Tampere University

Carlos Tiu

MAKE IT MAKE SENSE

An Exploration of AI-Mediated Individual Sensemaking in Education Through the Lens of Cognitive Mechanisms

Master's Thesis
Faculty of Information Technology and Communication Sciences
Examiner: Thomas Olsson
Second Examiner: Aino Ahtinen
December 2025

ABSTRACT

Carlos Tiu: Make it Make Sense: An Exploration of AI-Mediated Individual Sensemaking in Education Through the Lens of Cognitive Mechanisms
Master's Thesis
Tampere University
Faculty of Information Technology and Communication Sciences
December 2025

Sensemaking is defined as the way individuals interpret, connect, and relate new information around themselves with their existing knowledge and past experiences to build up their understanding of reality. Because it is fundamentally connected to learning, it is an important lens with which to understand a student's studying and analysis process. Zhang and Soergel's model of individual sensemaking was used as a foundation for understanding sensemaking in this study because it outlines sensemaking activities in general through the cognitive mechanisms employed to instantiate, tune, and adjust an individual's knowledge structure.

In light of the rise of AI use by both students and educators for a myriad of purposes, this study aims to observe AI-mediated sensemaking for the specific task of academic reading, a common task required of the average student, especially in higher education. By doing so, the study also seeks to observe whether Zhang and Soergel's model still appropriately captures sensemaking when AI-mediation is involved, or whether there are now changes in the type, frequency, and order of activities that were not apparent before.

A contextual inquiry and interview was conducted on 5 masters students in Tampere University's Human-Technology Interaction program to understand the sensemaking activities they would perform with and without AI mediation, such as the use of LLMs and text overlays such as Semantic Reader. Screen recordings of the tasks, transcripts of think-aloud and interview dialogue, as well as note-taking artifacts of all participants were collected and analyzed. Results showed that while there appeared to be no new mechanisms being used, participants tended to offload certain lower-level mechanisms such as summarization, key item extraction, categorization, and restatement. The users' overall sensemaking loop also appeared to put a greater emphasis on establishing a clear frame of mind to compare and evaluate AI responses with their own understanding of the information being collected.

The work illustrates not only a snapshot of current use cases and capabilities for AI but also the implicit and explicitly stated preferences of the users in higher education towards AI tools. While AI can already be used as a tool that offers more structured feedback and opportunities to extend learning, other factors such as the nature of the task given to participants, time pressure, interest in the papers being read, and prior knowledge of the content also informed the level of effort and focus participants felt was necessary to dedicate to the tasks. Further research can be pursued towards looking at AI-mediated group sensemaking, a common occurrence especially for learning in classroom or lecture-based environments.

Keywords: Artificial Intelligence, Sensemaking, Education, Contextual Inquiry, Interviews, MAXQDA, Large Language Models, Cognitive Mechanisms, Critical Thinking, Human-AI Collaboration, Prompting, Learning, Knowledge Building

The originality of this thesis has been verified using the Turnitin Originality Check service.

1. USE OF AI IN THESIS

I have utilised AI tools in my thesis:

☑ No
☐ Yes

I acknowledge that I am fully responsible for the entire content of my thesis and accept accountability for any violations of ethical standards in publications.

PREFACE

I would firstly like to thank my supervisor, Thomas Olsson, for his guidance, communication, and flexibility with the entire thesis work process, and to my fellow colleagues at Tampere University's Human-Technology Interaction program. It's been an absolute pleasure to learn and grow with you all and I know the future holds great things for us.

This paper is dedicated to the patience and support of my whole family, my aunts and uncles, and everyone who encouraged me to never give up on my dreams. It truly takes a village to raise a child, and I humbly offer my work as the fruit of all our combined effort and perseverance.

Finally, I would also to specifically dedicate this paper to Lena, who looked after me through the worst sickness I have experienced in recent memory, and without whom I would be physically unable to complete this project on time.

Tampere, 17 December 2025

Carlos Tiu

CONTENTS

1. INTRODUCTION

Sensemaking has been a difficult term to pin down due to its wide use in varying disciplines, but all fields generally agree upon its concept as the way for how individuals, groups, organizations, or communities are able understand, interpret, and integrate their experiences into their current knowledge. Approaches towards sensemaking as well as the frameworks for how they are conceptualized have been highly popularized in human-computer interaction (HCI), cognitive systems engineering, knowledge management, communications studies, and library and information studies (Urquhart, 2025).

The primary definitions of each of these fields also give some insight into each of their perspectives: Russell's HCI work focuses on developing a cost structure for sensemaking as “the process of searching for a representation and encoding data in encoding data in that representation to answer task-specific questions” (1993), Snowden and Kurtz created the Cynefin framework which focuses on 5 dimensions of known, knowable, complex, chaos, and disorder to assist with strategic decision-making for organizations (2003), Weick's take on organizational sensemaking revolves around the collaborative aspect of creating mutual understanding, paying attention to an existing social context and its continued creation, as well as power dynamics (2001), whereas Dervin's Sense-Making model looks at an individual creating and applying sense in a given context to address a physical or cognitive gap (1999). Because sensemaking concerns itself at its core with processing and structuring information, it has often had touchpoints in education, with general agreement on the need for more research strategies that were focused on education and did not overly draw on perspectives from other branches of sensemaking that were quite popular at the time, namely organizational sensemaking, institutional theory, and information science (Urquhart, 2025). Recent studies have now focused on a range of problems that are more nuanced to key actors in the educational sphere, such as understanding the sensemaking processes of school and district leaders in Maine as a driver of institutional policy change in the wake of COVID (Yoshizawa, 2025) to evaluating more theoretical frameworks of sensemaking in faculty members to understand how it impacts their roles and priorities in an experiential education setting (LaCroix, 2024). Amidst an increasingly technology-integrated landscape, there has also been investigation into the effect of AI-mediated sensemaking on students and what kinds of interactions and tensions arose from a critical evaluation of their use (Sivola et al., 2025).

Given such differing perspectives on the nature and application of sensemaking, Zhang and Soergel's comprehensive model of individual sensemaking serves as a valuable way to bridge the gap between the sensemaking frameworks of all these major disciplines. They claimed that the sensemaking process can be generalized into an iterative loop of activities, with each prior model simply differing in terms of (1) which activities are involved, (2) which cognitive mechanisms are used to achieve each activity, and (3) the frequency and order with which each activity is performed (2014, 2020).

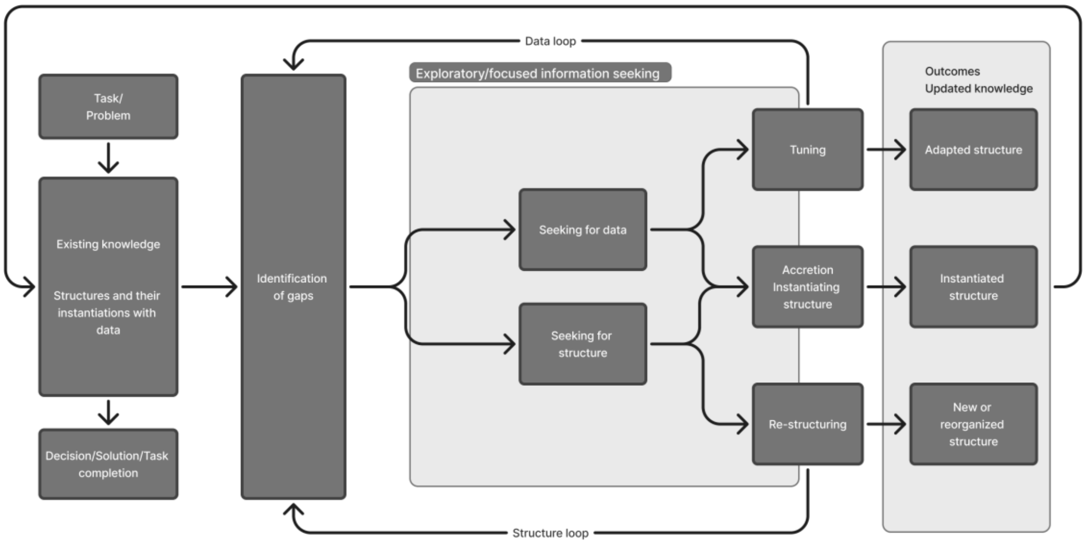


Figure 1. *Adaptation of Zhang and Soergel's Comprehensive Model of Information Searching and Sensemaking, 2020.*

Zhang and Soergel's model generally has six kinds of activities, as well as a seventh set of general control and command activities that may occur throughout the sensemaking process, such as clearly defining a problem or goal, interacting with information, giving and receiving feedback, taking note of where they are in the process, and deciding what they should do next.

The first activity, task analysis and determination of information gaps, is about deciding what the individual wants or needs to know more about, getting an initial impression of the gap between what their current knowledge is and what they want to know, determining if there is an ideal or expected audience that they want to present their results to ahead of time to frame their research process and what information they look for, or trying to imagine what an appropriate solution or answer to their inquiry might look like.

The second activity, information seeking, mainly deals with filling the gaps in information that have been identified. During these activities, the gap tends to manifest in two forms: missing data or missing structure. Missing data means that there the individual needs to acquire new information that they do not have yet, such as the answer to a question they already have in mind, whereas missing structure refers more to the way the individual's existing info pieces of information relate to one another and whether they need to be related to other data they were previously not connected to. This activity also includes the method of searching for the missing information, planning out a search strategy, deciding whether to search in a general or specific manner, looking for sources of information or extracting information from the sources found, and finally, evaluating the information they find.

The third activity in the model is making sense of the information and can involve a set of processes or activities performed when analyzing the data. The sensemaking process is quite iterative since the individual needs to understand the information they are taking in, connect it with their current information or mental structure, and then determine whether the mental model or new data they acquired sufficiently answers the problem they set out to solve, or if there are any inconsistencies or sources of tension in the information they collected. If such tensions exist, the first three activities may be returned to in order to reclarify or adjust the goal, search for new information, or evaluate information. The cognitive mechanisms involved in sensemaking are also primarily used in this activity. One approach is to generate sense through inductive reasoning, such as identifying important content from a source based on how its text is read, while the other is a more deductive, top-down approach, like defining, relating, organizing, and classifying concepts based on a pre-existing structure. After this, the information that has been made sense of is consumed and integrated into the individual's cognitive model, which is the fourth activity, and then usually receives feedback from the application of their new knowledge towards the goal or problem, which is the fifth activity. Lastly, the individual can also reflect on their process, think about how their new knowledge is different from how it was at the beginning of the sensemaking process, and perhaps even relate or expand this to larger themes or to other fields of thought not initially considered.

The reason that there are different sub-activities under each of the activities just mentioned is because different sensemaking theorists outline different steps based on the respective interests of their field, so the main considerations and activities they decided to focus on also differed in turn. This supports the note of the authors in the model that mentions that specific stages or paths were not outlined due to the varying paths and iterations that could take place, as a variety of different tasks and the person-to-person preferences for how they prefer to individually approach learning the subject matter might also affect their process. Therefore, the use of the word *activity* as opposed to *step*, for example, is intentional here – sensemaking is very much a non-linear process that is planned, executed, and monitored in real time as we decide to do it, so we cannot expect to predict what steps might be taken next, but we can perhaps think about what activities might tend to change or be prioritized in the context of certain problems or goals being set by the individual.

While the paths through the sensemaking loop may differ from individual to individual, Zhang and Soergel also found that the cognitive mechanisms employed could be tied to different processes in the loop. This meant that there were relatively predictable sequences of mechanisms used one after another, though which sequence was to happen was largely dependent on a number of extraneous factors, an important one being whether the individual was working with the top-down knowledge structure or the bottom-up manner of bringing together relevant material with new information to form and instantiate a structure. Again, the authors stress that the process is nonlinear and skipping of steps and entering and exiting from various points is to be expected (2020).

Table 1. *Zhang and Soergel's adapted list of cognitive mechanisms tied to key processes in the individual sensemaking loop, 2014.*

Cognitive mechanisms by functionality	Short definition	Data vs. structure driven	Processes supported
Processing new data			
Key item extraction	Identifying important terms/phrases in a text	Data	Acquiring pieces of data to process immediately or keep in mind for later, feeds into many other processes

Restatement	Rephrasing texts into a more accessible or shorter form	Data	Refine the results of key item extraction
Judgment /evaluation	Forming critical opinions toward information being processed	Data	Feeds into many other processes, particularly checking fit
Summarization	Giving a brief statement of the main points	Data	
Generalization	Making claims about groups based on a sample	Data	Tuning, reorganizing, or building structure
Schema induction	Discovering regularities in co-occurrence of phenomena	Data	
Examining concepts			
Definition	Defining different aspects of a concept	Structure	Instantiating structure
Specification	Specifying as conditions or requirements of a problem or task	Structure	Tuning, reorganizing, or building structure
Examining relationships			
Comparison	Comparing a concept/fact to other concepts/facts to recognize similarities and differences	Both	Instantiating structure
Analogy & metaphor	Arguing from similarity in known respects to similarity in other re-	Both	Tuning, reorganizing, or building structure

	spects		
Stereotyping	Using stereotypes (fixed and oversimplified ideas of people or things)	Both	Tuning, reorganizing, or building structure
Classification	Relating a concept to a broader conceptual category	Both	Tuning, reorganizing, or building structure
Explanation-based mechanisms	Examining the causal connections of two phenomena	Both	Tuning, reorganizing, or building structure
Inference	Drawing a conclusion on the basis of circumstantial evidence and prior conclusions	Structure	Tuning, reorganizing, or building structure
Examining anomaly			
Elimination	Eliminating concepts or facts that are not relevant	Structure	Instantiating and tuning structure
Semantic fit	Examining whether a concept appears to fit reasonably into the knowledge structure as a whole	Both	Instantiating, tuning, reorganizing, and building structure
Socratic dialogue	Critical dialogues to facilitate the awareness of inconsistencies	Neither	Instantiating, tuning, reorganizing, and building structure

Given that the actual processing and sense-making activities are often subconscious, the cognitive mechanisms act as a signpost to help observers understand to some degree how the individual's mental schema was being reinforced or constructed. For example, a student recalling a specific concept they found relevant from a reading would fall under key item extraction, then providing specific examples of this concept would be categorized as specification.

While this model is comprehensive and does condense and incorporate many of the prevailing theoretical models of sensemaking, its general applicability also gives it a lack of tangible grounding in task-specific scenarios. The authors themselves name a few extensions they would like to be done off of their research, such as the development of “sensemaking assistants” that utilize machine learning and AI to partially automate certain cognitive tasks for users in the hope of giving them more time to focus on the more personally-satisfying parts of the sensemaking process, such as analyzing the content of key sources rather than having to find which sources are relevant, or structuring and visualizing the data instead of collecting it (2020). Through their more recent research in trying to understand which cognitive mechanisms are used during sensemaking, for what purposes, and in what defined order (if at all), their findings supported pre-existing themes in previous literature about the strong influence of prior knowledge in how individuals would search for information, the importance of resolving and incorporating conflicting information in sensemaking, and the nature of sensemaking as fundamentally creative and focused on extracting new insights and reflecting on new knowledge as opposed to simply just bringing together and condensing information.

Given this prior research, it appears timely with the growing widespread use of AI particularly in education or for educational purposes to pursue further investigation into how these tools shape and assist our sensemaking process. The objective of this study is to add to the nascent but growing body of research in AI-mediated sensemaking by using Zhang and Soergel's model as a frame of comparison to actual sensemaking activities in an educational context, namely the reading and understanding of academic literature. Specifically, we will be paying attention to the type, frequency, and order of cognitive mechanisms learners use, and compare this behavior in AI-mediated and unmediated settings. The results of this study benefit the field of HCI in two valuable directions: firstly, by evaluating the existing high-level theoretical framework used in sensemaking and offering potential improvements to its ability to robustly describe a variety of sensemaking activities, and secondly, by adding to our understanding of how AI affects the way we make sense of the information around us and how we can best utilize it to improve educational outcomes.

1.1 Research Questions

1. Does Zhang and Soergel's model contain all the sufficient[1] activities or cognitive mechanisms for sensemaking when AI-mediation/assistance is involved?
   a. If so, does the frequency, emphasis, or order of activities performed by users differ meaningfully from when they perform the task without AI mediation?
   b. If not, what new activities, processes, or cognitive mechanisms must now be given more consideration when performing AI-mediated sensemaking?

---

[1] The word *sufficient* is key here – as noted earlier, because there are many ways to go about sensemaking, we are only interested in the activities and cognitive mechanisms that satisfy the conditions to enable sensemaking rather than identifying which of them are **absolutely necessary** for it (since there is still contention on whether or not this is even possible).

2. LITERATURE REVIEW

2.1 A Brief Overview of AI

To get a better understanding of AI-mediated sensemaking, we must begin with a clear explanation of the characteristics, capabilities, and limitations of AI, as well as commonly used AI tools in the educational space and their respective use cases.

The European Union Commission defines AI as "systems that display intelligent behavior by analyzing their environment and taking actions – with some degree of autonomy – to achieve specific goals" (2018). Over the last decade, these systems have been growing in accuracy and computing power driven by the exponential increase of data availability and investment from governments and corporations due to their large commercial and industrial success (Boucher, 2020). Also known as predictive AI, these models are trained by providing them with large amounts of data in various formats like text, images, and audio. Over systemic repetitions, a large population of these models are given slight algorithmic differences to induce are compared to an expected correct result, and the models that display the lowest amount of errors (the difference between the model's output and the expected output) are selected, retrained on more data, and given more slight variations until a model with high amounts of probabilistically accurate output is obtained. This is a simplification of how AI models are trained and optimized, but this level of explanation should be suitable for the purposes of this paper.

When arranged by their functionalities, current versions of AI can be sorted into three groups: reactive machines, limited memory, and theory of mind AI (Dhokare and Gaikwad, 2021). Reactive machines are not capable of "learning" from their experiences and are only capable of making decisions based off the information being shown to them in the present moment. They can be given a set of prior conditions or ways of responding to input, but they are incapable of using past input to adjust or improve their future responses. Limited memory AI is the category that most current forms of AI belong to. These models improve from reactive machines by gaining the ability to learn from historical data and input as well as using a technique called deep learning to improve its outputs the more information is fed into it. In the future, developers of AI seek to grant it empathetic thinking and self-awareness, hence the name theory of mind, but these features do not exist yet and have their own set of benefits and concerns for their users and society at large.

The increased use of AI in the modern day also has significant implications to the way we produce knowledge, critique ideas, and interpret data and patterns depending on the type of AI being used (International Science Council, 2025). For example, descriptive AI analyses large datasets much like the predictive AI discussed earlier, but instead of attempting to suggest answers or forecast outcomes, it focuses on interpreting data to identify themes or patterns already existing in it. Generative AI (known commonly as GenAI for short), on the other hand, creates new content from the data in various forms, or even simulations of prompted scenarios. Large language models, or LLMs, are popular examples of GenAI that can understand and create their own sequences of sentences by being capable of understanding how previously entered characters relate to one another and the entire string of text that was provided as a whole. LLMs also differ mainly in terms of whether they allow external users or individuals outside of the companies that create and maintain the models to have access to their LLM's weights (the numerical parameters what determine what information is prioritized or focused on) or source code (International Science Council, 2025). One last type of AI worth noting is causal AI, which provides reasoning and cause-and-effect relationships from the results of quantitative studies. This differentiation between types of AI will be useful for our later understanding of AI's use in general education, as the paradigm of each kind of AI tool might serve different functional purposes in a user's learning process.

A common thread that arises throughout the literature and is also present in this study is the inherent limitation of AI and how its imperfections affect users' experiences with them. While a formal definition has not yet been established, the phenomenon of AI models generating inconsistent information that does not match the source it draws from is commonly referred to as a “hallucination”, although other alternative terms that have more specific definitions do exist (Maleki et al., 2024). These inconsistencies, dubious claims, or inaccurate answers can have serious consequences if not spotted and checked for and can sometimes fool users into thinking that the output of the AI application is more believable than it may sometimes claim to be (International Science Council, 2025).

2.2 AI in Education

In this section, the papers found could be sorted into two categories. Firstly, we looked at studies that classified types of AI in terms of their use cases in education, as well as an analysis of the levels of involvement and human-AI collaboration achieved in an educational context. The second set of articles helped characterize current knowledge on how to maximize the effectiveness of AI tools and what techniques and practices improve its usage.

Zhai et al. performed a content analysis from studies involving AI in the educational sector and found three different classifications for each of their respective research questions (2021). Firstly, the development layer referred to the creation of educational tools such as intelligent tutoring systems and electronic assessments using AI, and mainly discussed what kinds of algorithms would be best for programming and creating these tools. The second layer, extraction, focused on improving learning outcomes and retention through methods such as artificial neural networks that would provide immediate feedback to students, modelling and argumentation tools to help students understand explanations behind the answers they learn, and AI that could help teachers provide personalized support to students that would match their preferred learning styles and pace. The last dimension, application, looked at the formats for how these educational tools could be presented, such as using extended reality (XR for short), role-playing, and gamification to be areas where AI integration could be particularly beneficial. Meanwhile, Zhao's work on collecting studies of AI-assisted assessments for higher education revealed 69 different AI tools over 81 articles, with six distinct categories of application (2024). Similar to Zhai et al., intelligent tutoring and personalized learning showed up as another popular category, surpassed only in frequency by the category of educational robots and chat assistants (which in this study refer to LLMs such as ChatGPT, Claude, or Perplexity) as well as the category of automated assessment and feedback, another category this paper has in common with Zhao's identified categories. These two literature reviews help emphasize the increasing popularity and use of personalized learning, immediate and accurate feedback, and convenient and flexible AI chatbots as the main use cases of AI in education.

In terms of the kinds of interactions or the relationships learners could have with AI, a number of different frameworks were found. One example, the Passive-Participatory (or #ppAI6) framework, looks at the level of creative engagement learners are able to reach with AI that they could not do by themselves. The levels range from purely passive consumption of AI-generated content without any understanding taking place between the learner and the AI tool, up until the sixth level of expansive learning which has the AI tool making sharp and helpful critiques of the learner's rhetoric, providing holistic recommendations, and deepening the user's ability to problem-solve (Romero, 2025). Romero also conducted a literature review to situate the current capacity of AI tools in education to meet maximum levels of creative engagement, and found that no tools were able to surpass the fourth level of collaborative content creation, defined by the working on solutions in a smaller group setting that is assisted or facilitated by AI but has its final products driven by individuals and not by the AI tool itself. The research also showed that a vast majority of tools for creative engagement for learners tended to remain at the second level of interactive consumption (where the AI tool adapts to the learner's input but does not require creative input from them) because the ways in which they are allowed to interact are rigid and already predetermined by the tool, like in the case of an Intelligent Tutoring System (or ITS for short).

Table 2. *An adaptation of Romero's Passive-Participatory AI framework, also known as #ppAI6. (2025)*

Passive consumer	Interaction	Individual content creation	Collaborative content creation	Participatory knowledge co-creation	Expansive learning supported by AI
The learner consumes AI-generated content without understanding how it works	The learner interacts with AI generated content. The AI system adapts to the learner's actions.	The learner creates new content using AI tools.	A team creates new content using AI tools.	A team creates content thanks to AI tools and the collaboration of stakeholders in a complex problem.	In formative interventions supported by AI, participants' agency may expand or transform problematic situations.

This research is useful for helping to moderate expectations about the capability of AI tools to assist with the sense and meaning-making process that is inherent to learning, and that with the tools we currently have, we are not expecting AI to act as an equal or a peer that is capable of complex problem-solving, but we can make observations about the quality of learning we are able to extract for individuals from the range of an interactive, responsive tool that gives and receives input to a supportive driver of peer learning.

The idea of AI as a peer learner is also supported by Yan's research on creating a conceptual framework for human-AI interaction in an educational context, called the APCP framework (2025). The model's conception is strongly informed by the arrival of agentic AI, a form of AI that can understand and execute context-specific tasks independently, use complex reasoning, and act collaboratively with other AI applications. In light of this, Yan chooses to model their framework in terms of AI agency, starting with AI as an “Adaptive Instrument” that focuses on automating clerical tasks of lower importance to free up the human user to focus on more creative or higher-order tasks like analyzing or evaluating the material. This level is followed by AI as a “Proactive Assistant” that can now offer suggestions and flagging areas for concern or improvement but still ultimately leaves the final judgment to the human, which is a similar description to Romero's fourth level of creative engagement discussed earlier. Once AI reaches the level of a “Co-learner”, it is now capable of sharing meaningful parts of the workload for a task, engaging in meaningful dialogue and learning from the feedback provided to it by the human, and providing genuinely novel and insightful insight to the human user. To do this, some further development has to be undertaken in explainable AI, which is the sub-field of making AI applications or tools whose reasoning process and ways of arriving at a decision or recommendation are understandable to human users so that the solutions that they offer can be justified with the same logical steps that a human would use for their own solutions (Khosravi et al., 2022). The last level, AI as a “Peer Collaborator”, is a type of AI that does not exist yet, which is a similar trend with the past frameworks discussed. In this level, AI is practically on an equal level of reasoning as another human and is mainly used in educational settings to contribute and guide a discussion with meaningful input. However, Yan notes that particular care must be taken to create an AI that is able to offer useful levels of pushback and disagreement to avoid the risk of creating automated devil's advocates who contend with the points of learners just for the sake of doing so.

Table 3. *An adapted version of Yan’s APCP Framework for AI Agency in Human-AI Collaborative Learning, 2025. AI agency increases from the left side, human-exclusive, to the right side, fully distributed agency between the human and AI.*

Name	Adaptive Instrument	Proactive Assistant	Co-learner	Peer Collaborator
Goal	Task automation	Scaffold reflection	Joint inquiry	Collaboration
AI’s Role	Reactive tool	Monitor-and-suggest	Co-construct knowledge	Socio-cognitive process catalyst
Interaction	Command-and-execute	Guided partnership	Dialogic co-creation	Team-simulative
Human’s Role	Operator	Strategist	Collaborator	Equal Partner

Yan finally argues that while AI does not possess human cognition and beliefs, it is still capable of being a functional and effective collaborator because it can be programmed to imitate and follow certain behaviors that make for good teamwork, such as turn-taking, adopting a role in a team, or making its own individual contributions to a discussion (2025). This final point has a useful bearing on this study in showing that the main obstacles for AI as a useful educational tool do not lie in its lack of human behaviors and reasoning but rather a lack of explainability. Without this, human collaborators are less likely to understand how the AI arrived at the insights it contributes and lose trust in its suggestions regardless of their novelty.

Lin et al.'s study served as a meaningful reference point for observing sensemaking in an educational context by looking at AI-assisted metacognitive strategies by postgraduate students, specifically on how they integrated the use of AI with understanding academic literature. After having participants select and read an academic paper of their choice and communicate with an AI chatbot throughout the entire process, they organized the AI-assisted strategies these students used into five categories: planning, monitoring, evaluation, support, and prompting (2025). This provided a more granular view of the individual considerations and steps each user was taking when deciding what to prompt the AI to do, evaluating its responses, and figuring out what to do next. This sequence of activities shares some similarities with Zhang and Soergel's model, where planning matches the task and gap identification activity, evaluation matches the sensemaking activity, and prompting has some overlap with the iteration of going back to information searching to supplement oneself with new information after having determined what to look for. This study will build off of this research by evaluating user activities through the lens of sensemaking as well as making tweaks to our methodology to explore in what situations users choose to actively use AI prompting rather than requiring them to use it at all times.

In summary, while the papers have different levels of focus and ideas on how to classify AI tools and human involvement with them, they share a common understanding on both the current capabilities of AI as a collaborative tool as well as its current limitations for generating novel and insightful contributions to individual and group educational settings.

2.3 Maximizing AI Effectiveness

Given the settings and use cases of AI in education, an additional set of studies were reviewed to investigate the settings, practices, and ways of working with AI that affected learning performance.

In response to the increased use of GenAI and LLMs in educational settings, research is being conducted to observe if there is any relationship between using these tools and the user's level of critical thinking, or whether the current use of AI does not go beyond offloading lower-level cognitive tasks, similar to Yan's points about AI agency (2025). A study by Nasr et al., aimed to uncover the self-reported perceptions of students on their own GenAI usage and their ability to think critically, as well as testing whether or not their expectations matched these predictions when using ChatGPT to assist with their coursework (2025). After the participants' chat history was collected, semi-structured interviews were conducted and interpreted through by using Garrison's Practical Inquiry Cognitive Presence model (1999) to assess the presence of critical thinking during the tasks. The model breaks down critical thinking into four main quadrants: triggering events, the gap-identifying and problem-framing stage; exploration, the information-searching and clarification stage; integration, the knowledge comparison and synthesis stage; and resolution, the reflection and application stage. In an initial survey, the participants felt that ChatGPT was most useful for the exploration and integration stages, particularly for the follow-up prompting they would receive to extend their knowledge, brainstorm, or sharpen their understanding of a concept, but had much more neutral responses regarding their trust in the tool's conclusions. When their chat scripts were analyzed for critical thinking, half of the responses were at the level of moderate critical thinking and included three out of the four cognitive presence stages, which is characterized by having a "collaborative dialogue" that built on earlier questions, critiqued answers, and refining subsequent questions. Finally, the study highlights the importance of guided AI usage in improving a learner's educational outcomes. The researchers noted two distinct interaction pathways for their participants: one that positioned the user in a more passive role that would prompt the LLM to spark their creativity and ask it for answers, and a more active role that enabled the user to have a richer level of participation by having a back-and-forth conversation with the LLM to refine their ideas.

These results also align with Jacobsen and Weber's research on investigating the effect of prompt quality on an LLM's feedback. Their results showed that a prompt that included details such as the role of the person asking the prompt, the medium of the answer's presentation (such as academic writing or a letter to a close friend), and a clearly stated goal strongly outperformed more basic prompts and enabled the LLM to give answers that matched or exceeded the feedback from domain experts who were given the same prompt to answer (2025). In other words, the existing literature highlights the importance of preparation and guidance in getting the most out of using AI, and that learners who have a clearer understanding of their tools' capabilities and how to use them are also likely to critically think when utilizing them.

Another interesting effect of AI on student experience is its effect on self-reported motivation as shown in a study by Ward et al., attempting to evaluate the impact of AI on learners' study habits and general interest in learning. 71 university students were asked to rate AI's impact on their motivation on a 5-point Likert scale with 5 as the Very Positive rating and received an average score of 4.17 (2025). While the authors offered some suggestion as to why AI tools improved motivation, such as increased engagement, faster feedback, and a personalized experience, the relative weight of each of these factors as well as a clear causal link between them and the reported scores was not established in the study. This is valuable for the research as Zhang and Soergel note the importance of motivation as an element of sensemaking, in particular "activating a sense of curiosity about the world" (2014), and both motivation and curiosity can build off one another to drive an individual to want to learn more about a topic or continue to ask and refine their questions to obtain more holistic understanding.

A final effect of AI on educational outcomes to consider is the matter of how easy it actually is to achieve a strong level of human-AI collaboration, especially when the learner may have varying opinions about the accuracy and trustworthiness of the AI's responses. Chen et al. recruited 47 expert trained tutors and 48 non-tutors to evaluate different types of AI-generated praise responses for a student's performance in a simulated mathematics task. The explanations for the AI's praise had three levels of variation: a simple label of whether or not the AI thinks the response was correct or wrong, an additional feature of highlights in the praise responses that help support the AI's decision, and finally, generated reasoning text to further explain its decision (2025). The experiment found that the AI generated explanations affected the tutors on the dimensions of accuracy, reliance, and time efficiency: firstly, no significant differences in evaluation accuracy were observed between the 'AI-only' generated responses and the human-AI collaborated responses. Secondly, the novice tutors were more likely to trust the AI responses and benefit from its already respectable baseline accuracy, while the experts were more likely to critique these responses and occasionally ended up suffering accuracy issues as a result. The researchers also found that tutors tended towards over-relying on the AI explanations whenever the textual reasoning paragraph was included but were not as prone to this issue when only inline highlighting was used. Lastly, tutors were actually performing their evaluations faster without the assistance of AI possibly due to having to process the additional cognitive load of evaluating how much they could trust the AI's responses, which meant that for this specific kind of task, tutors ended up spending more time to complete the activity while not achieving any significant accuracy gains in return. While the study was limited to the specific task of evaluating praise, the research still highlights potential backfire effects in human-AI collaboration that can result from a lack of understanding of the AI's capabilities, even when- and sometimes especially when- the user themselves is knowledgeable about the task being performed.

2.4 AI-assisted Sensemaking

Because sensemaking is an integral part of learning, research into understanding the learning process through the lens of sensemaking is becoming an emerging topic of interest. In this section, we seek to enumerate other studies investigating similar outcomes as us and to highlight how our work will build off of theirs in the conclusion.

Chen et al. observed how high school students integrated ChatGPT into their knowledge-building practices over a 14-week program with rigorous instructional guidelines on how to use GPT and what questions the students could ask themselves as they completed each of the assigned tasks over each week (2023). These tasks were part of two distinct phases in the study that would highlight how differently set goals and objectives would change which knowledge-building steps students would do independently, which ones they would do with the assistance of GPT, and in what order these steps would be completed. In Phase 1, students had to critically evaluate an essay generated by ChatGPT and use their questions about the essay to participate in critical discourse and build new knowledge. Here, knowledge creation began with new ideas directly from AI, which students then independently participated in discourse about, framed a problem from the text with the help of AI, and then came up with new ideas from the questions they asked. In Phase 2, students were asked to prompt ChatGPT with questions, participate in a dialogue with it, share ideas with one another, and then synthesize the conclusions they gained from their peer sharing in their own essay about the topic. While the order of steps that were taken was changed, the use case of ChatGPT in both phases was similar – a generator of ideas and framer of problems. Other knowledge building steps, such as critical discourse, evaluating the value of ideas, and coming up with more refined ideas based on initial thoughts were exclusively performed by the student.

Meanwhile, Silvola et al.'s research focused on understanding AI-mediated sensemaking in higher education, specifically for the task of completing academic writing assignments. To answer their three research questions, they observed points of ambiguity students faced when integrating AI into their sensemaking process, identified sensemaking practices specific to using AI during academic writing tasks, and noted the roles and purposes they assigned to GenAI as a sensemaking tool (2025). To accomplish this, 22 2^{nd}-year students from an educational sciences course in university were asked to formulate an essay question, prompt out a ChatGPT essay based on this question, write another essay themselves to answer the question they came up with, and to finally reflect on the essay-writing process and how it differed to or was influenced by GenAI. Student essays were graded, and a qualitative analysis was performed on the final reflection to observe trends in the participants' sentiments towards AI. The value of this paper was highlighting tensions, and sensemaking strategies students used that were specific to AI. Three main categories of difficulty arose from the analysis: firstly, confusions regarding the clarity, depth, or creativity of the AI-generated text, secondly, a growing awareness of the limitations of the AI to perform at a certain capability for the given task, and lastly, arising ethical concerns about the use of AI for the task. These points of tension were mainly addressed by cross-checking generated content and comparing it to other sources, adjusting their prompt output in the hopes of getting a better or more accurate answer, and rethinking their workflow or process for what stages of their task they would use AI in their work. In conclusion, the research adds to the field by highlighting specific concerns that students have when using AI in their academic work and being able to enumerate tangible strategies that could be used to resolve these points of tension.

2.5 Conclusion of Related Literature

In conclusion, there are three important takeaways from the literature: an understanding of the current capabilities, limitations, and benefits of using AI, the current use cases and levels of synergy we can achieve with human-AI collaboration, and the strategies that have been employed to maximize the gains from using AI tools in education. Because the field of AI-mediated sensemaking is now more relevant than ever, this study aims to understand and model this process using an already-existing framework for individual sensemaking as a foundation, and then determining if the cognitive mechanisms employed during AI-mediated sensemaking change or are used in a different order.

There are three ways our study aims to build on the existing research, particularly from the literature on AI-mediated sensemaking. Firstly, in Chen et al.'s study, by the nature of the program the students were a part of and the tasks they were instructed to do, the knowledge-creating steps students participated in were quite homogenous because the instructions of tasks as well as the questions to ask themselves when performing the tasks were explicitly given to them. In our study, we would like to simply give participants an overarching task or goal in the form of an expected output and then observe the set of self-determined steps they take to achieve that goal without any further guidance. The second adjustment this study will do is to observe AI-assisted sensemaking relative to a baseline of unassisted sensemaking, as Chen et al.'s research looks at differing processes with AI-assisted knowledge building when the task changes, but does not compare this to a case in which the students are not using AI to accomplish the task at all. Thirdly, while Silvola et al. were able to identify sensemaking processes specific to AI, our study aims to broaden these processes in general sensemaking language so that it becomes easier to compare what sensemaking activities in AI-unmediated sensemaking might be affected or replaced when AI is used for the exact same task. This general language takes the form of Zhang and Soergel's cognitive mechanisms (2020) and allows us to label and then map these activities like we would for any other task that involved sensemaking. If through our research we find that there are cognitive mechanisms that are AI-specific, this would also serve as an answer to our research questions in the sense that current models for individual sensemaking must be adjusted when AI mediation is involved. In a sense, this research aims to bridge the best qualities of the two studies by comparing AI-mediated sensemaking to non-mediated sensemaking in the context of reading academic literature, and representing those differences in the form of a process or user flow model to visualize what activities are taken, who or what performs them, and in what order.

3. METHODOLOGY

This section outlines how the research was conducted, participant sampling, a brief overview and justifications of the methods and questions used, and a final note on the limitations of the study.

3.1 Experimental Design

Participants were 5 second-year students at Tampere University's Human-Technology Interaction master's program, ranging from ages 23-29. Students were invited to participate in the study through messages in the university intranet and local Telegram groups for students in the major, and the selected participants were the first ones to volunteer. The plan for the research was to interview at least five participants and not more than ten to balance a diversity of responses with workload and time constraints. All the participants were required to have had previous experience reading and analyzing academic literature, as well as having some experience using any AI tools. Initial background questions revealed that participants would read around 2 academic papers in a week, a number that was set to grow as all of them were beginning their thesis projects and were consuming a larger number of papers for their research, topic scoping, and literature reviews. All five participants identified as women, but this gender bias was also reflected in the distribution of the program's cohort.

After an initial interview to gather the background information above, a contextual inquiry was used to investigate the research questions because it would enable us to investigate the participants' user flow and general 'ways of working', as well as to incentivize users to also share their thoughts about their process. A contextual inquiry was also preferable to a usability study because the main research questions do not have to do with the advantages or disadvantages of AI tools or their features in a vacuum, but rather how these features as they currently exist affect the way users analyze, interpret, and make sense of the information from academic literature. As an overview, the study was split into four main parts: the background interview, two sensemaking/analysis tasks, and a post-study interview. After completing each task, the participants were also given a set of post-task questions about their analysis process. In both tasks, participants were given academic papers in their field of study and were asked to read the paper and take notes on it such that they would be capable of writing a report or making a presentation about how the content of the paper relates to their prior knowledge and offer implications from the paper on their subject domain. The participants were given different papers because changing the organization, content, and writing of each piece while keeping all the topics still under the general umbrella of human-computer interaction would help introduce some variation of cognitive mechanisms used by participants, and therefore also give us more scenarios to view how AI-mediated reading affects each individual's sensemaking process.

Table 4. *Participant Academic Papers for Unmediated and AI-Mediated Sensemaking Tasks.*

Paper title	Task type	Participant
A User Interface Study on Sustainable City Trip Recommendations	Unassisted	P1
User-Centered Design with AI in the Loop: A Case Study of Rapid User Interface Prototyping with “Vibe Coding”	AI-assisted	P1
Privacy or Transparency? Negotiated Smartphone Access as a Signifier of Trust in Romantic Relationships	Unassisted	P2
Fact Checking Chatbot: A Misinformation Intervention for Instant Messaging Apps and an Analysis of Trust in the Fact Checkers	AI-assisted	P2
Assessment of Forward Head Posture and Ergonomics in Young IT Professionals – Reasons to Worry?	Unassisted	P3
How do digital threats change requirements for the software industry?	AI-assisted	P3
Can GPTs Evaluate Graphic Design Based on Design Principles?	Unassisted	P4
Phish Phinder: A Game Design Approach to Enhance User Confidence in Mitigating Phishing Attacks	AI-assisted	P4
Influencing Incidental Human-Robot Encounters: Expres-	Unassisted	P5

sive movement improves pedestrians' impressions of a quadruped service robot		
From Cloud to Edge: Rethinking Generative AI for Low-Resource Design Challenges	AI-assisted	P5

However, in the second task, the participant would be allowed to use any AI tools of their choice to process and analyze the paper but would have to be using the AI tool Semantic Reader by default. Semantic Reader is an AI-powered reading overlay that seeks to improve the reading experience for academic papers by improving the usability, readability, and knowledge building functions of a basic document interface. Two specific features to pay attention to are the AI-generated highlights that flag important information to the reader, as well as the clickable-in-document citation cards that summarize and explain the literature that gets cited within the paper being read (Lo et al. 2024). The value of the highlights is that they are color-coded and categorized according to the methods, process, and results of the paper, which help users not just skim a paper faster when necessary but also do so in a way that ensures that they catch all important information. Meanwhile, the in-document citation cards help a reader situate a paper in the context of its field by providing the user with AI-generated summaries of its related literature within the UI of the paper itself. This saves the user the work of having to scroll down to the references section, search the paper in a new tab, and read the abstract of the paper to make the connections between the two works, which can provide compounded gains in time as more papers are read efficiently. For the purposes of this experiment, the two papers were read by participants independently, so the full benefit of the Semantic Reader to connect information between papers was not utilized; however, the main use of the tool for the study was to expose all the participants to a baseline level of AI-mediated reading in the second task.

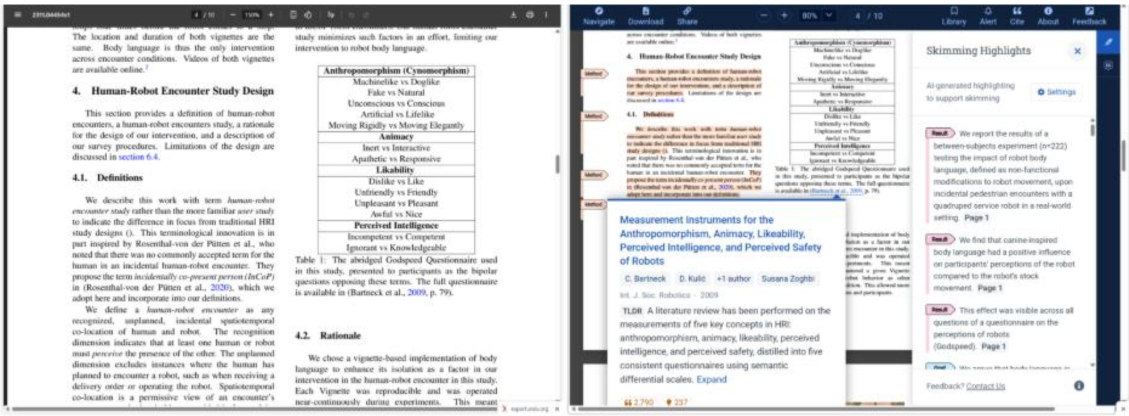


Figure 2. *A comparison of a paper viewed in ACM Digital Library's Semantic Reader (right) with its non-assisted version (left). Methods in the paper are highlighted with orange tags, a list of skimming highlights is present for the reader to individually scroll through on the right-side column, and in-text citations can now be opened into a summary card within the same document view.*

Semantic Reader was also chosen as the default AI tool to expose the participants to because it was a more passive tool based on the classifications of the CPAP framework (2025) as well as the Passive-Participatory AI framework (2025) discussed in the literature. Given that the literature showed that most AI tools used in education existed at this level of interaction, as well as the fact that the Semantic Reader was a relatively less-known tool compared to LLMs like ChatGPT or Claude, it also had the advantage of having all participants be equally familiar with using its features. Participants were given a brief tutorial on the features and uses of Semantic Reader prior to beginning the second task so that they would have sufficient understanding of the tool, but because its main use was to highlight relevant information, none of the participants felt the need to tweak or adjust the settings of the tool during the task. LLMs were also offered as an optional tool to use during the task so that we could also ask participants under what circumstances they would plan to use AI in the first place when reading academic literature, as well as when they felt they would not need to use it at all.

Participants were also asked to imagine as if they were doing these tasks in the context of an exercise session for one of their classes, and to better reflect this environment, were also given a 45-minute time limit to complete each task. To balance the time limit, the length of the papers (which varied from seven to twelve pages and had different formats and text sizes), and the effort that would be required to analyze all the information for the study within the scheduling constraints set for the thesis, participants were not required to write an actual report or presentation. Participants were also given the freedom to use any note-taking application they preferred for the task, and 3 participants used FigJam (a common mind-mapping and sticky note-style application for visualizing notes in the design industry) while 2 participants used Google Docs.

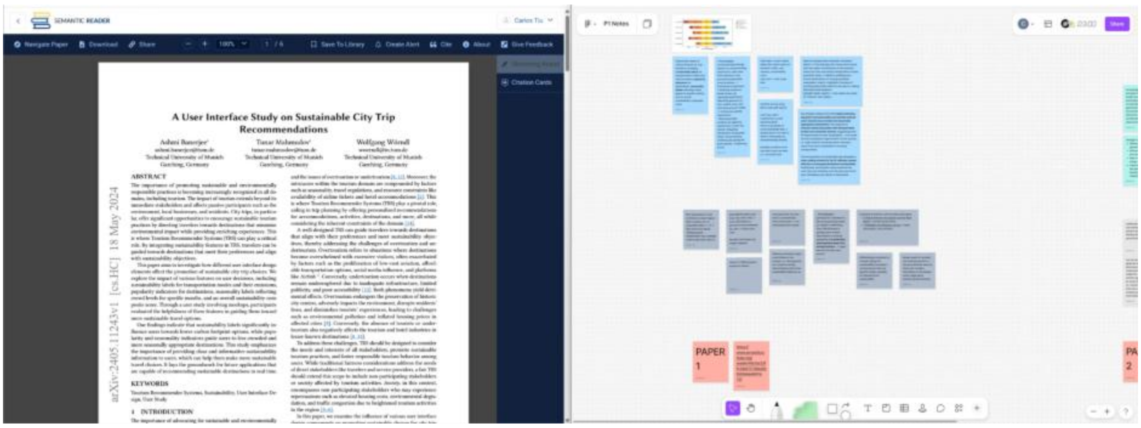


Figure 3. *A screenshot of Participant P1's screen after performing the first unassisted academic reading task. The experiment was conducted on a widescreen monitor to enable the user to take notes and read the paper in full size to minimize tab switching and to help the user keep track of their ideas concurrently.*

Right before the contextual inquiry was conducted, participants were given a briefing on thinking aloud during the study. While all participants had prior experience thinking aloud from classes in interviewing and usability testing, they could not constantly share their thoughts either due to having to exert most of their focus and attention into understanding the literature. This was accounted for in three ways: firstly, the interviewer was free to interrupt or to ask questions during the task when noticing something interesting or to confirm their idea of the participant's process or mental model. Secondly, the post-task interview asked participants about whether they had a set process for how they would analyze the paper, how they resolved any moments of uncertainty or confusion during the task, and to explain their notes as they made them and once they finished them. These questions served as a useful comparison between the participants' statements during the task itself, and how they felt about their process once they had completed the task. Lastly, the final interview questions asked more explicit questions to the participants about whether they experienced any differences in their process or thinking during the AI-mediated task and the non-mediated task, if they felt that using Semantic Reader and AI affected their reading process, as well as if they considered any other factors that affected their overall experience. This helped us identify any potential confounding variables and identify the relative impact of AI against other factors in their analysis process.

To record the experiment, three types of data were collected. Firstly, a screen recording was taken of the entire experiment to observe note-taking process as they built up, formatted, and refined their notes, as well as where their cursor would follow along for the text and which sections of the paper they were spending a lot of time on. Secondly, the audio of the screen recording was also transcribed, and the audio and video feeds were coded in MAXQDA for thematic analysis. Thirdly, the note artifacts themselves were also saved to compare the insights the participant captured between both tasks and how they were structured.

3.2 Analysis Methods

As the interview was being conducted, live notes were taken of the participants, focusing in particular on logging a timestamp of the interview, a corresponding action or statement from the participant that was interesting or insightful, and whenever relevant, a corresponding code of the action or statement based on which of Zhang and Soergel's cognitive mechanisms it falls under. We also noted any questions to ask the participants once they were finished with a task, as well as to save them later for closer inspection.

19.50	P4 deciding whether to write down challenges (types of phishing) or not	Semantic fit Specification Key item extraction
23.20	Rearranging taken notes + fitting them into new categories	Classification
24.30	Asking question about organization of notes: the further right the boxes are, the more specific the points become So far, no notes have been given that offer personal thoughts, ideas, or opinions about the paper	
26.10	"Do I really need to know about this in detail or not?"	Elimination
28.25	For a user study, P4 would go more in depth with the flow and specifics, but otherwise, would not → because of the task	
33.20	"This paper doesn't have good structure" NEED TO ASK about expected structure for the participant ANALYSIS: something also to be said about cognitive dissonance and additional friction in sensemaking when the structure of the information being presented does not necessarily align well with the user's cognitive model	

Figure 4. *An excerpt of the live coding table, containing the timestamp, activity/statement, and cognitive model categories. After being reviewed alongside the screen recording, cells that corresponded with key themes in the study were highlighted green for easier recall and identification.*

After the interview, the screen recordings and transcripts were reviewed in MAXQDA, a software tool for qualitative analysis. The live-coded notes were also cross-referenced with the study footage so that the codes would be added into the MAXQDA transcript as well. Finally, after the study is completed and the thesis is submitted, the recording data and transcripts will be appropriately disposed of to protect user privacy.

4. RESULTS AND DISCUSSION

There are three sections to cover in the analysis. The first part covers extraneous findings and other insights that arose through thematic analysis, as well as additional factors that must be taken into consideration when putting the main findings into perspective. The second section begins to cover the main findings and how they answer our research questions, and the third section looks at the implications of these findings on the original model of sensemaking and provides recommendations for future work.

4.1 Non-AI Related Factors Influencing Sensemaking

One theme that arose from the thematic analysis was the importance of other factors that impacted each participant's work process. These factors must be given ample consideration so as not to explain all the differences in how participants acted in both tasks purely to AI-mediation. The table below highlights the list of prominent factors that participants brought up, rated from the most frequent to the least frequent.

Code System	P1	P2	P3	P4	P5
Factors influencing analysis process					
Reason for academic reading/Task requirement	10	2	4	4	3
Set routine/process for paper analysis	2	3	1	5	5
Expected content based on sections of the paper	1		2	4	1
Time limit/time pressure	2	3	2	1	5
Having AI as part of the reading process	1	1	2	2	5
Interest in topic	2	3	2	1	1
Technical/domain-specific terminology	2	3	2	1	
Amount of prior knowledge on the topic		3	1	3	
Capabilities/tools in the reading interface	1	1			1
Equipment				1	

Figure 5. *Code system of factors that influence how each participant analyzed the academic paper. Higher numbers/redder cells indicate that this issue was brought up by that respective participant more.*

4.1.1 Reasons for Academic Reading

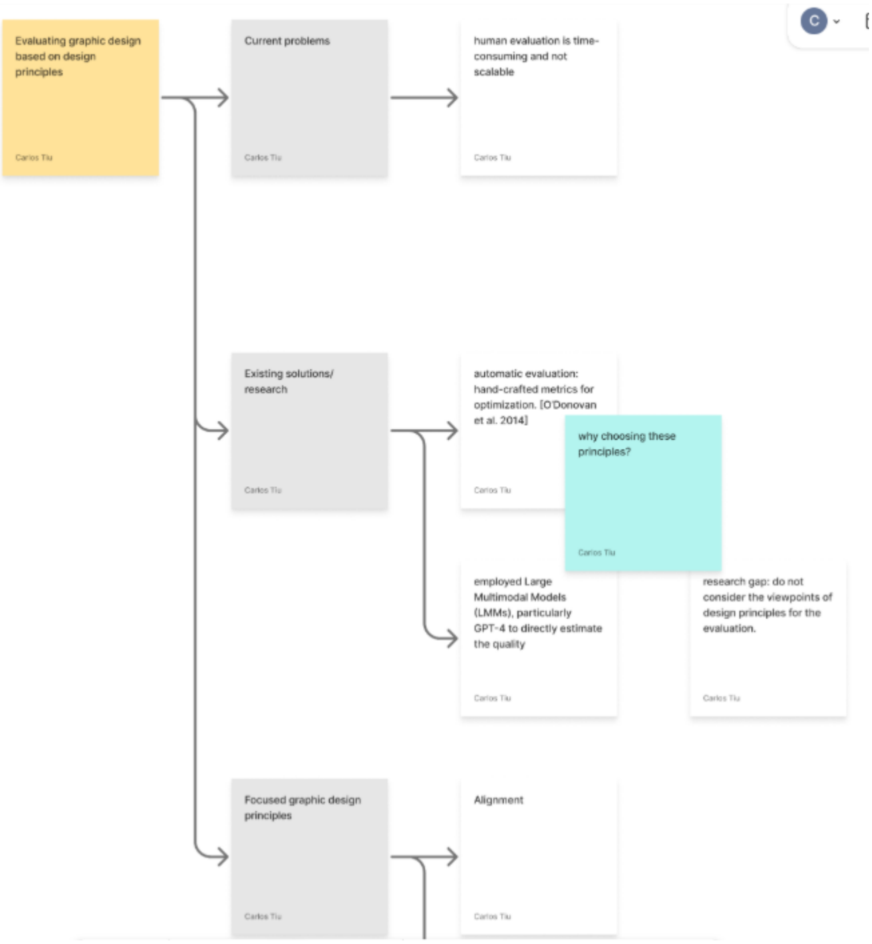
Firstly, we can see that a participant's purpose for reading an academic paper heavily informs how they choose to analyze a paper, and in turn, what details they would care about trying to make sense of. For example, P1 noted that in the context of a school assignment, they were coming into the task with the expectation of putting in "low to moderate effort", and that they “would just read without in-depth reading and just go ahead and take notes”, but would proceed in a more grounded and focused manner were they to do a thesis. In a similar respect, P3 stated that “If I already know what I want from the paper, one specific thing from the paper, I would not even read the paper”, referring to the fact that they had to read both papers in the task in full and understand what they were talking about instead of skimming through the paper to find a conclusion or take note of what method was used in the study. This factor aligns well with the sensemaking model because setting up a task/problem is an important starting point for helping the individual establish what their knowledge or structure gap is between what they want to know and what they currently know. It then follows that if the task were to change from a higher order thinking task such as critically evaluating the content of a paper to a lower order thinking task such as finding and summarizing the conclusion of the same paper, the corresponding sensemaking activities the user would plan to take would also change.

4.1.2 Set Routines/Process

The participants’ set routines of how to read and analyze a paper were also strong factors in their sensemaking process. When asked at the end of each non-mediated task, all participants mentioned that they had established practices for not only how they would read and what sections they would focus on, but also how they would take note of relevant information. For example, P2 formatted their notes in both papers by making relevant categories based on their opinion on the topic, then spending the entire task extracting key information, paraphrasing it when necessary, and placing them as bullet points under these categories. They did not record any other personal thoughts or opinions on the information they recovered after this. Alternatively, P4 focused on creating a more intricate and connected web of information points of the paper by using FigJam to construct a skeleton of the research's outline, then adding sticky notes of their opinions after the paper was appropriately represented. The implication here is that when the papers being read were organized in ways the users did not expect, this tended to cause additional tension, difficulty with understanding, and general negative affect towards the content of the paper.

Prior knowledge
- Comparison are between female and male
- Sharing pins/patterns is way easier
- Consensual is fine
    - share devices and credentials
    - Mutual Tracking App 73.7
    - socio-economic status and sharing behaviour
    - only acceptable with negotiated consent and trust, reciprocity of access, context and frequency
    - Boundaries are dependent on relationship
        - explicit or implicit boundaries
        - Most people agree that access and usage behaviors should be negotiated, mutual, and consensual.
    - safety or well-being, or a need to verify they were not engaging in problematic behaviors:

- Non-consensual is not fine
    - violating their partners' privacy and betraying their trust
    - 'snooping' behaviors amongst couples
    - prevalence of intimate partner monitoring
    - concerned about infidelity
    - blurry line between "justifiable" or benign non-consensual access and abuse



Figures 6 and 7. *Visual artifacts of P2 (top) and P4's (bottom) note-taking process demonstrating differing approaches to paper analysis and by convention, their underlying sensemaking process.*

At times, these preferences also appeared inane but were necessary to the participants so that they could analyze the paper properly. At the beginning of the first task, P5 noted that they had to make the color of their sticky notes blue because it was a color that they felt helped them remember information better. The other two participants chose colors for sticky notes purely based on contrast (such as making it clear that information taken from the paper and ideas and opinions were categorized differently), but this practice was so routine to P5 that to them, not following this basic but important step would worsen their experience for the task.

It is also possible to visualize this difference in style and process by examining a diagram of the cognitive mechanisms each participant used throughout their tasks.

Code System	P1	P2	P3	P4	P5
Zhang and Soergel's listed cognitive mechanisms					
Examining anomalies/inconsistencies					
Socratic dialogue	1	3	2		1
Semantic fit	3	7	2	3	1
Elimination	2	5	2	1	
Examining relationships					
Inference	5	1	1	1	2
Explanation-based mechanisms	2	3	2	1	
Classification	5	4	3	4	3
Stereotyping		1			
Analogy and metaphor					
Comparison	2	3	1		
Examining concepts					
Specification	1	3	1	2	1
Definition	2	2	1		4
Processing new data					
Schema induction	4	1	1	2	1
Generalization					
Summarization	5	3	1		
Judgment/evaluation	6	8			
Restatement	6	4	1	1	7
Key item extraction	6	11	7	7	5

Figure 8. *Cognitve mechanisms used by each participant for both tasks. Note that not all mechanisms could be noted due to either time constraints or ambiguity during coding regarding whether the action of the participant counted as an explicit application of a cognitive mechanism. In an ideal setting, a second coder could also be added to reduce bias, but this was not possible in the context of the academic requirements.*

For example, P5 put a strong emphasis on having clarity on all relevant terminology as they were writing their notes, and was the only one of the participants to actually look up the definition of words they were unfamiliar with on a Google search instead of relying on piecing together the meaning of the word through contextually reading it in the paper. This is reflected in their tags for the *definition* cognitive mechanism as well as in their transcripts (ex: “Yeah, I just Google the words and find the meaning so it's just easier for me to understand like, oh, okay, this means that then I can like fully understand the sentence.", Task 1 Transcript). Alternatively, P1 was one of the participants who tended to use *restatement, judgment, and summarization* a lot because they would paraphrase or shorten excerpts from the text as they input them into their notes, and then put an additional arrow comment to share their thoughts on the text or offer alternative explanations to claims the writers made (ex: “I'm going to put a note: “probably would be more accurate to get real data, i.e. conversion rate”, P1 recording). It is important to note however that the cognitive mechanisms ultimately only represent the surface-level expressions of the users’ thoughts, which is to say that there is definitely some unobserved reasoning, information retrieval, and connections between ideas that are being made by the participants that we cannot capture with the study. These associations between the participants’ existing knowledge and the information they are processing are often made subconsciously and therefore not as likely to be expressed, but there is still value in being able to show that the differences in how individuals like to think and process their information can be reflected in the cognitive mechanisms they use.

4.1.3 Time Pressure

Another major consideration the participants had during the study was the limited time they had to read, analyze, and make note of the paper. For example, P2 noted that they had not read two academic papers deeply in a row in a while, and so the experimental setting induced some stress onto them that they were not fully prepared for. Given that the AI-mediated task was always done second for all participants, the effort they had spent in the first task might have also influenced how motivated they were to read and analyze another paper after having just read one previously, but a common piece of feedback from all 5 participants is that their notes and analysis process were strongly framed by the time limitation. For example, P4 noted that when reading an academic paper with a more open time window, their routine would be to read through the paper one more time after making the notes, make any adjustments that were needed, and only then start would they start writing reflections about the paper. In this regard, the study also looked at what the participants’ priorities were when reading a paper, such as making sure that they had full understanding of the content, or acquiring the minimum amount of knowledge needed to be able to start offering their judgments and recommendations.

4.1.4 Interest and Prior Knowledge

The last group of non-AI related factors to consider revolve around the user's interest in the subject matter being read as well as their background knowledge on the topic. Background knowledge had different effects on participants depending on the level of familiarity they had with the content of the paper. When the level of knowledge was too low due to either difficult jargon or just the style of writing in the paper, user analysis suffered and frequency of cognitive mechanism use decreased. For example, P3's AI-assisted task paper “reminded [them] of IELTS where the participants try to write complicated on purpose just to get more scores”, and they opted to upload the paper to Claude to summarize and identify key items in the paper for them. P2 also had a similar experience where they encountered numerous bar graphs but could not interpret their values, and after uploading the results and visualizations to GPT but failing to get a trustworthy answer, decided to skip all of the results section altogether. On the other hand, having utmost familiarity with the content tended to result in lower interest for the participant because they did not feel like they had anything new to glean from the text. One example would be P1's first task paper regarding creating an application for city trip recommendations. Given their prior experience working in the design industry, they had “similar work experiences to do, so [they] could just write about that, but implications on your field of study is a bit tough because [they] don't find anything groundbreaking.” However, a proper level of background knowledge helped spark interest in the topic and helped users want to bridge their current understanding with the insights in the paper. One such case would be P5’s first task about improving human-robot interaction with quadripedal robots. Upon learning that the robot in question was a Spot robot that they had worked with in a previous class, they felt that “I know something about it previously, so I'm not completely going into this research paper clueless.” These findings match Zhang and Soergel's task analysis activity under the subtask of “activating a sense of curiosity about the world” (2014), and highlights the role of interest and knowledge in motivating participants to want to engage meaningfully with the academic text in the first place.

In summary, the other factors that participants were affected by have interleaved relationships with how they went through each paper and which cognitive mechanisms they used. It is important to recognize that while using AI may have meaningful effects on the sensemaking process, it must be understood in the context of other aspects of the learning setting that could improve or reduce the gains users can extract from its use.

4.2 The Role of AI in Sensemaking

In this section, we will answer how AI-mediated reading affects either the type, frequency, or order of cognitive mechanisms used.

4.2.1 Frequency of Mechanisms Used

The table below shows a comparison of which mechanisms were used between the two tasks – unmediated and AI-mediated, respectively. While it cannot be guaranteed that the difference in cognitive mechanism use between the two tasks is due purely to the effect of AI mediation as established earlier, we argue that some of these changes can be explained by using the lens of human-AI collaboration and its role that emerged through the tasks. Among all the cognitive mechanisms, *elimination, comparison,* and *specification* stand out for having occurred in more instances overall in the AI-mediated task than in the non-mediated task, even if only slightly. This becomes even more salient when considering that the number of cognitive mechanisms recorded under the AI-mediated tasks is significantly less at 45 compared to the mediated tasks at 78. Looking through additional codes in the study, we can see that these cognitive mechanisms also tend to match the instances when participants were using AI, including both Semantic Reader and LLM use. Elimination is used when the sensemaker determines that certain information is not relevant to their inquiry, and excludes it from their research. In this case, it was used by P2 when determining that certain Semantic Reader highlights were incorrectly labeled, and after seeing it happen one more time, they stopped paying as much attention to the highlighted sections.

Figure 9. *Code system differences of participants in unmediated (left) and AI-mediated (right) academic reading settings. P1's recording is not included because their study was recorded in one full take, whereas the other participants had their tasks recorded separately.*

P3 also used elimination in the AI-mediated task, specifically when they were deciding which information was irrelevant from Claude's response. After deciding that they were only interested in understanding the main idea and results of the paper, they ignored the other feedback Claude gave because “there wasn't anything special about the methods.” Given that most LLMs tend to suggest additional follow-up prompts or extend their answers beyond the initial prompt, it would make sense that eliminating irrelevant feedback would be more important under AI-mediation.

Comparison also increased for P2 in the AI-mediated task, as they specifically mentioned in which cases they would consider using AI to help them accomplish a reading task and how they would use such tools. They put particular emphasis on comparing their understanding of the AI's output with their own interpretation of the text (“And what I would do is I would go through the paper still, try to kind of like compare the answer that AI gave me versus what is available on the paper.”), and when they started to run into inconsistencies between the parts of the paper they did understand and the AI's interpretation, they would always favor their own over the AI's. This behavior was also similarly imitated by P3 when they prompted Claude to summarize their assigned paper for them, but they added that they did not feel that such a comparison was as necessary “Because [they] felt like [they] understood the article by summarizing with Claude.” In this situation, P3 did not have any useful interpretation to have of the paper by themselves and therefore deferred to Claude's interpretation.

Lastly, *specification* refers in this case as a counterpart to the cognitive mechanism of *generalization,* where the sensemaker further explains or contextualizes a concept by giving specific examples or enumerating applications of a general principle. In the case of this study, participants tended to specify exact outputs from an LLM as well as to ask it to elaborate on answers it provided. For example, P3 asked Claude to explain a bullet point it provided about “designing for longer product lifecycles”, while P5 would provide further explanation or connected examples to Semantic Reader's highlighted goal sections.

4.2.2 Order of Mechanisms Used

Of the three kinds of ways AI-mediated reading could affect sensemaking, it is most difficult to determine its impact on the order of mechanisms. One reason is because the literature has established that the nature of sensemaking is to jump back and forth between activities and drop out at any points, which means that certain combinations of tasks could have been equally likely in both mediated and unmediated scenarios. Another reason is that the changes in order could have also been explained well by the earlier considerations about differences between each participant's approach to paper analysis, such as P4's breadth and outlining-focused approach favoring *classification* and *schema induction* versus P2's opinionated sorting approach which focused on *key item extraction* and *judgment*. A final explanation worth mentioning is that it was simply logistically difficult to sort through combinations of cognitive mechanisms and identify whether they were different from another set of actions later on in the interview because the participant could now use AI, or if it was because they were performing an entirely different task. For example, we would naturally expect that participants trying to understand a conclusion of a paper would be performing different sensemaking activities from critiquing that same paper's methodology, so we cannot say with confidence that AI has a strong influence on the order of mechanisms used.

4.2.3 Type of Mechanisms Used

While we do not think that any new cognitive mechanisms were used by participants during the study, we think that given the current capabilities of AI, there are certain tasks that AI tends to be used for or offloaded to based on the codes, namely *key* *item extraction, summarization, classification,* and *restatement*. Key item extraction refers to identifying important information within the text, while classification refers to creating or sorting concepts and information being collected into groups based on similar characteristics. These two mechanisms are automatically present in the AI-mediated tasks by default because they fall under Semantic Reader's features of highlighting relevant data and then classifying it as a goal, method, or result for the user. However, key item extraction, summarization, and restatement were also used when participants prompted LLMs for insight about the papers they were reading. P2 prompted GPT to help them analyze a graph and interpret its results, while P3 prompted Claude to summarize and explain key information from a paper, as well as to rephrase its content with more reader-friendly language. These results match the conclusions by past literature, but the note artifacts of the participants shed additional light on the role other mechanisms play now that AI takes on some of the cognitive load.

Table 5. *Zhang and Soergel's cognitive mechanisms relabeled based on whether they were used exclusively by a human (blue) or collaboratively between a human and AI (green). Mechanisms highlighted in red were not used during the study.*

Inductive (data-driven, bottom-up)	Structure-driven (logic-driven, top-down)	Both or neither
Key item extraction	Definition	Comparison
Restatement	Specification	Analogy
Judgment or evaluation	Explanation-based mechanisms	Classification
Summarization	Elimination	Stereotyping
Schema induction	Inference	Semantic fit
Generalization		Socratic dialogue

In some participant notes, such as P1's, they mentioned that "having the Semantic Reader to highlight information for me helped me focus more on thinking about the paper rather than having to understand the paper", which is reflected in their notes having more reaction and opinion text in the AI-mediated task compared to their unmediated task notes. Meanwhile, P5 completely restructured their notes in the AI-mediated task into Goals, Methods, and Results columns to better fit the Semantic Reader's categories, even considering that it may be a better way to take notes compared to their original process of putting down sticky notes based on the chronological order of the paper. This feedback suggests that AI offloading some of the initial data processing with identifying important information, giving the excerpt a basic theme or category, or simplifying larger amounts of text into easier to process chunks frees up the user to focus on cognitive mechanisms that require more critical evaluation, such as creating a schema to determine how to compare different concepts, or creating causal links between prior knowledge and new information. This is not to say that AI tools could not have been used for these higher-level cognitive tasks, but it is telling to note that none of the participants considered using the tools for these purposes even when they had the option to do so.

On a final note, we posit that the cognitive mechanism of *socratic dialogue* does change slightly when the user is undergoing AI-mediated sensemaking. Our research shows that users put a greater emphasis on establishing a clear frame of mind about their knowledge, as well as the knowledge they identified in the paper they were reading. This frame became more important because users would constantly compare and contrast their understanding of the paper with the AI’s interpretation of the paper in a back-and-forth fashion, whether that was in the form of Semantic Reader categorizing a statement that appeared inconsistent, or if it was an LLM making potentially unfounded claims about a paper's content. Our theory as to why this comparison became more frequent and explicit lies in the fact that the reasoning or processing used by the AI models is unknown to the user, therefore if it ever came to a point where they would have to choose between their understanding of the information and an AI's interpretation, they would always favor their own vision unless they themselves could not offer a reasonable alternative explanation. Therefore, we suggest that while AI mediation might not add new cognitive mechanisms to the sensemaking process, it might add or make apparent some specific steps in the overall sensemaking loop.

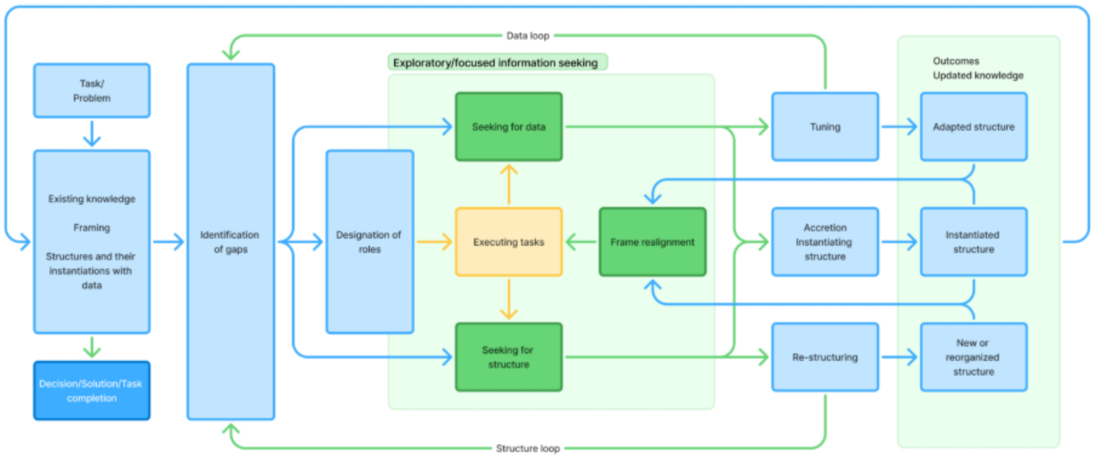


Figure 10. *A slightly tweaked version of Zhang and Soergel's model of individual sensemaking, accounting for explicit activities involving AI-mediated processing and collaboration. Blue tasks and arrows signify that those tasks and information come exclusively from the human; green tasks and arrows involve human-AI collaboration, and yellow tasks and arrows refer to tasks only performed by AI.*

In our new model of individual sensemaking, most of the steps and paths remain the same as Zhang and Soergel's model, save for the addition of four new steps. Firstly, the step of identifying existing knowledge and structures has the added term of framing in line with earlier findings about the individual having to develop stronger self-awareness about their own mental models and how they view the world and the information around them. Secondly, 'Designation of roles' is added as a step after gap identification for the sensemaker to determine what tasks they want to offload to the AI, upon which it executes its commands, extracts this information, and processes it as the sensemaker concurrently conducts their own knowledge/structure searching. After a structure is either tuned, restructured, or instantiated, the final additional step of frame realignment happens. In this step, the human takes their new or adjusted mental model and attempts to impart it to the AI model. The reason why the arrow moving from frame realignment to task execution is collaborative and not solely based on the human is because the ability of the AI to correctly “understand” and adopt the new mental model is also dependent on its weighting, characteristics, and internal capabilities. It is also possible that the realignment does not require the AI to understand the human's mental model at all so long as it is able to provide information that the human can critique and evaluate to help them build their own knowledge.

Despite the inclusion of AI as a “dialogue partner” of sorts, we would not treat this as an example of collaborative or group sensemaking but rather an additional source of information or structure that an individual can access to build knowledge. This is due again to the nature of AI as an abductive inference tool as well as a “black box” that does not arrive at its conclusions through deduction but rather a series of numerical weights that are not visible nor as explainable to the average user. This is not to say that the results of AI prompting can never be correct – the literature clearly shows that proper guidance can greatly improve the outcomes from AI use. Rather, the model simply visualizes these additional guardrails that users in the study took to minimize AI dependence during sensemaking, establish clear roles and use cases for the tool, and iterate on its feedback to obtain more accurate answers.

5. CONCLUSION AND RECOMMENDATIONS

There are three important things to take away from the study. Firstly, there are multiple factors that affect how sense-making happens, and we cannot simply reduce the observed behaviors by participants to the presence or absence of AI. Taking prior knowledge, level of interest in the topic, and the task requirement into consideration are important when looking at sensemaking in academic reading. Despite this, we can say that there is a change in AI-mediated sensemaking. While the cognitive mechanisms involved in the process do not change themselves, the lower-order mechanisms such as summarization, restatement, and categorization do get offloaded to AI, even if done so as a later option when the human is not capable of reaching a desired answer or level of understanding by themselves. Lastly, we also found that other activities in the entire sensemaking loop have become more important. Learning and framing not only the individual's goal for reading but also their understanding of the content and how it fits into their mental model becomes key, especially when it comes to evaluating AI responses and determining if they fit into their established vision.

We would also like to lastly discuss a few limitations and recommendations for future work. Firstly, my results are a timebound examination of AI's current capabilities. As future models with improved “reasoning” and explanatory power are created, AI's main role in collaborative learning with humans may change. Some individuals may be more inclined to use other features of AI that are not as common right now, and the tools themselves may find more formal or integrated use in educational systems. Such changes are likely to once again change the role of AI in individual sensemaking, and new work would have to be done to observe these developments over time.

Secondly, the study focuses on individual sensemaking and not group sensemaking, a common and important form of knowledge building to take place in educational setups. Looking at the role AI plays in helping individuals build knowledge, share it between other human groupmates or classmates, and form their own personalized reflections from a learning experience could be an interesting area for future research.

Lastly, the study also focused on unmoderated AI use as much as possible to see what individuals would prefer to do when they were trying to accomplish tasks exactly the way they saw fit to. Another interesting angle to the study would be requiring the participants to prompt an LLM in deliberate ways, or to use multiple other tools in conjunction with another as some participants suggested they would when performing a literature review. This could help examine the current limits of skillful and guided prompting in extracting the best possible responses from AI, and seeing how sense-making would change when users were tasked to engage with AI as much as possible when building knowledge.

6. APPENDIX

6.1 Interview Questions

https://proximal-lighter-ccb.notion.site/Interview-Questions-2ccf23cb427f80c2a9ddd88d1f42113c?source=copy_link

Part 1: Contextual Information

- To start off, I would like to know more about your background. How old are you?
- What is your level of English proficiency?
- What is your field of study and one thing about it that you've enjoyed learning in that domain?
- To participate in this study, you needed to have recent experience reading academic literature in an educational context. What were the common reasons or circumstances for why you had to read these kinds of work?
- How often do you have to read and make sense of academic literature during your studies?
- What tools do you normally use to record, organize, and make sense of academic literature? What do you like or dislike about them?
- Do you commonly use AI tools for any part of your studying process?
    - If so, what kind of tools do you use?
    - What kind of tasks do you use them for?
    - How often do you use these tools?

Part 2: Unassisted Reading + Sensemaking Questions

- What were the most significant or insightful takeaways you got from the paper, including but not limited to the conclusion?
- Did you have a set process in mind for how you were going to analyze the paper?
    - Were you able to follow that process? Why or why not?

    - If so, what were the key steps you took to accomplish the task?
    - If not, what did you do first and how did you decide what to do next?
- Did you have any moments of uncertainty or confusion during this task? If so, how did you resolve them?
- Could you briefly explain your notes to me?
    - Are there any areas or sections I should focus on?

[This process is repeated for Part 3, AI-Mediated Sensemaking]

Part 4: Comparisons/Overall Thoughts

- How was your experience of performing the tasks?
- Were there any techniques for reading/understanding the material that you were trying to make sure you used?
    - What about for the notes you took?
- What factors, if any, affected how you chose to go about reading the literature and taking notes?
- Did using AI tools like the Semantic Reader affect your reading process?
- If this task was more complex and required you to construct a literature review and analyze content from multiple other academic sources besides the paper you read, what other steps do you think you would take or put greater emphasis on?
- Do you have any final thoughts about the use of AI tools with reading in an educational context?

6.2 MAXQDA Code System

Code System

Code System	Memo	Frequency
Code System		449
Zhang and Soergel's listed cognitive mechanisms		
Examining anomalies/inconsistencies		

	Socratic dialogue	Thinking aloud, talking to self, critical evaluation of concepts and facts	7
	Semantic fit	Examining the reasonableness of a fact, concept, or relationship as it relates to the meaning of other concepts in the knowledge structure	16
	Elimination	Excluding facts, concepts, or relationships that are not applicable	10
Examining relationships			
	Inference		10
	Explanation-based mechanisms	Any kind of strategy or thinking process that tries to establish causality between concepts and prior knowledge	8
	Classification	Relating a concept to a broader conceptual category and grouping sufficiently alike concepts	19
	Stereotyping		1
	Analogy and metaphor		0
	Comparison		6
Examining concepts			
	Specification	Explicitly stating details about concepts and relationships as the counterpart (or reverse) mechanism of generalization - Specifying a concept with instances or examples - Specifying a claim or principle with examples	8
	Definition	Explaining or identifying different aspects of a concept, such as purpose, function, and use	9
Processing new data			
	Schema induction	Identifying similar elements/categories for two or more related concepts, phenomena, or situations	9
	Generalization	Transforming specific data into a general claim that focuses on key concepts and relationships, recognizing trends or patterns	0
	Summarization	Reducing complexity by focusing on main points and omitting all but important details	9

Judgment/evaluation	Forming critical opinions towards information being processed and connecting it with existing knowledge AKA forming and giving opinions	14
Restatement	paraphrasing into shorter form, less formal language to examine the same facts from a new perspective Focusing on facts that are essential to the sensemaker	19
Key item extraction	processing text to identify key concepts as expressed by words or phrases	36
Reading, processing, and notetaking techniques		
Rereading		1
Flagging/bracketing		2
Skimming		9
Sticky note	All forms of key item extraction get put into notes, but not the other way around	2
Reading aloud		2
Cursor to follow text		1
Bolding (+) (+)	11/28/2025 09:36 - carlos Merged with code Note taking > Bolding in notes 11/30/2025 03:31 - carlos Merged with code Reading, processing, and notetaking techniques > Font changes	7
Reading back-to-back		1
Bullet points		4
Factors influencing analysis process		
Reason for academic reading/Task requirement		23
Set routine/process for paper analysis		16
Expected content based on sections of the paper		8

Time limit/time pressure		13
Having AI as part of the reading process		11
Interest in topic		9
Technical/domain-specific terminology		8
Amount of prior knowledge on the topic		7
Capabilities/tools in the reading interface		3
Equipment		1
Effect of AI use on notes		1
AI tools in academic reading		5
Experiences of using AI tools		2
Claude		5
Consensus		2
Semantic Reader		17
AI tool use cases	Only for stated use cases of AI tools. If the participant starts explaining WHY they prefer or choose to use these tools, it goes in "Benefits of using AI tools".	37
LLM prompts/questions		8
Issues with AI use		9
Desired improvements for AI tools		7
Human-AI collaboration		4
Conventional/Non-AI tools in academic reading		
Types of tools		0
Google searching		1
Obsidian		2
Microsoft Word/Google Docs		6
Microsoft Excel/Google Sheets		7
FigJam		9
University platform		1
Benefits		1

Downsides		3
Background Info	Mainly for information about the participant	
Frequency of using AI tools per paper		6
Frequency of reading academic literature		7

7. REFERENCES

[2509.16772] AI Knows Best? The Paradox of Expertise, AI-Reliance, and Performance in Educational Tutoring Decision-Making Tasks. (n.d.). Retrieved November 26, 2025, from https://arxiv.org/abs/2509.16772

Artificial intelligence: How does it work, why does it matter, and what can we do about it? | Think Tank | European Parliament. (n.d.). Retrieved November 18, 2025, from https://www.europarl.europa.eu/thinktank/en/document/EPRS_STU(2020)641547

COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE EUROPEAN COUNCIL, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS Artificial Intelligence for Europe (2018). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52018DC0237

Dervin, B. (1999). On studying information seeking methodologically: The implications of connecting metatheory to method. *Information Processing & Management*, *35*(6), 727–750. https://doi.org/10.1016/S0306-4573(99)00023-0

Fu, Y., & Hiniker, A. (2025). *Supporting Students' Reading and Cognition with AI*. https://doi.org/10.48550/ARXIV.2504.13900

Gaikwad, A. (2021). A STUDY OF ARTIFICIAL INTELLIGENCE TYPES OPPORTUNITIES and CHALLENGES.

Garrison, D. R., Anderson, T., & Archer, W. (1999a). Critical Inquiry in a Text-Based Environment: Computer Conferencing in Higher Education. *The Internet and Higher Education*, *2*(2), 87–105. https://doi.org/10.1016/S1096-7516(00)00016-6

Garrison, D. R., Anderson, T., & Archer, W. (1999b). Critical Inquiry in a Text-Based Environment: Computer Conferencing in Higher Education. *The Internet and Higher Education*, *2*(2), 87–105. https://doi.org/10.1016/S1096-7516(00)00016-6

Kurtz, C. F., & Snowden, D. J. (2003). The new dynamics of strategy: Sense-making in a complex and complicated world. *IBM Systems Journal*, *42*(3), 462–483. https://doi.org/10.1147/sj.423.0462

LaCroix, E. (2025). Making sense of experiential education in Canada: The four lenses of faculty sensemaking. *Higher Education*, *89*(3), 611–628. https://doi.org/10.1007/s10734-024-01238-6

Lin, H., Chen, Z., Wei, W., & Lu, H. (2025). GenAI Tools in Academic Reading: A Study on AI-Assisted Metacognitive Strategies and Emotional Reactions. In E. C. K. Cheng (Ed.), *Innovating Education with AI* (pp. 98–112). Springer Nature. https://doi.org/10.1007/978-981-96-4952-5_7

Lo, K., Chang, J. C., Head, A., Bragg, J., Zhang, A. X., Trier, C., Anastasiades, C., August, T., Authur, R., Bragg, D., Bransom, E., Cachola, I., Candra, S., Chandrasekhar, Y., Chen, Y.-S., Cheng, E. Y.-Y., Chou, Y., Downey, D., Evans, R., … Weld, D. S. (2024). The Semantic Reader Project. *Commun. ACM*, *67*(10), 50–61. https://doi.org/10.1145/3659096

Maleki, N., Padmanabhan, B., & Dutta, K. (2024). AI Hallucinations: A Misnomer Worth Clarifying. *2024 IEEE Conference on Artificial Intelligence (CAI)*, 133–138. https://doi.org/10.1109/CAI59869.2024.00033

Mustafa, M. Y., Tlili, A., Lampropoulos, G., Huang, R., Jandrić, P., Zhao, J., Salha, S., Xu, L., Panda, S., Kinshuk, López-Pernas, S., & Saqr, M. (2024). A systematic review of literature reviews on artificial intelligence in education (AIED): A roadmap to a future research agenda. *Smart Learning Environments*, *11*(1), 59. https://doi.org/10.1186/s40561-024-00350-5

Nasr, N. R., Tu, C.-H., Werner, J., Bauer, T., Yen, C.-J., & Sujo-Montes, L. (2025). Exploring the Impact of Generative AI ChatGPT on Critical Thinking in Higher Education: Passive AI-Directed Use or Human–AI Supported Collaboration? *Education Sciences*, *15*(9), 1198. https://doi.org/10.3390/educsci15091198

Romero, M. (2025). From Consumption to Co-Creation: A Systematic Review of Six Levels of AI-Enhanced Creative Engagement in Education. *Multimodal Technologies and Interaction*, *9*(10). https://doi.org/10.3390/mti9100110

Sensemaking, Sensegiving, and the Challenges of Making School Changes for Students' Socioemotional Wellbeing—Lauren Yoshizawa, 2025. (n.d.). Retrieved October 24, 2025, from https://journals.sagepub.com/doi/10.1177/0013161X251316587

Sense-making, sensemaking and sense making—A systematic review and meta-synthesis of literature in information science and education: An Annual Review of Information Science and Technology (ARIST) paper—Urquhart—2025—Journal of the Association for Information Science and Technology—Wiley Online Library. (n.d.). Retrieved October 2, 2025, from https://asistdl.onlinelibrary.wiley.com/doi/full/10.1002/asi.24866?utm_source=chatgpt.com

The Promises and Pitfalls of Large Language Models as Feedback Providers: A Study of Prompt Engineering and the Quality of AI-Driven Feedback | MDPI. (n.d.). Retrieved November 21, 2025, from https://www.mdpi.com/2673-2688/6/2/35

The Semantic Reader Project – Communications of the ACM. (n.d.). Retrieved October 15, 2025, from https://cacm.acm.org/research/the-semantic-reader-project/

Vieriu, A. M., Petrea, G., Vieriu, A. M., & Petrea, G. (2025). The Impact of Artificial Intelligence (AI) on Students' Academic Development. *Education Sciences*, *15*(3). https://doi.org/10.3390/educsci15030343

Ward, B., Bhati, D., Neha, F., & Guercio, A. (2025). Analyzing the Impact of AI Tools on Student Study Habits and Academic Performance. *2025 IEEE 15th Annual Computing and Communication Workshop and Conference (CCWC)*, 00434–00440. https://doi.org/10.1109/CCWC62904.2025.10903692

Weick, K. E. (with Internet Archive). (1995). *Sensemaking in organizations*. Thousand Oaks : Sage Publications. http://archive.org/details/trent_0116403577194

Yan, L. (2025). *From Passive Tool to Socio-cognitive Teammate: A Conceptual Framework for Agentic AI in Human-AI Collaborative Learning* (No. arXiv:2508.14825). arXiv. https://doi.org/10.48550/arXiv.2508.14825

Zhai, X., Chu, X., Chai, C. S., Jong, M. S. Y., Istenic, A., Spector, M., Liu, J.-B., Yuan, J., & Li, Y. (2021). A Review of Artificial Intelligence (AI) in Education from 2010 to 2020. *Complexity*, *2021*(1), 8812542. https://doi.org/10.1155/2021/8812542

Zhang, P., & Soergel, D. (2014). Towards a comprehensive model of the cognitive process and mechanisms of individual sensemaking. *Journal of the Association for Information Science and Technology*, *65*(9), 63. https://doi.org/10.1002/asi.23125

Zhang, P., & Soergel, D. (2020). Cognitive mechanisms in sensemaking: A qualitative user study. *Journal of the Association for Information Science and Technology*, *71*(2), 158–171. https://doi.org/10.1002/asi.24221

Zhao, C. (2024). AI-assisted assessment in higher education: A systematic review. *Journal of Educational Technology and Innovation*, *6*(4). https://doi.org/10.61414/jeti.v6i4.209