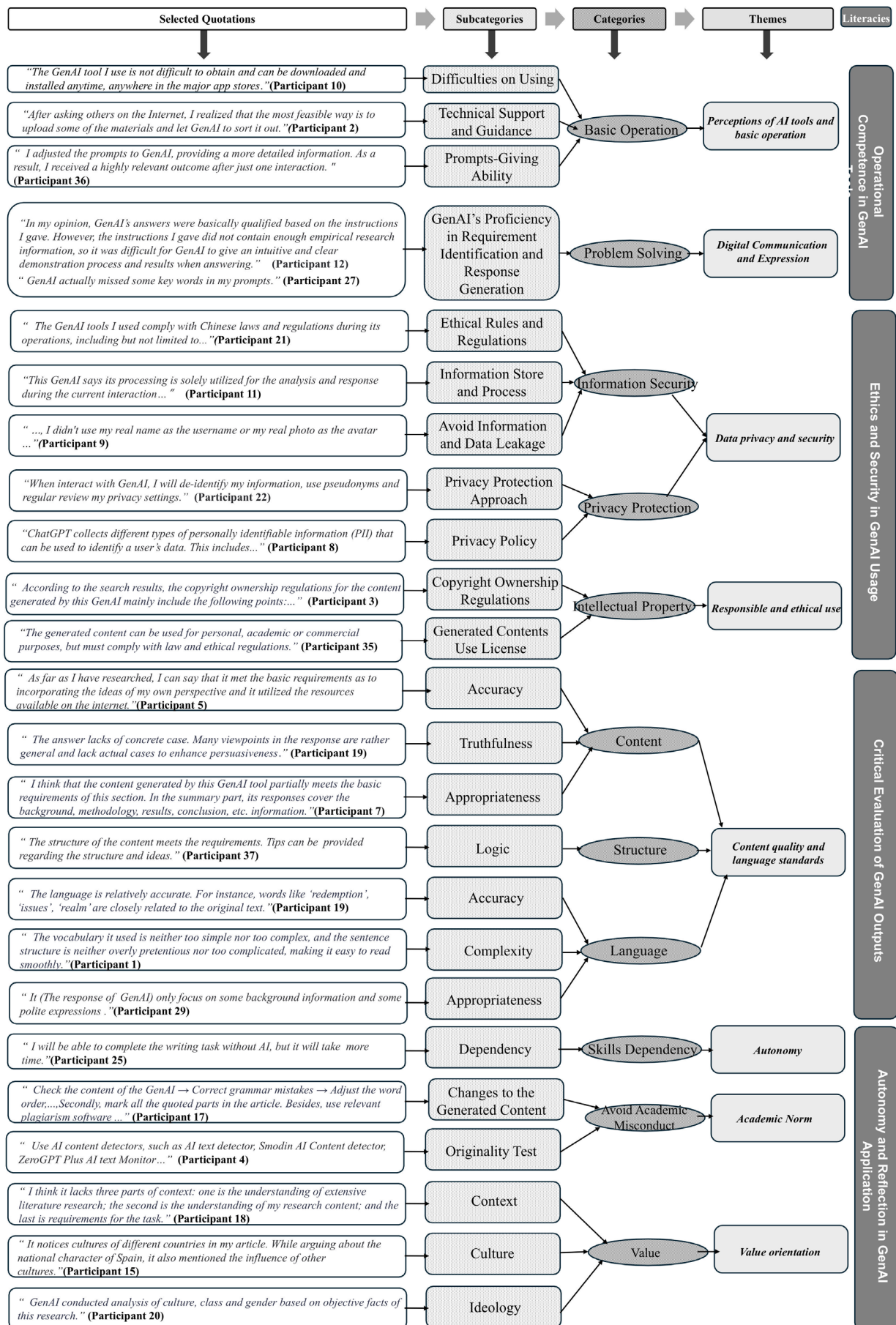
Selected Quotations → **Subcategories** → **Categories** → **Themes** → **Literacies**

"The GenAI tool I use is not difficult to obtain and can be downloaded and installed anytime, anywhere in the major app stores." **(Participant 10)** → Difficulties on Using

"After asking others on the Internet, I realized that the most feasible way is to upload some of the materials and let GenAI to sort it out." **(Participant 2)** → Technical Support and Guidance

" I adjusted the prompts to GenAI, providing a more detailed information. As a result, I received a highly relevant outcome after just one interaction. " **(Participant 36)** → Prompts-Giving Ability

→ Basic Operation → *Perceptions of AI tools and basic operation*

"In my opinion, GenAI's answers were basically qualified based on the instructions I gave. However, the instructions I gave did not contain enough empirical research information, so it was difficult for GenAI to give an intuitive and clear demonstration process and results when answering." **(Participant 12)**
" GenAI actually missed some key words in my prompts." **(Participant 27)** → GenAI's Proficiency in Requirement Identification and Response Generation → Problem Solving → *Digital Communication and Expression*

Operational Competence in GenAI

" The GenAI tools I used comply with Chinese laws and regulations during its operations, including but not limited to..." **(Participant 21)** → Ethical Rules and Regulations

"This GenAI says its processing is solely utilized for the analysis and response during the current interaction..." **(Participant 11)** → Information Store and Process → Information Security

" ..., I didn't use my real name as the username or my real photo as the avatar ..." **(Participant 9)** → Avoid Information and Data Leakage

→ *Data privacy and security*

"When interact with GenAI, I will de-identify my information, use pseudonyms and regular review my privacy settings." **(Participant 22)** → Privacy Protection Approach

"ChatGPT collects different types of personally identifiable information (PII) that can be used to identify a user's data. This includes..." **(Participant 8)** → Privacy Policy → Privacy Protection

" According to the search results, the copyright ownership regulations for the content generated by this GenAI mainly include the following points:..." **(Participant 3)** → Copyright Ownership Regulations → Intellectual Property → *Responsible and ethical use*

"The generated content can be used for personal, academic or commercial purposes, but must comply with law and ethical regulations." **(Participant 35)** → Generated Contents Use License

Ethics and Security in GenAI Usage

" As far as I have researched, I can say that it met the basic requirements as to incorporating the ideas of my own perspective and it utilized the resources available on the internet." **(Participant 5)** → Accuracy

" The answer lacks of concrete case. Many viewpoints in the response are rather general and lack actual cases to enhance persuasiveness." **(Participant 19)** → Truthfulness → Content

" I think that the content generated by this GenAI tool partially meets the basic requirements of this section. In the summary part, its responses cover the background, methodology, results, conclusion, etc. information." **(Participant 7)** → Appropriateness

" The structure of the content meets the requirements. Tips can be provided regarding the structure and ideas." **(Participant 37)** → Logic → Structure → *Content quality and language standards*

" The language is relatively accurate. For instance, words like 'redemption', 'issues', 'realm' are closely related to the original text." **(Participant 19)** → Accuracy

" The vocabulary it used is neither too simple nor too complex, and the sentence structure is neither overly pretentious nor too complicated, making it easy to read smoothly." **(Participant 1)** → Complexity → Language

" It (The response of GenAI) only focus on some background information and some polite expressions ." **(Participant 29)** → Appropriateness

Critical Evaluation of GenAI Outputs

" I will be able to complete the writing task without AI, but it will take more time." **(Participant 25)** → Dependency → Skills Dependency → *Autonomy*

" Check the content of the GenAI → Correct grammar mistakes → Adjust the word order,...,Secondly, mark all the quoted parts in the article. Besides, use relevant plagiarism software ..." **(Participant 17)** → Changes to the Generated Content → Avoid Academic Misconduct → *Academic Norm*

" Use AI content detectors, such as AI text detector, Smodin AI Content detector, ZeroGPT Plus AI text Monitor..." **(Participant 4)** → Originality Test

" I think it lacks three parts of context: one is the understanding of extensive literature research; the second is the understanding of my research content; and the last is requirements for the task." **(Participant 18)** → Context

" It notices cultures of different countries in my article. While arguing about the national character of Spain, it also mentioned the influence of other cultures." **(Participant 15)** → Culture → Value → *Value orientation*

" GenAI conducted analysis of culture, class and gender based on objective facts of this research." **(Participant 20)** → Ideology

Autonomy and Reflection in GenAI Application

FIGURE 4 | Legend on next page.

SRJ group expressed a generally positive view of GenAI in responding to prompts, believing it had the ability to identify and address the needs outlined in the prompts. In contrast, the SSRJ group held a more negative perspective, arguing that GenAI's performance needs improvement.

This discrepancy may stem from the distinct designs of SRJs and SSRJs. With SRJ's detailed questions and guidance, SRJ group users were more likely to systematically optimise their interactions with GenAI tools, leading to clearer, more specific instructions and stronger performance. In contrast, while SSRJs foster open-ended reflection, they lack the scaffolding to guide users in refining their prompts, resulting in less effective GenAI outputs. These findings highlight how the reflection journal format directly influences users' ability to craft high-quality prompts, suggesting that structured frameworks offer a strategic advantage in cultivating the skills needed to maximise GenAI's potential for L2 writing development.

4.2.2 | Dimension 2 Ethics and Security in GenAI Usage: From Rules to Responsibilities

In *Ethics and Security in GenAI Usage*, qualitative analysis of the two groups' journals revealed two main themes: (1) *data privacy and security* and (2) *responsible and ethical use*. Under the first theme, participants from both groups engaged and reflected on aspects of GenAI-generated content. However, participants in the SRJ group conducted a detailed investigation and explanation of the ethical guidelines, data storage and processing methods, and copyright regulations for GenAI tools, which were absent in the SSRJ group. For example, as Participant 9 from the SRJ group stated:

> I adopted several measures to prevent information and data leakage. First, I used a secure and reliable network connection to avoid public Wi-Fi or other potentially risky networks, thereby safeguarding my personal information and data from exposure. Second, when registering, I didn't use my real name as the username or my real photo as the avatar. Also, I didn't provide real information when presented with options regarding more personal details. Finally, I deleted my search history immediately after getting answers to my commands to prevent it from being stolen.

Most of the SRJ participants shared similar detailed actions for protecting the privacy and information, this suggests that participants in the SRJ group, guided by the structured questions, were able to systematically explore the settings of the GenAI tools concerning data security, privacy protection and intellectual property, thereby gaining a deeper understanding of the ethical and legal requirements associated with their use. In contrast, the SSRJ group, lacking such targeted questions, did

not address these aspects, indicating that participants may not have been sufficiently guided to investigate or reflect on the configurations and implications of GenAI tools.

Regarding the second theme, *responsible and ethical use*, this was reflected in both groups' responses concerning taking measures to protect data privacy and information security. However, the SRJ group demonstrated more comprehensive practices in responsible and ethical use, as they reported that they searched and elaborated detailed ethical rules and information store regulations of the GenAI they used:

> I searched its (Skywork AI) settings, and it presented that its operation follows the basic laws, and it does not store the specific information and data I input. When I ask a question, it processes my input in real-time to understand my needs and generate an appropriate response. The processing is only used for the analysis and response of the current interaction. Once the interaction ends, the relevant information will no longer be retained.
>
> (Participant 7 from the SRJ group)

The SRJ group, compared with the SSRJ group, was led by more detailed question prompts regarding the ethical settings of GenAI. Therefore, participants from this group made a greater effort to revisit and examine the privacy policies and related documentation of the GenAI they used. Consequently, their reflection journals contained more comprehensive and detailed responses of this theme.

4.2.3 | Dimension 3 Critical Evaluation of GenAI Outputs: Truth in Detail, Contents in Depth

In *Critical Evaluation of GenAI Outputs*, a main theme emerged from participants' journals: *content quality and language standards*, which reflect the quality of responses from content, structure and language perspectives. Under this theme, both groups acknowledged that the responses of GenAI tools demonstrated basic logical consistency:

> Each GenAI can give relatively complete and logical responses to each segmented part of the content. This is especially true when pointing out the deficiencies in the article, where the answers provided by the GenAI are quite objective.
>
> (Participant 17 from the SRJ group)

> The structure of the content meets the requirements. Tips can be provided regarding the structure and ideas.
>
> (Participant 37 from the SSRJ group)

TABLE 3 | Three key solutions for avoiding academic misconduct in the SRJ group.

Solution	Examples in RJs	Participant no.
Identification of GenAI-generated content	Use various GenAI content detectors, such as AI text detector, Smodin AI content detector and ZeroGPT Plus AI text monitor, to assist in identifying whether GenAI has generated the text	4, 15
Verification of plagiarism on large-scale data platforms	Conduct plagiarism checks on big data platforms like CNKI. These platforms can compare the paper with a vast database to detect replicated parts and calculate the duplication rate	9, 15
Assessment of content originality	Compare the content with the original GenAI-generated text to identify the specific parts where the content has been directly lifted from the GenAI tool	2
	Consult supervisors for a professional review of the paper	9

TABLE 4 | Three key value orientation aspects of GenAI tools arising from the SRJ group.

Key aspect	Examples in RJs	Participant no.
Research context	Cannot provide content with sufficient contextual information (e.g., limited understanding of research background, research goals and expected results)	1, 18
Culture: Partial cultural integration	1. Acknowledges socio-historical contexts (e.g., Western industrialisation, Spanish national identity) but lacks depth 2. Fails to address specific cultural differences (e.g., China vs. Spanish-speaking countries in political discourse)	9, 15, 17
Ideology: Underdeveloped ideological engagement	Outputs focus on 'objective' cultural analysis (e.g., class disparities, gender roles) but neglect political/ideological influences (e.g., government policies)	4, 20

However, the SRJ group further noted that GenAI responses showed good logical coherence on individual questions, but there were issues with logical organisation when summarising ideas. Additionally, the groups differed in their evaluation of the writing style of GenAI tools. The SRJ group highlighted the high quality of GenAI's language expression. For instance, Participant 19 from the SRJ group mentioned that GenAI tools can present ideas in relatively plain and easy-to-understand language. The participant added that while maintaining readability, it also uses some professional terms and more standardised expressions, demonstrating a certain level of professionalism. In contrast, the SSRJ group found that the content generated by GenAI lacks sufficient academic rigour. Participant 29 stated that it mainly focuses on presenting background information and using polite expressions, but the language it generates is frequently inadequate in terms of academic tone.

4.2.4 | Dimension 4 Autonomy and Reflection in GenAI Application: From Academic to Values

In *Autonomy and Reflection in GenAI Application*, three themes emerged: *Autonomy*, *academic norms* and *value orientations*. *Autonomy* focuses on students' reliance on GenAI tools and their ability to independently complete tasks. Participants reported that they could complete writing on their own, ensuring both quality and quantity, and meet deadlines; however, they

acknowledged that using GenAI tools would be of some assistance to them (e.g., Participant 25). Regarding academic norms, the SRJ group demonstrated a more comprehensive reflection, encompassing the identification of GenAI-generated content, verification of plagiarism on large-scale data platforms and assessment of content originality (Table 3). These specific measures were not addressed in the SSRJ group.

For the third theme *Value orientation*, both groups demonstrated concerns. However, the SRJ group, guided by more detailed prompts, offered a more comprehensive evaluation of GenAI-generated content, considering research context, cultural background and ideological alignment (Table 4). This underscores the influence of SRJs in fostering deeper and more nuanced critical analysis of the output of GenAI tools.

5 | Discussion

Through a mixed-methods approach, this study has investigated the impacts of SRJs and SSRJs on enhancing postgraduates' GenAI literacy in an L2 academic writing context. After the eight-week intervention, participants' GenAI literacy improved significantly in overall and dimensions 1–3, but the SRJ and SSRJ groups did not differ statistically. Our qualitative research, however, revealed that the SRJs fostered greater development in some aspects, including methodical

adherence to ethical standards, more critical assessment of GenAI results, and autonomy to consider approaches to solve tool constraints. These findings offer insightful perspectives on GenAI literacy development in L2 academic situations and are consistent with other research on reflective practices (e.g., Alt and Raichel 2020; Brookfield 2017).

Specifically, our quantitative analysis reveals the significant functions and effects of reflection journals in developing participants' GenAI literacy. The results of the two-way mixed MANOVA (and ANOVA) analysis indicated a significant effect of Time on participants' overall and three-dimensional GenAI literacy scores, suggesting that the use of reflection journals had a positive impact on their GenAI literacy. These findings confirm the scaffolding and catalytic roles of reflection journals in the context of writing, as shown by Li and Hebert (2024) and Zhang and Hyland (2023), whereas also highlighting their pedagogical value as metacognitive tools. This supports the claims of Dubiner (2018) and Allan and Driscoll (2014) that reflection helps learners bridge the gap between theory and practice.

Regarding the comparison between SRJs and SSRJs, the results revealed no statistically significant differences in their impacts in enhancing participants' GenAI literacy across the four dimensions. The absence of significant differences between the SRJ and SSRJ groups echoes Callens and Elen's (2015) conclusion that structured and semi-structured formats can yield comparable outcomes when aligned with learning objectives.

Qualitatively, SRJs demonstrated superior efficacy in guiding participants to investigate GenAI-related ethical rules, data security protocols, and strategies to mitigate academic misconduct (e.g., plagiarism checks and originality reviews). This aligns with UNESCO's (2024) emphasis on cultivating the competencies of *Human-centred Mindset* and *Ethics of AI*, which require deliberate scaffolding. For instance, SRJ participants' systematic exploration of GenAI tools' privacy settings and citation practices resonates with the call of Yang et al. (2024) for explicit training in ethical risk mitigation. In contrast, SSRJ participants, whereas acknowledging GenAI's accessibility advantages, seldom delved into ethical or technical configurations, highlighting a potential gap in semi-structured designs.

These findings have several implications for educators. First, reflection journals can be effectively incorporated into L2 academic writing courses to enhance students' GenAI literacy. Educators should be aware of the different strengths of SRJs and SSRJs. SRJs can be used to provide structured guidance, especially when introducing new concepts or when students need to develop a more systematic understanding of GenAI-related issues (Crane and Sosulski 2020). SSRJs, on the other hand, can promote creativity and independent thinking, allowing students to explore their unique experiences and perspectives (Alt and Raichel 2020). SSRJs could be supplemented with optional ethical guidelines or exemplars to mitigate gaps in critical reflection. Second, educators should design reflection journal tasks

that target specific GenAI literacy dimensions based on their students' needs (Dyment and O'Connell 2011). For example, if students struggle with data privacy and security, tasks can be designed to encourage in-depth reflection on these aspects.

For learners, the study highlights the importance of reflective practice in improving GenAI literacy. Students can better understand how to use GenAI tools effectively and responsibly in their academic writing by reflecting. They can also learn from their experiences, identify areas for improvement and develop strategies to address challenges. In this research, we found that students' self-monitoring/self-regulation abilities play a vital role in their performance of reflective practice. According to Ghanizadeh's (2017) study, self-monitoring could have a positive and significant impact on students' critical thinking, as it also significantly and positively influences their reflection. Therefore, for students with weak self-regulation skills, SSRJs may not be appropriate compared with SRJs.

6 | Conclusion

This study firstly proposed a theoretical framework for GenAI literacy among postgraduate students in L2 writing contexts. Then, through an eight-week intervention experiment, it demonstrated that reflection journals (both SRJs and SSRJs) significantly enhanced students' GenAI literacy in L2 academic writing. Although no significant difference was found between the two groups, qualitative analysis revealed distinct characteristics of the two types of journals in students' GenAI literacy development.

Despite the study's strengths, there are also several limitations. First, the sample size ($N = 39$) and homogeneity (predominantly Chinese postgraduates) may limit generalisability. Future studies should consider incorporating more students from different levels and disciplines. Second, the eight-week intervention period precludes conclusions about long-term retention of GenAI literacy. Future studies could adopt longitudinal designs to track sustainability and explore cultural variations in GenAI literacy development. Additionally, combining SRJ and SSRJ elements—such as core ethical prompts with open-ended reflection—might harness the strengths of both approaches, a hypothesis that warrants empirical testing.

Future research can extend the present work by further exploring the influence mechanisms of SRJs and SSRJs in GenAI-assisted L2 academic writing. For instance, adopting a process-oriented methodology, such as capturing micro-level reflective moments, can illuminate exactly how SRJs and SSRJs mediate the transition from noticing a GenAI-literacy gap to enacting a refined writing strategy.

Author Contributions

Danyang Zhang: research design; data collection; writing – original draft; Literature review and analysis. **Lanyu Wen:** research design; formal analysis; writing – original draft. **Junjie Gavin Wu:** writing – final version.

Ethics Statement

All participants voluntarily participated in this research and were well-informed about the detailed research design. They were informed that they could withdraw from the research at any time. Their data were anonymised and were securely protected on the first author's PC. There are no competing interests between the researchers and participants.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

Allan, E. G., and D. L. Driscoll. 2014. "The Three-Fold Benefit of Reflective Writing: Improving Program Assessment, Student Learning, and Faculty Professional Development." *Assessing Writing* 21: 37–55. https://doi.org/10.1016/j.asw.2014.03.001.

Alt, D., and N. Raichel. 2020. "Reflective Journaling and Metacognitive Awareness: Insights From a Longitudinal Study in Higher Education." *Reflective Practice* 21, no. 2: 145–158.

Bitchener, J., and H. Basturkmen. 2006. "Perceptions of the Difficulties of Postgraduate L2 Thesis Students Writing the Discussion Section." *Journal of English for Academic Purposes* 5, no. 1: 4–18.

Boyatzis, R. E. 1998. *Transforming Qualitative Information: Thematic Analysis and Code Development*. Sage.

Braun, V., and V. Clarke. 2006. "Using Thematic Analysis in Psychology." *Qualitative Research in Psychology* 3, no. 2: 77–101. https://doi.org/10.1191/1478088706qp063oa.

Brookfield, S. D. 2017. *Becoming a Critically Reflective Teacher*. John Wiley & Sons.

Callens, J. C., and J. Elen. 2015. "Does a Structured Methodology Support Pre-Service Teachers More to Reflect Critically Than an Unstructured?" *Reflective Practice* 16, no. 5: 609–622.

Choy, S. C., J. S. C. Yim, and D. S. Sedhu. 2019. "Pre-Service Teachers' Reflection on Reflective Practices: A Malaysian Perspective." *Universal Journal of Educational Research* 7, no. 12A: 18–26.

Crane, C., and M. J. Sosulski. 2020. "Staging Transformative Learning Across Collegiate Language Curricula: Student Perceptions of Structured Reflection for Language Learning." *Foreign Language Annals* 53, no. 1: 69–95.

Dewey, J. 1997. *How We Think*. Dover Publications.

Dubiner, D. 2018. "'Write It Down and Then What?': Promoting Pre-Service Teachers' Language Awareness, Metacognitive Development and Pedagogical Skills Through Reflections on Vocabulary Acquisition and Teaching." *Language Awareness* 27, no. 4: 277–294.

Dyment, J. E., and T. S. O'Connell. 2011. "Assessing the Quality of Reflection in Student Journals: A Review of the Research." *Teaching in Higher Education* 16, no. 1: 81–97.

Flavell, J. H. 1979. "Metacognition and Cognitive Monitoring: A New Area of Cognitive–Developmental Inquiry." *American Psychologist* 34, no. 10: 906–911.

Ghanizadeh, A. 2017. "The Interplay Between Reflective Thinking, Critical Thinking, Self-Monitoring, and Academic Achievement in Higher Education." *Higher Education* 74: 101–114.

Kong, S. C., M. Y. W. Cheung, and O. Tsang. 2024. "Developing an Artificial Intelligence Literacy Framework: Evaluation of a Literacy Course for Senior Secondary Students Using a Project-Based Learning Approach." *Computers and Education: Artificial Intelligence* 6: 100214.

Lee, K.-W., K. Mills, P. Ruiz, et al. 2024. "AI Literacy: A Framework to Understand, Evaluate, and Use Emerging Technology." Digital Promise. https://digitalpromise.org/2024/06/18/ai-literacy-a-framework-to-understand-evaluate-and-use-emerging-technology/.

Li, A. W., and M. Hebert. 2024. "Unpacking an Online Peer-Mediated and Self-Reflective Revision Process in Second-Language Persuasive Writing." *Reading and Writing* 37, no. 6: 1545–1573.

Liu, M., L. J. Zhang, and D. Zhang. 2025. "Enhancing student GAI literacy in digital multimodal composing through development and validation of a scale." *Computers in Human Behavior* 166: 108569. https://doi.org/10.1016/j.chb.2025.108569.

Miao, F., W. Holmes, R. Huang, and H. Zhang. 2021. *AI and Education: A Guidance for Policymakers*. UNESCO Publishing.

Minott, M. A. 2008. "Valli's Typology of Reflection and the Analysis of Pre-Service Teachers' Reflective Journals." *Australian Journal of Teacher Education* 33, no. 5: 55–65.

Ng, D. T. K., W. Luo, H. M. Y. Chan, and S. K. W. Chu. 2022. "Using Digital Story Writing as a Pedagogy to Develop AI Literacy Among Primary Students." *Computers and Education: Artificial Intelligence* 3: 100054.

Ou, A. W., B. Khuder, S. Franzetti, and R. Negretti. 2024. "Conceptualising and Cultivating Critical GAI Literacy in Doctoral Academic Writing." *Journal of Second Language Writing* 66: 101156.

Piaget, J. 1953. *To Understand Is to Invent*. Grossman (French: Ou va I'education?, 1948).

Schryen, G., M. Marrone, and J. Yang. 2025. "Exploring the Scope of Generative AI in Literature Review Development." *Electronic Markets* 35: 13.

Su, J., and W. Yang. 2023. "Unlocking the Power of ChatGPT: A Framework for Applying Generative AI in Education." *ECNU Review of Education* 6, no. 3: 355–366.

Tzirides, A. O. O., G. Zapata, N. P. Kastania, et al. 2024. "Combining Human and Artificial Intelligence for Enhanced AI Literacy in Higher Education." *Computers and Education Open* 6: 100184.

UNESCO. 2024. *AI Competency Framework for Students*. United Nations Educational, Scientific and Cultural Organization.

Warschauer, M., W. Tseng, S. Yim, et al. 2023. "The Affordances and Contradictions of AI-Generated Text for Writers of English as a Second or Foreign Language." *Journal of Second Language Writing* 62: Article 101071.

Widad, A., and G. Abdellah. 2024. "Reflecting on Soft Skills Development Sessions: Utilizing Reflective Journaling to Enhance Nursing Students' Soft Skills." *Teaching and Learning in Nursing* 19, no. 2: e344–e349.

Williams, R., S. Ali, R. Alcantara, T. Burghleh, S. Alghowinem, and C. Breazeal. 2024. "Doodlebot: An Educational Robot for Creativity and AI Literacy." In *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, Association for Computing Machinery, 772–780.

Yang, Z., J. G. Wu, and H. Xie. 2024. "Taming Frankenstein's Monster: Ethical Considerations Relating to Generative Artificial Intelligence in Education." *Asia Pacific Journal of Education*: 1–14. https://doi.org/10.1080/02188791.2023.2300137.

Zhang, D., Y. Wu, and J. G. Wu. 2025. "Exploring Language Learners' New Media Literacy: Instrument Development and Validation." *Innovation in Language Learning and Teaching*: 1–18. https://doi.org/10.1080/17501229.2025.2473597.

Zhang, Y., M. Li, Y. Chieh, and S. Han. 2022. "On a Path to Becoming More Self-Regulated: Reflective Journals' Impact on Chinese English as a Foreign Language Students' Self-Regulated Writing Strategy Use." *Frontiers in Psychology* 13: 1042031.

Zhang, Z. V., and K. Hyland. 2023. "Student Engagement With Peer Feedback in L2 Writing: Insights From Reflective Journaling and Revising Practices." *Assessing Writing* 58: 100784.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Appendix S1:** GenAI literacy questionnaires. **Appendix S2:** SRJ and SSRJ. **Appendix S3:** The detailed anonymised participants' information.