

Exploring the Benefits and Limitations of AI Summarization Tools in Teacher Education

Todd Jobbitt

Associate Professor, Hankuk University of Foreign Studies

ABSTRACT: Reading response journals used in combination with GenAI summary apps can help students better understand their own thinking about and approaches to using AI for classroom learning and teaching. The purpose of this research is to raise awareness of AI summarization tools for academic content as well as gain insight into teacher-trainee perspectives on the process, of which there is a notable gap in the research. Graduate school teacher-trainees wrote and compared an individual summary to an AI-generated summary on the same article. Qualitative data was collected in the form of reading response journals and a small set of reflection questions about the AI app used. Results showed that while students appreciated the efficiency and ease of use of AI summary apps, such apps can produce overly-broad and less detailed information. While AI summary apps are a practical tool that can aid students' learning needs, the use of reading journals or reflective journals on AI tools can help students better discriminate about the role of AI in their future learning.

KEYWORDS: *AI summary applications, response journals, teacher trainees, perceptions*

I. INTRODUCTION

The last few decades have brought major advances in artificial intelligence (AI), specifically in language and literacy classrooms. Georjeva (2025) notes that a long-term goal of AI “has always been to design student-centered, interactive, and personalized learning materials that enable students to study independently” (p. 90), and thus expand students' language-based knowledge and skills. To this, Chandel and Lim (2024) add multiliteracies – learning that “includes students' ability to interpret the information embedded in multimodal texts and communicate their ideas across a range of modes, including the L2” (p. 31). With so many changes in technologies at play, a brief overview of the history of AI can help clarify the present drive for AI in learning classrooms and the reasons for this study.

AI as an academic field began roughly in the 1950's, with several periods of growth ('AI summers') and sluggishness ('AI winters') – periods of time where AI failed to meet expectations – through the 1990's (Toosi, et al. 2022). The expansion of data and computing power in the early 2000's and deep learning (DL) development in the 2010's meant that dataset models could better mimic the human brain, which allowed DL models to power generative AI (Gen AI). Gen AI is defined as “an artificial intelligence field that concentrates on generating new and original information by machine learning on massive databases of experience” (Aydin and (Add 'and'; delete astericks) Karaarslan, 2023, p. 119). It is seen as a robust type of intelligence due to past shifts toward more statistics-based algorithms, and “...the convergence of parallel processing, higher memory capacity, and more massive data collection (e.g. big data)” (Toosi et al, 2022; p. 11). Thus, since 2022-2023, generative AI (Gen AI) apps for conversation and text using Large Language Models (LLMs), like ChatGPT, Claude, and Google Gemini, have seen an exponential growth in usage by the general public and especially by students and teachers. (See Aydin and Karaarslan (2023) for a helpful overview of ChatGPT fundamentals).

Although there has been exponential growth in the use of generative AI tools, it is also true that research on how graduate students in particular perceive AI's role in reading comprehension and summarization has been less than robust. This author often trains students in summarizing, as that skill has long been a trial for most Korean graduate students. In Korea, there is a traditionally strong focus from middle-school through high school for students to study intensively for the SAT-like *soonung* test, to the detriment of reading and writing skill development. So, with this background in mind, the thought then arose of how best to boost students' awareness of AI-summary apps and attain practice with them, aside from Chat GPT. Rosvelj et al (2025) noted that “teachers and students also started using ChatGPT immediately after its release” due to its “possibilities” of use or task flexibility, and this author's students were also familiar with ChatGPT, but less so with other AI summary apps (p. 4), which again may seem ironic, as “summarizing texts” is a major use of ChatGPT by students in higher education (HE) (Rosjelj et al, 2025, p. 3). To be fair, it is likely the type of summarizing that

is the challenge for students; such task work might impact students' AI competency or AI literacy as well.

The major aim of this study is to examine how students' perceptions of self-written summaries of academic articles were similar or different from AI-generated summaries on the same article. This study is grounded in Vygotsky's sociocultural theory of learning (SCT), exemplified in its three main parts: social interaction, cultural tools, and mediation in cognitive development (Vygotsky, 1978). Accordingly, SCT views learning as taking place in a social context where a lesser-skilled learner attains higher levels of understanding with the help of a more experienced learner through guided or scaffolded support and mediational means. This type of scaffolding is 'near' to a learner's developing skill and so can be achieved with minimal guidance. Vygotsky called this learning closeness a 'Zone of Proximal Development.'

Within SCT, AI tools can be conceptualized as cultural artifacts that mediate learning experiences (Negi, 2020), and tools like academic texts and technology can be viewed as mediational tools that can assist a learner "...to resolve a problem or achieve the [learning] target" (Rahmatirad, p. 25). Summarization and reflection tasks can be helpful, as learners can modify and clarify their own understanding of the text to improve comprehension. Response journals can aid in this process of making-meaning through active analysis and reflection, whereas AI-generated summaries can add an external layer of mediation with a focus only on content. Students can contrast AI summaries against their own interpretations in order to expand their comprehension of such text, as well as ascertain the benefits and limitations of using AI applications for summary writing. Other studies using similar research methods performed in this study on this topic were not found by this author. In addition to a base curiosity as to which AI apps students would use, the research questions for this study were:

1. How do graduate students perceive the differences between their own reading response journals and AI-generated summaries of the same article?
2. In what ways do students describe the benefits and limitations of using AI applications for summarizing academic readings?

II. LITERATURE REVIEW

Recent developments in AI technology have brought tremendous change to higher education classrooms, and teachers must meet these changes head-on. As Negi (2020) put it, "modern challenges...must reflect in the instruction-learning processes inside as well as outside the classroom" (p. 349). One challenge at present is that educator-training in AI tools presently lags behind that of student-learning (Tan et al, 2024), and yet non-stop advances in AI-enhanced methodology, pedagogy, and tools demand that educators keep up-to-date with such advances (Wang et al, 2023; Jobbitt, 2025). Such progression comes at a cost, as students, including teacher trainees, may not be fully aware of how to fully exploit or use an AI app to their advantage, let alone have the opportunity or wherewithal to formulate their own opinions on how such an app can be of benefit. For example, Nazim and Alzubi (2025) declare that, "existing literature underscores the potential for AI to improve teaching and learning outcomes in higher education... [but also] ...indicates persistent challenges concerning training, integration, and awareness" (p. 4). In the words of Alwaqadani (2025), it is therefore "essential to understand teachers' perceptions to effectively harness [AI's] potential" (n.p.) for effective education – this being one of the main reasons for this present study.

Perceptions on the benefits and limitations of AI tools inevitably arise from the usage of such tools. According to Ravselj et al (2025), there are many advantages to using generative AI like ChatGPT, for example, because it offers significant applications in higher education such as "providing continuous, on-demand support, personalized tutoring, enhanced revision tools, and accessibility aid..." as well as having the capability to generate practice questions, summarize content, and assist in academic writing (p. 4). A report by Attewell (2025) reiterated these results, but also noted that while exposure to AI tools is helpful, the skills needed to successfully operate such generative-AI tools will be needed by students, with students emphasizing "the importance of generative AI-ready skills relevant to their future industries" including "a clear expectation by students/learners for comprehensive generative AI integration across education" (p. 2).

There has been research that has examined how students' perceptions of usefulness or intentions may weight against the actual use of AI or the perceived costs of using AI tools. A recent study by Chan and Zhou (2023) with 405 students examined the relationship between student perceptions and their intention to use generative AI in higher education. The results showed that "while it is important to provide students with basic knowledge about GenAI, such as its definition, limitations, and benefits, this alone is not sufficient to foster their intention to use it" and that steps to promote AI literacy should be enhanced. (n.p.). An online-only study with 669 university students by Stritto, et al (2024) looked at additional factors. Specific aspects were: defining generative AI, awareness and suggested use, current degree of use, interest, guidance, career and education impacts, trust, emotions, policies, concerns, hopes, and other perspectives and issues. The results were illuminating, showing that while participants knew a lot about generative AI, a majority had not actually used AI tools online; that campus-based students used AI more frequently; and, that "most participants were skeptical

about the accuracy, trustworthiness, and reliability of generative AI tools” (p. 7), but acknowledged the necessity of using AI in professional pursuits; that the policies and clarity about AI use online were muddled or not addressed at all; and finally, a range of often conflicting emotions, hope and concerns were expressed about AI, with the majority of these being negative or pessimistic in relationship to future uses of generative AI.

A study by Wang, et al (2021) looked at how teachers in higher education perceived a variety of factors, like self-efficacy (SE), anxiety (AN), perceived usefulness (PU), and perceived ease of use (PEU), among others, to predict teacher intentions to use AI-based applications in their classroom teaching, keeping in mind that teaching with AI technology is still in its nascent stage. **The findings showed that “university teachers’ SE would have positive effects on their perceived ease of use, perceived usefulness and attitude toward AI technologies” (p 124).** In other words, how these teachers felt about their ability to use the technology positively impacted how much they would use it and what they thought about it.

A few recent studies have looked at perceptions on AI by students in higher education (HE). Research from Chan and Hu (2023), for example, looked at 399 undergraduate and post-graduate students’ perceptions of Gen AI technologies in higher education. They found that students liked that AI can: personalize learning support and resources; aid in writing and brainstorming; assist in research, data collection, and analysis; create artwork and facilitate “the production of multi-media”; as well as help with routine repetitive administrative tasks (p. 10). **These findings are similar to other research on GenAI apps in this HE student population.** Last, a study by Tierney, Peasey and Gould (2025) examined the perceptions of HE students on their learning experiences with AI. The authors notes that “the student voice is virtually absent from this discourse” and that “few studies use qualitative approaches to glean student perspective on AI in HE (p. 3) – fewer still, perhaps, with teacher-trainees who are learning in HE – but a main point made by Tierney et al (also by Chan and Hu, 2025) is that can be “difficult to plan and mitigate issues on AI in HE, or support engagement with AI, without first understanding the perspectives of students. This point ties in directly with this author’s research on teacher-trainee perspectives.

Use of Reading Response Journals

The use of reading response journals (RRJs) is a qualitative research (QR) process, and a brief explanation of QR can be elucidating. Comprehensive in its overview and guidelines of qualitative research methodology, Lim (2025) conceptualizes qualitative research as having the capability to offer, “a unique lens through which to explore and interpret the complexities of social phenomena” as a complement to QR (p. 200). There is no technology other than AI that has recently captivated both the private and professional sectors of humanity quite like AI has in the last half-decade in particular. This makes the reasons for QR noted by Lim – necessity (addressing complex social phenomena), importance (generating rich insights and human understanding), relevance (connecting research to real-world issues), and urgency (responding to rapid social change), as very appropriate and relevant to this study. By definition, a RRJ is an informal communication “between two or more people about something one person has read about” (Fulps & Young, 1991).

Recent literature on the use of RRJs in HE for graduate school of education teacher trainees is somewhat limited in depth and is skewed toward a young learner (YL) emphasis. In a study by Gudowski (2021) on emergent readers’ use of talk and RRJs to comprehend text, results showed that both had a positive impact on motivation and reading comprehension. Buus (2005) performed a study on upper-elementary students in a literature-based reading program where RRJs were used to assist in students’ reading comprehension. Results were mixed, with various factors affecting students’ RRJ and reading performance, those mainly being motivation, attitudes toward reading, and effort involved in writing. Extending the YL focus to K-6th-grade teachers, Sage (2010) detailed a 10-session professional development project to guide YL teachers “through the process of implementing an effective literacy instructional practice with their students” (p. 99) via employing “literature-response journals in the classroom over time.” These types of research articles apparently encompass the breadth of RRJs in education at present, aside from this author’s prior work (Jobbitt, 2023) which focused on the use of RRJs related to literature circles assignments.

AI vs Human Summarization

It is well-known in recent scholarship that summary writing is one of the main uses of ChatGPT by students. Summary writing is one way for students to increase reading comprehension (Radmacher & Latosi-Sawin, 1995; Sucita & Hukom, 2022). Summarization can be defined as “the extraction of essential information from one or more information sources (Hahn & Mani, 2000, as cited in Fitria, 2023). Since the arrival of ChatGPT, however, more AI writing assistance applications have been developed, specifically in the PDF-summarizer domain. Despite this, even quite-recent research on AI summary apps dealt a dearth of mentions. For instance, a 2023 survey of generative AI apps by Brizuela & Garrido-Merchan featured text-to-video, text-to-3D, text-to-itinerary, and text-to-code apps, but no separate text-to-summary or PDF-summarizer apps. Further, a study on the uses of AI in education (AIEd) reported by Crompton and Burke (2023) found that the most common use of of AIEd in HE was in assessment and evaluation; within AIEd, for ‘generating texts’ the most common uses were ‘single question’ and ‘full test’. Two likely reasons could be: first, the lengthy review-and-revision time needed for publication; and second, the explosion of AI summary apps mainly occurred in

early-to-mid-2023 (including apps by Adobe Acrobat, ChatGPT Code Interpreter, Claude, Wondershare PDFelement, Humata, QuillBot, etc).

A few recent studies, however, were surprisingly helpful and relevant to this author's research. Fitria (2023), presented a compendium of information on the teaching of manual and automated summarization – sans any student subjects – and defined two types of summarizations (text, automatic), along with steps on how to complete both. Research by Youn et al (2025) in a writing-based communication course had students compare their own summary to AI-generated summaries on the same reading text; however, the reading text type was unclear to this author.

Conversely, there are also some limitations to the implementation or use of AI tools. Nazim and Alzubi (2025) state also that while the merits of AI are recognized, it can also be true that “comprehensive inquiries into the practical efficacy and utilization of AI tools in diverse educational contexts” has a clear deficiency (p. 5). Shen et al (2023) note that some areas of writing, like creative writing and summarization, are undergoing much change due to the introduction of AI technology. Shen et al (2023) argue that a new framework for writing may be necessary due to expository writing's requirement for new information in real-world tasks in which “the unique characteristics of expository writing open up new opportunities for designing AI support” (p. 3). Such quick progression of AI tools can lead to gaps in research on how pre-service and in-service teachers perceive AI's role in reading comprehension and summarization.

III. METHODOLOGY

Research design

The methodology for this small-scale study follows a qualitative design process utilizing reading response journals (RRJs) to gather 11 student responses in the form of observations, feelings, attitudes, and beliefs. There were several steps for students to complete. Students read an academic article of their choice related to the field of reading education (lexis, extensive reading, etc.) and then completed a RRJ, which consisted of four paragraphs:

- A student summary paragraph of an academic article (on reading education)
- An AI-generated summary paragraph of the same academic article
- A reaction paragraph to the academic article (Would it work in the trainees' context?)
- A reflection paragraph on the article (Was it easy or difficult to read? Why/Why not?)

Students first completed one RRJ without an AI summary in order to build familiarity with the assignment format. Then, a second RRJ was completed that included the four paragraphs noted above. A final step was the completion of a question set about using the AI application. These questions are listed in Table 1.

Table 1. Reading Response Journal question set

Table 1. RRJ questions for students

1. What app did you use? Was it your first time using an AI app?
2. How was the experience? (Was it helpful or not? Why/Why not? Would you use it again?)
3. How was the app's summary different from your summary? How was this helpful to you?
4. Would you recommend this app (or another app) to a classmate or not? Why/why not?

Students were asked to write at least 50 words per answer. The received responses were either in question-answer or short essay formats.

Participants

The participants in this study were 11 teacher trainees in a Seoul-based graduate school of education program. All students were English education majors aged from early-20's to early-30's. Three students were male; eight were female. Teaching experience ranged from 'none' to full-time academy or public-school instructors with almost an even split between these two. The terms 'teacher trainees', 'students', and 'trainees' are used synonymously.

Data collection

For this study, three separate points of data were collected from students: 1) Reading response journals (human-generated summaries + reactions + reflections); 2) AI-generated summaries (via each student's chosen application); and, 3) written reflections on the process of using the AI summary app. Due to the fact that different articles were read by the trainees, the summaries were not cross-checked for summary response length as different articles were of obviously different page counts, meaning different summary-response lengths by default. Of note is the small sample size for this study. Although small in number, the data gleaned from students is believed to be adequate when considering *information power* (Malterud et al, 2016), which “indicates that the more information the sample holds, relevant for the actual study, the lower [the] amount of participants is needed” (p. 1), other design and methods factors notwithstanding. The study had a narrow aim (student

perceptions of AI-generated versus self-generated summaries) with a specific homogenous sample (graduate students in English Education). Acquired responses were categorized into relevant themes utilizing as much data as possible to increase the sufficiency of the data set.

Findings

This section presents the student RRJ results based on the above criteria. As noted, students chose different articles, so the best-representative summaries are illustrated in Tables 2 and 3. The first research question, (RQ1), was, “How do graduate students perceive the differences between their own reading response journals and AI-generated summaries?” is supported by RRJ Question #3 and #4. Students’ answers revealed relevant several themes, which include ‘accuracy and completeness’, ‘style and voice/perspective’, and depth of engagement’. RQ1 themes are presented in Table 2 for easier inspection.

Table 2. Perceptions of differences between Student and AI Summaries (RQ1)

Table 2. Main perceptive differences: Student vs. AI Summary apps (RQ1)		
Theme 1	Description	Example Student Quote
Accuracy & Completeness	AI captured main ideas but often missed supporting details	[The AI] “quickly highlighted key points...(but) small but important examples were missed” (S1)
Accuracy & Completeness	AI gave broad overviews	“The app allowed me to quickly grasp the structure and main points of the paper, and it simplified some of the more complex parts” (S3)
Accuracy & Completeness	AI glossed over details that students found helpful	[The AI’s] “summary was more descriptive and more factual, focusing mainly on listing the five success factors without much explanation or examples. My own summary included more details from the interviews and case studies” (S7)
Theme 2	Description	Example Student Quote
Style & Voice/Perspective	Students described their journals as personal and reflective; AI was more neutral or formal	“I could feel that the writing was logically organized and well-structured. I make the mistake of making my text too long because I don’t want to miss any detailed points” (S10)
Style & Voice/Perspective	Student responses used simpler language whereas AI used more formal language	“My summary...uses easy and simple sentences to talk about the text. On the other hand, the AI summary uses more formal and professional words. It focuses on big ideas like technology, multiculturalism, and 21st-century skills, and it sounds more like a research paper or a report” (S4)
Style & Voice/Perspective	AI was more objective; student responses were more subjective	“The AI-generated summary was more formal and focused on the overall structure of the research. In contrast, my own summary focused more on the learner’s experience and the interactive nature of the reading” (S3)
Theme 3	Description	Example Student Quote
Depth of Engagement	The AI gave opportunities to engage in deeper processing	“The AI also made good questions about how to improve literacy and how to support multicultural education and communication. When I clicked the questions, it gave more information and helped me think more deeply.” (S5)
Depth of Engagement	The AI and students discerned key areas as being different in importance	“AI wrote more specifically about the part related to strengthening independence, and I wrote more specifically about the part related to types of inquiry.” (S7)

Depth of Engagement	The AI gave opportunities to engage in deeper processing	<i>"I found [the app] especially helpful—not only did it summarize the article, but it also created quiz questions based on the content. That was interesting and helped me check my understanding."</i> (S9)
---------------------	--	---

Table 2 shows three major themes from student responses. Under RQ1, the first theme of 'accuracy and completeness' indicated that while students really liked the ability of the AI summary app to give a clear overview of the journal article, they noted that the app did not include many details that the students themselves would have included in their own summary. This can indicate that students using AI summary apps may not be obtaining a thorough understanding of the article's information.

The second theme under RQ1 of 'style and voice/perspective' shows how formal AI application summaries can be as opposed to student's use of more informal language. Students generally described their summary as being more personal and subjective, whereas the AI app summary tended to be more objective and formal in tone.

Students noted positively that the AI summarization apps did allow for them to process the content on a deeper level, which presented a third major theme for RQ1 of, 'depth of engagement.' AI apps offered additional background information to students in the form of related questions or focused on other areas of article content that showed the student another perspective.

Results from the second research question (RQ2), *"In what ways do students describe the benefits and limitations of using AI applications for summarizing academic readings?"* also showed a variety of themes. Some themes were positive while others were less so. Student responses were mostly brief, but supportive. These results are listed in Table 3 with following explanations.

Table 3. Benefits and limitations of using AI apps for summarizing academic readings (RQ2)

Table 3. Perceived Benefits and Limitation of AI Summarization (RQ2)		
Theme 1	Description	Example Student Quote
Efficiency & Clarity	AI offered speed and concise main points	<i>"It helps save time and organizes ideas clearly."</i> (S1) <i>"It can definitely save us time by providing very accurate information."</i> (S10)
Efficiency & Clarity	AI offered speed and clarity of content	<i>"It can save time and make complex content more accessible. The app allowed me to quickly grasp the structure and main points of the paper, and it simplified some of the more complex parts."</i> (S3)
Theme 2	Description	Example Student Quote
Confidence & Support	Students used AI to check or validate their comprehension	<i>"[My summary] ...includes a little more detail for the method section and the quantitative information. This was helpful because I could check that the overall findings were not mistaken but it also allowed me to realize that I should practice writing more concisely."</i> (S8)
Confidence & Support	Students use of different AI apps promotes AI literacy	<i>"The experience of trying different AI apps is helping me understand which tools suit my learning style best. Without this class, I probably wouldn't have tried using these tools on my own."</i> (S9)
Theme 3	Description	Example Student Quote
Risk of overreliance	Concern that AI summaries discourage active engagement	<i>"I would remind [peers] that it's important to think critically and not rely entirely on the AI. It's best used as a support tool, not a replacement for personal understanding."</i> (S3)

Theme 4	Description	Example Student Quote
Perceived affordances of AI tools		<i>"...these applications can be effective tools when you want to get an overview of the thesis or find specific information." (S2)</i>
Perceived affordances of AI tools		<i>"...the numbered reference to the PDF attached to the answers makes it easier to find the main points they are referring to. So, I would use this app again due to its wide variety of specified tools." (S8)</i>
Perceived affordances of AI tools	The AI app's tools boosted in-text navigation	<i>The app showed the important points with headings for each section. It also had a cool feature where I could click a number next to the summary and go to the page where the sentence came from. (S4)</i>

As seen in Table 3, four major themes arose from RQ2. The first theme of 'efficiency and clarity' was a comment on how fast the AI apps summarized the pdf articles, and how logical the apps organized the data. This is a practical benefit of using the most recent technology that is available to students around the world in HE, but one caution could be that there is a lack of critical perspective from AI apps. As Gen AI use becomes more widespread, it is likely that even narrow AI summary apps will increase in performance and reliability. Until then, it is important to acknowledge that not all tools are perfect, and imperfections can lead to "skepticism and criticism regarding how it (AI) is incorporated in qualitative research tasks such as critical reviews, conceptualizations, and various forms of content analysis" (Miloyan et al., 2019; Moreno & Redondo, 2016 as cited in Christou (2023). Time will tell as to how well and useful AI apps develop.

The second theme of 'confidence and support' is an emergent theme in the research literature in both online and face-to-face courses. In a study by Jin, et al (2023), perceptions on the usefulness of AI applications of 16 online university students were examined. The researchers noted that such learners can have difficulty regulating their own learning. The results showed that AI applications designed to support self-regulated learning (SRL) were of benefit to learners, but that AI apps were "not useful in supporting motivational regulation" (p. 16). To be fair, the study was not on specific AI applications, but only on the students' perceptions of AI apps as being useful.

'Risk of overreliance' was identified as the third theme. Although it is hard to find such research on AI summarization tools specifically, research on the overreliance on AI dialog systems (GenAI) is very common. Zhai et al (2024) reviewed studies on the effects of over-reliance on AI dialog systems (like ChatGPT, Bard, Copilot, etc.) on university students' cognitive abilities. The results y showed that "overreliance on AI dialogue systems can significantly impact decision making, critical and analytical thinking abilities by fostering dependency and potentially diminishing individual judgement skills" (p. 16). In other words, this means a lack of originality, creative thinking, comprehension and originality, among other concerns, may arise from overuse of AI apps. Conversely, a study by Silitubun (2023) on 200 assorted university students looked at how AI apps can influence students' self-confidence and academic performance. Results showed that students can be successful in their academic performance and gain self-confidence when AI apps are used to support their learning aims.

'Perceived affordance of AI tools' was the fourth and final theme pulled from student responses to RQ2. S8 noted above that in addition to the app allowing here to engage in deeper learning (RQ1, theme 3) and perceived usefulness of the functional features or tools of the app, the app used

...was easy to use because it allows you to directly select specific sections from the PDF and then automatically generate certain pieces of information that you should understand about that section without any self-made prompts. Overall, [it] was helpful, especially due to its easy access and detailed tools. (S8)

This implies that students not only valued the AI summarization apps for their ability to generate content, but also for the organizational tools which facilitated their learning of the text through novel features within the app to more deeply engage with the source material. A study by Lee, et al (2024) of 80 Korean university students examined their perceptions of AI-based writing tools (Google Translate, Naver Papago, and Grammarly). The results indicated that these writing tools (apps) might improve English language-learners (ELLs) writing skills. There was no analysis of in-app tools from any of these three writing tools; again, research on specific in-app tools was difficult to find.

Reflection paragraphs

Several key reflection paragraphs are included in short form as a type of supporting evidence for the

above research questions. It was felt best to include authentic data on how students engaged personally with the texts – their difficulty or ease of comprehension when reading – and to include student perspectives on that process. The several extracts shared below show how students described the challenge of engaging with dense academic texts, particularly when encountering new and unfamiliar vocabulary and technical terminology. Student 10 noted:

The article was not that hard for me to read in the beginning, but while I was reading it, I came across many words that I didn't know before, so even though the content was pretty short, it took me a long time to read because I had to look up the vocabulary often. Also, since my reading skills are not very great so that I used AI to help me when I couldn't understand some certain parts. After reading the article carefully once, I tried to focus on what the author was emphasizing and what the main message was. (S10)

Student 10 shows awareness of how difficult the reading was for her, and presents some strategies to gain further comprehension of the text. For instance, S10 used AI when the text proved too difficult for them to understand. Student 5 shared how the learning of new vocabulary was achieved during their reading of the article.

The article had some difficult words like "instantaneous," "sequentially," and "pervasiveness," but I used a dictionary and could understand the main ideas. I have already done literature circles in class, so I know how they work and why they are good. Because of this, the article was easier for me to read. I was surprised that the old literature circle could change by using technology and teaching 21st-century skills. Even though some words were hard, the article was interesting and taught me new things. (S5)

AI translation tools are often used by students as aids to comprehending reading texts, but a danger can be overreliance. A response from Student 3 illustrates the opposite approach – the lack of use of an AI summary application but instead powering through:

The article was informative but not very easy to read. It included a lot of technical terms and detailed explanations, especially about research methods and game design frameworks, which made some parts challenging to follow. However, the overall idea became clearer as I read more — especially the connection between storytelling and extensive reading. It's getting easier to understand how game-based learning can support reading skills, but the academic tone and structure still require careful reading. (S3)

These student accounts illustrate a common theme of 'navigating difficulty through strategies and support mechanisms.' Such efforts highlight a clear interchange between each student's individual effort and their use of external tools in receptive skill development.

A final piece of supporting evidence comes from a reaction paragraph. In their RRJs, students were to 'react' to the article in terms of how the topic could fit (or not) into their Korean teaching context – would it work, or not? Why or why not? One example is shown from Student 3, who read an article about literature circles, and then wrote:

I found the approach refreshing and meaningful. Letting learners help create stories for extensive reading seems like a great way to boost motivation, especially for adults tired of textbook learning. Personalized content clearly made reading more enjoyable and less stressful. In Korea, this method may face challenges due to the test-focused system and time constraints for teachers. Still, in private academies or after-school programs, it could work well. With proper support, this kind of interactive reading could help learners break away from rote learning and engage more deeply. (S3)

S3 identified why the article was enjoyable to her and perhaps beneficial to several teaching contexts, but recognized limitations of this approach within Korean public school's more traditional classroom context.

Table 4 includes the AI summary applications used by the students in this study.

Table 4. AI Applications used by teacher trainees

Table 4. AI Applications used by Teacher Trainees for AI Summary paragraphs			
Student #	App & link	Web definition	Frequency of use
1	Openread	Interactive paper platform	First time
2	ChatPDF	Document assistant	First time
3	Monica	All-in-one AI assistant	First time
4	ChatPDF	Document assistant	First time
5	ChatPDF	Document assistant	Occasional
6	Recall	Knowledge management system	Twice
7	ChatPDF	Document assistant	Occasional
8	Scispace	Research toolkit	First time
9	MyMap	Idea mapping app	First time
10	ChatPDF	Document assistant	First time
11	Notion	Productivity assistant	Near-daily

A distinction needs to be made about AI app classifications and use. While all of the apps in Table 4 are generative, their scope and purpose are narrower than that of LLM-based chatbot apps, which have a broader scope and purpose. Whereas GenAI chatbot apps can perform PDF analysis, a dedicated PDF app is regarded as having better analysis due to its specialized features. A generative AI app like ChatGPT is a broad, general-purpose chatbot better-suited to generating content across many topics. This author surmises that a majority of students in this study likely used ChatPDF due to familiarity with ChatGPT, which is the most popular GenAI app at present.

IV. DISCUSSION

This study explored how graduate teacher-trainees' perceived the usefulness and limitations of AI summary applications as compared against their own self-generated summaries. Overall, the reflections noted that students valued the AI apps for their timeliness in summarizing and for the brevity in transmitting the main ideas. However, there were limitations noted, such as a lack of detail (which can affect evaluation and comprehension); reliability (which may lead to confusion or obfuscation of comprehension); and over-generalization (which may increase risk of misinterpretation). These findings suggest that human-generated summarization generates richer content (in details and examples) and offers a fuller interpretation of academic texts. These findings further represent some distinct challenges for AI summary applications at present, but with ever-increasing large-language model (LLM) processing, productivity and effectiveness of dedicated AI-summarization apps (including ChatGPT-like apps) like those used in this study may change in the near-future (Malec, 2024).

The results of this study are supported partially by other recent scholarship, which also focused on the use of AI in HE. For example, related in topic (GenAI summarization apps like ChatGPT) but different in the research population (30 undergraduate students), Hutapea et al (2024) examined the role of AI applications in supporting students' ability to understand complex academic texts (textbooks, journals, and conference papers). Results illustrated that "AI applications have significant potential to improve students' learning efficiency and comprehension of academic reading materials" (p. 744) "by reducing "the cognitive load on students, allowing them to focus more on critical analysis and reflective thinking about the material being studied (p. 739). Some results from Hutapea et al (2025) parallel themes from this author's findings, mainly: that overreliance on AI technology can result in diminished critical thinking, so such tools should be considered as supportive in use, not a replacement of such; that the use of AI apps can boost students' confidence and aid in comprehension of academic text, as well as save time.

The implications of this study's findings could offer helpful suggestions to teacher trainees or professional teachers who desire to increase their awareness and use of AI tools in an accessible format (like RRJs) that is both subjective and achievable.

The results of this study may prompt teacher trainees or in-service teachers who are hesitant to use AI, or further prompt those who have just begun to do so. Only through the use of such apps can their awareness of AI tools expand. Using RRJs can be an approachable foundation upon which to document experiences with AI learning tools. RRJs are personal and not overly difficult to use with measured practice. Research scholarship on how GenAI is perceived by students (undergraduate or post-graduate) is increasing (across mixed majors and departments), but there is a clear lack of research on post-graduate teacher-trainee perceptions of AI tools in the field of AI education. These research findings hopefully contribute to this area.

There were two main limitations in this study. The first was the very small-sample size of 11 students. This meant that less data was collected, which consequently restricted the generalizability of the findings. Also, the differences in articles read, and AI summary apps used, by students indicate a lack of standardization or comparability of results.

Future studies in select areas could be helpful to cover gaps in this research. For example, the RRJs used were one-off examples, but a long-term study using repeated journaling on student experiences with AI tools could lead to a better understanding of how deeply students' own metacognitive growth develops. A second gap is that trainees used different AI apps, each with distinct features; thus, there was no opportunity to compare trust or strategy use across multiple uses of the same AI tool. Future research could have students use and compare several AI apps one-by-one in order to ascertain which features are of the most benefit to students' academic reading as well as summarization abilities.

V. CONCLUSION

The purpose of this study was to better understand teacher-trainee perspectives on differences between their own reading response journals (consisting of a summary, reaction, and reflection paragraph, plus a small question set on the AI app itself) and AI-generated summaries. It also examined how trainees described the benefits and limitations of such apps.

The findings reveal that while students appreciated the accuracy of the app to capture main points and its efficiency ('speed') of analysis, they were less impressed with its ability to give details and its lack of

nuance in the results given. Secondly, although the AI summaries were well-organized, the language used was more formal, stiff, or objective when compared to students' more casual or rich, subjective language. Finally, students noted that the AI tool was useful in checking their comprehension of the text and helped promote curiosity in the novelty of AI tools, but there was a concern that the AI summaries would discourage more active engagement from learners if used often.

These insights imply the need for greater AI literacy in teacher education programs – for both student learning needs and professional teacher needs – as teacher education in this area is lagging (Najmiddinova, 2025). The use of reflective practices (like reading journals) may be able to assist trainees in evaluating the role of AI in education more critically.

AI tools are being developed at a rapid pace and hold great promise in helping learners personalize their learning experiences. The variety of AI tools available – summary apps and GenAI – and their ease of use, along with well-structured response journal activities can assist teachers in evaluating the benefits and limitations of AI tools in general and boost students' reading and summarization skills. Complemented with increased and measured AI literacy development, great promise may hold precipitous advances in check for current and future generations of students. However, the limitations of AI tools necessitate that those students judge carefully about which tools to apply, when to use such tools, and to not rely on them for the sake of convenience.

REFERENCES

- [1] Alwaqdani, M. (2025). Investigating teachers' perceptions of artificial intelligence tools in education: potential and difficulties. *Education and Information Technologies*, 30(3), 2737-2755.
- [2] Attwell, S. (2025). *Perceptions of AI 2025*. <https://www.jisc.ac.uk/reports/student-perceptions-of-ai-2025> Accessed: 2025-9-10
- [3] Aydın, Ö., & Karaarslan, E. (2023). Is ChatGPT leading generative AI? What is beyond expectations?. *Academic Platform Journal of Engineering and Smart Systems*, 11(3), 118-134.
- [4] Buus, L. M. (2005). Using reading response journals for reading comprehension. *Networks: An Online Journal for Teacher Research*, 8(1), 122-122.
- [5] Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, 20(1), 43.
- [6] Chan, C. K. Y., & Zhou, W. (2023). Deconstructing student perceptions of generative AI (GenAI) through an expectancy value theory (EVT)-based instrument. *arXiv preprint arXiv:2305.01186*.
- [7] Chandel, P., & Lim, F. V. (2024). Generative AI and literacy development in the language classroom: A systematic review of literature. *Ubiquitous Learning*, 18(2), 31.
- [8] Christou, P. A. (2023). A critical perspective over whether and how to acknowledge the use of artificial intelligence (AI) in qualitative studies. *The Qualitative Report*, 28(7), 1981-1991.
- [9] Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: the state of the field. *International journal of educational technology in higher education*, 20(1), 22.
- [10] Fitria, T. N. (2023). Using Manual and Automatic Summarization: What Should Students Consider in Writing a Summary?. *English Language and Education Spectrum*, 3(2), 85-105.
- [11] Fulps, J. S., & Young, T. A. (1991). The What, Why, When and How of Reading Response Journals. *Reading Horizons: A Journal of Literacy and Language Arts*, 32 (2). Retrieved from https://scholarworks.wmich.edu/reading_horizons/vol32/iss2/3
- [12] Georgieva, V. (2025). Embracing the Inevitable: Tips for EFL and ESP Learning with AI. *E-Learning Innovations Journal*, 3(1), 89-115.
- [13] Gozalo-Brizuela, R., & Garrido-Merchán, E. C. (2023). A survey of Generative AI Applications. *arXiv preprint arXiv:2306.02781*.
- [14] Gudowski, Samantha, (2021). "Using student talk and reader response journals with emergent readers" (2021). Theses and Dissertations. 2868. <https://rdw.rowan.edu/etd/2868>
- [15] Hutapea, E., Hutabalian, R., & Hartati, R. (2024). Summarizing AI Application on Student Learning Efficiency in Understanding Academic Reading Materials. *Indonesian Journal of Education and Development Research*, 3(1), 737-745.
- [16] Jin, S. H., Im, K., Yoo, M., Roll, I., & Seo, K. (2023). Supporting students' self-regulated learning in online learning using artificial intelligence applications. *International Journal of Educational Technology in Higher Education*, 20(1), 37.
- [17] Jobbitt, T. (2023). Raising awareness of extensive reading programs and research for South Korean graduate school teacher trainees. *International Journal of Advanced Research*. 11(09), 382-387.
- [18] Jobbitt, T. (2025). A Case Study: Using Project-Based Learning with AI-Enhanced Learning Applications for Lesson Planning. *International Journal on Studies in English Language and Literature*. 13(8), 1-12.

- [19] Lee, Y. J., Davis, R. O., & Lee, S. O. (2024). University students' perceptions of artificial intelligence-based tools for English writing courses. *Online Journal of Communication and Media Technologies*, 14(1), e202412.
- [20] Lim, W. M. (2025). What is qualitative research? An overview and guidelines. *Australasian Marketing Journal*, 33(2), 199-229.
- [21] Malec, M. (2025, August 10). *Large language models: What you need to know in 2025*. HatchworksAI. <https://hatchworks.com/blog/gen-ai/large-language-models-guide/>
- [22] Malterud, K., Siersma, V. D., & Guassora, A. D. (2016). Sample size in qualitative interview studies: guided by information power. *Qualitative health research*, 26(13), 1753-1760.
- [23] Najmiddinova G. (2025) AI-Enhanced Methodological Training for Future Teachers. *Journal of Multidisciplinary Sciences and Innovations*. 1(3), 248-253
- [24] Nazim, M., & Alzubi, A. A. F. (2025). Empowering EFL teachers' perceptions of generative AI-mediated self-professionalism. *PLoS One*, 20(6), e0326735.
- [25] Negi, S. (2020). Constructivist approach of Vygotsky for innovative trends in learning and teaching. *International Journal of Advanced Academic Studies*, 2(1), 349-353.
- [26] Radmacher, S. A., & Latosi-Sawin, E. (1995). Summary writing: A tool to improve student comprehension and writing in psychology. *Teaching of Psychology*, 22(2), 113-115.
- [27] Rahmatirad, M. (2020). A review of socio-cultural theory. *Siasat*, 5(3), 23-31.
- [28] Ravšelj D, Keržič D, Tomaževič N, Umek L, Brezovar N, A. Iahad N, et al. (2025) Higher education students' perceptions of ChatGPT: A global study of early reactions. *PLoS ONE* 20(2): e0315011. <https://doi.org/10.1371/journal.pone.0315011>
- [29] Sage, Lisa L. La Rue, "Beyond just reading the words: Interacting with text using literature response journals" (2010). Graduate Research Papers. 3281. <https://scholarworks.uni.edu/grp/3281>
- [30] Shen, Z., August, T., Siangliulue, P., Lo, K., Bragg, J., Hammerbacher, J., ... & Sontag, D. (2023). Beyond summarization: Designing ai support for real-world expository writing tasks. *arXiv preprint arXiv:2304.02623*.
- [32] Silitubun, E. (2023). The Relationship Between Student Self-confidence, AI Support, and Academic Achievement: A Study in the Psychology of Motivation and Learning. *Bulletin of Counseling and Psychotherapy*, 5(3), 305-313.
- [33] Stritto, M. E. D., Underhill, G. R., & Aguiar, N. R. (2024). Online students' perceptions of generative AI. Oregon State University Ecampus Research Unit. <https://ecampus.oregonstate.edu/research/publications/>
- [34] Sucita, D. P., & Hukom, S. J. (2022). Writing Summary as An Alternative Strategy to Comprehend the Text. *MATAI: International Journal of Language Education*, 3(1), 33-43.
- [35] Tan, X., Cheng, G., & Ling, M. H. (2025). Artificial intelligence in teaching and teacher professional development: A systematic review. *Computers and Education: Artificial Intelligence*, 8, 100355.
- [36] Tierney, A., Peasey, P., & Gould, J. (2025). Student perceptions on the impact of AI on their teaching and learning experiences in higher education. *Research & Practice in Technology Enhanced Learning*, 20.
- [37] Toosi, A., Bottino, A. G., Saboury, B., Siegel, E., & Rahmim, A. (2021). A brief history of AI: how to prevent another winter (a critical review). *PET clinics*, 16(4), 449-469.
- [38] Wang, Y., Liu, C., & Tu, Y. F. (2021). Factors affecting the adoption of AI-based applications in higher education. *Educational technology & society*, 24(3), 116-129.
- [39] Youn, H., Park, N., & Murphy, R. M. (2025). Beyond writing a summary: Utilizing AI tools to enhance writing skills. [Abstract] *Communication Teacher*, 39(1), 49-55.
- [40] Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review. *Smart Learning Environments*, 11(1), 28.