



Co-constructing knowledge with generative AI tools: Reflections from a CSCL perspective

Ulrike Cress¹ · Joachim Kimmerle¹

Received: 5 June 2023 / Accepted: 20 September 2023 / Published online: 20 October 2023
© The Author(s) 2023

Abstract

Generative Artificial Intelligence (AI) tools, such as ChatGPT, have received great attention from researchers, the media, and the public. They are gladly and frequently used for text production by many people. These tools have undeniable strengths but also weaknesses that must be addressed. In this squib we ask to what extent these tools can be employed by users for individual learning as well as for knowledge construction to spark a collective endeavor of developing new insights. We take a social, collective notion of knowledge as a basis and argue that users need to establish a dialog that goes beyond knowledge telling (simply writing what one knows) and stimulates knowledge transformation (converting knowledge into complex relational argumentation structures). Generative AI tools do not have any conceptual knowledge or conscious understanding, as they only use word transitions and rely on probabilities of word classes. We suggest, however, that argumentative dialogs among humans and AI tools can be achieved with appropriate prompts, where emergent processes of joint knowledge construction can take place. Based on this assumption, we inquire into the human and into the AI parts of communication and text production. For our line of argument, we borrow from research on individual and collaborative writing, group cognition, and the co-evolution of cognitive and social systems. We outline future CSCL research paths that might take the human-AI co-construction of knowledge into account in terms of terminology, theory, and methodology.

Keywords ChatGPT · Conversational agent · Knowledge construction · Artificial intelligence

Introduction

ChatGPT is a conversational agent released by the US company OpenAI in November 2022. It is a generative AI tool that is based on large language models and pre-trained using a massive amount of data from the internet. The current version of ChatGPT (as of August 2023)

✉ Ulrike Cress
u.cress@iwm-tuebingen.de

Joachim Kimmerle
j.kimmerle@iwm-tuebingen.de

¹ Leibniz-Institut für Wissensmedien, Schleichstr. 6, 72076 Tübingen, Germany

makes use of GPT-4, a language model with more than a trillion parameters adjusted through machine learning. Soon after its release, ChatGPT received worldwide attention because its results are often impressive, and it is free and easy to use. ChatGPT is the most prominent of several generative AI tools that harvest masses of internet data and provide easily usable interfaces, allowing for language as both input and output. With these tools, users can ask any kind of question, and the tools provide written answers. Users can then re-formulate their questions, specify their expectations for answers, and build on former utterances.

These AI tools not only learn from their massive textual foundation but also from the users' reactions, so users can start an almost human-like conversation with the chat bot. Generative AI tools, however, do not possess any conceptual knowledge and have no conscious understanding of the world. They might produce text that might look adequate at first glance but lacks truth and validity or refers to entities that do not exist. With their own utterances, users need to provide relevant concepts as input, which the AI can use as a starting point for its output. These *prompts* influence the quality of the output and the fit to users' needs and expectations.

Current generative AI tools, however, are just a first impression of what AI-based conversational agents will probably become in the future. What do these new developments mean for learning, education, and knowledge-related processes? Do they, eventually, impede human learning, or promote it? In the corresponding debate the range of positions reaches from the demand for general bans to the ubiquitous use of such tools in education (for an overview see Kasneci et al., 2023; see also Cooper, 2023). We are convinced that CSCL research can make important contributions to this debate by asking the right questions and reflecting on what is already known. The general question that arises is: What is the human part of communication and text production, and what could be the AI part?

From a CSCL perspective, the interaction of multiple agents is relevant to communication, knowledge construction, and written composition. Communication in this context refers to an information-exchange process in which meaning is constituted and in which new knowledge can emerge in a group, which we refer to as knowledge construction. The concept of knowledge used here is therefore not a purely individual one, as it has a long tradition in philosophy and psychology, but a social concept of knowledge (for an overview see Oeberst et al., 2016). If AI in the form of large language models communicates so much like humans, isn't it also an agent (i.e., an acting entity) that can interact with human agents in a way that lets new knowledge emerge? Such an interaction would encompass true collaboration, that is, an interaction in which the actors involved share processes, goals, and outcomes (cf. Jeong et al., 2017)—even though the AI has no conscious mind of its own. What does this interplay between people and AI look like and how can it be described? How can the potentially complementary capabilities of humans and AI be leveraged (Dellermann et al., 2019)?

We want to contribute to this discussion from a CSCL and Learning Sciences perspective. We aim to stimulate the CSCL community to harvest decades of CSCL research and apply its insights to conversational AI tools. Which findings from CSCL research are applicable, what are potential pitfalls, and what are new research questions CSCL will have to deal with in the future? Our reflections mainly borrow from a cognitive-constructivist perspective, but selectively also touch socio-cultural, technological, and ethical considerations.

Writing and co-construction of knowledge

Key CSCL approaches consider writing not just as a product of a knowledgeable writer, but as a means for learning and knowledge communication. Scardamalia and Bereiter (1987) distinguish *knowledge telling* from *knowledge transforming*. After receiving a writing assignment, novice writers engage in knowledge telling in that they search their long-term memory for content that fits the appropriate genre and then start to write what they know in a relatively linear way. The writing of experts tends to encompass knowledge transforming, which is more complex. Here, the assignment creates two problem spaces, one regarding content, the other regarding rhetoric. While writing, writers repeatedly analyze the problem by considering content and rhetoric, which leads to continuous revisions and updated memory searches. This may not only create a coherent text, but also new knowledge in the individual writers. It transforms their knowledge into complex relational argumentation structures and can eventually induce conceptual change (Andriessen et al., 2013; Kimmerle et al., 2021). Research on self-explanation shows that the development of argument structures can be supported by prompts that enable learners to use typical components of argumentation, such as statements of theory and evidence, statements of alternative theories, and rebuttals and counterarguments against the original theory (Schworm & Renkl, 2007; Toulmin, 1958). In CSCL environments, scripts can be implemented, which include such prompts (Stegmann et al., 2007).

In their *knowledge building* approach, Scardamalia and Bereiter (2006) use these argumentation processes to support knowledge development and increase understanding in groups of learners. Learners exchange individual observations, hypotheses, or explanations for a given problem, such as the issue of how planes can fly. Prompts help them to classify their contributions and create an information structure in which the group proceeds to build knowledge by fading out hypotheses that are not well supported and strengthening those that have more validity. Here argumentation is not just a means for individual learning, but for collaborative knowledge progression, in which the whole group reaches improved understanding over time. Like other CSCL theories, knowledge building stresses that knowledge construction does not take place only in individual learners' minds. Stahl (2006) refers to similar knowledge progression in the *group cognition* approach. Through the interaction of its members, small groups can reach new understandings. Signs for such emergent processes are instances when people take up ideas from others (Suthers et al., 2010) or stimulate new and shared insights, for instance, by using analogies (Roschelle & Teasley, 1995). Going a step further, Trausan-Matu (2009) considers learning and understanding as a social and interindividual process in which multiple voices come into contact. Utterances of different contributors produce resonance and create a polyphony where "different voices singing variously on a single theme" (Bakhtin, 1986, p. 42; see also Koschmann, 1999). Following this idea, Trausan-Matu et al. (2014) opted for learning settings where this polyphony is supported through divergent and convergent phases of knowledge construction. This can be particularly productive when the participants begin to pull in the same direction in terms of argumentation.

The *co-evolution model* describes knowledge construction as a structural coupling of meaning-based cognitive and social systems (Cress & Kimmerle, 2008; Kimmerle et al., 2015). An individual's cognitions constitute an autopoietic cognitive system (Maturana & Varela, 1991; see also Varela et al., 1974): Cognition and knowledge-related processes stimulate further cognitions, perceptions, and interpretations. Social systems are formed by communication, in which utterances lead to further utterances, also constituting autopoietic

systems (Luhmann, 1986). Cognition and communication each have their own logic, but both use language as a medium and can influence each other. Cognitions (within cognitive systems) stimulate utterances (among people), and utterances, in turn, stimulate cognitions. This can be particularly the case when cognitions and utterances do not simply match, that is, when expectations are not entirely fulfilled but slightly violated. In this case, cognitive and social systems can *irritate* each other. This corresponds to the system-theoretical view of what Piaget (1977) calls *socio-cognitive conflicts* (see also Mugny & Doise, 1978). These mutual irritations may lead to individual learning and collective knowledge construction. Irritations and productive friction are processes that trigger knowledge development (Holtz et al., 2018).

Implications for collaborative learning with generative AI tools

To what extent do the considerations presented so far apply to communication and knowledge construction with generative AI models? How is cognition distributed when one of the interacting agents acts like a human individual but is a non-conscious entity (for an elaborate approach on distributed cognition in terms of information processing that is dispersed across humans, technologies, and their environment see Hollan et al., 2000; Hutchins, 1995)? Is the polyphony of different voices valid when some of these voices are not even human? Is it even communication if a language-producing system makes contributions but is not a cognitive system?

The preceding overview of some key insights from CSCL research raises several issues that need to be considered when reflecting on ChatGPT and similar tools for knowledge construction purposes. We argue that processes, which in previous research referred to human–human interaction exclusively, to some extent can be translated over to human–AI interaction: Based on their own knowledge and needs for information, humans phrase questions for an AI agent. They aim to use (supposedly) appropriate prompts that stimulate output from the agent to enter into a conversation. But since AI agents do not represent meaning-based cognitive systems, they are not meaning-constituting partners; they simply rely on word patterns and frequencies.

Thus, generative AI models can be used as answering tools that provide content to user questions, but they are different from search engines that deliver concrete references to links, texts, or pictures from the internet, which users can scrutinize for reliability and validity. Instead, generative AI tools are agents with vast amounts of associations, which construct new output for each interaction. Users' questions and statements trigger this text production. Unlike humans, AI agents do not conceptually understand this input or derive meaning. They provide answers to virtually every prompt, but validity is not reliable. With these strengths and weaknesses, they can indeed serve as interaction partners. This partnership, however, requires that users not only ask questions and receive information from an AI agent, but must also aim for a dialog in which both partners exchange information and stimulate each other's contributions and reflections. These considerations are in line with the ICAP framework (Chi & Wylie, 2014) that describes different modes of learners' engagement. This framework indicates that learners who are increasingly engaged with learning tools or material, move from a passive to an active, then constructive, and finally an interactive mode, accompanied by a rise of learning activities. The transferability and application of these modes to dealing with

generative AI should be relatively straightforward as experienced users of generative AI may reach an increasingly interactive mode over time.

For such a dialog between human users and AI agents, users must be aware that communication and text creation not only refer to content knowledge, but also to rhetorical knowledge. Writers have concepts about different genres that affect their writing. It makes a difference whether the goal of a discourse is drafting an argumentative essay or an election speech. Both differ in the kind of arguments being used, the choice of language, the complexity of sentences, or the length of text. However, not only humans possess and make use of rhetorical knowledge, but AI tools possess rhetorical capabilities as well. If asked for different genres, they also produce diverse types of text. To support joint processes of knowledge transforming (not just knowledge telling) optimally, users need to be able to make use of translations between a content problem space and a rhetorical problem space. Users need prompts for each of the problem spaces, for example, “Imagine you are an expert and have to explain to 10-year-old students how planes can fly” for the rhetorical problem space, or “Are there any facts that speak against this explanation of how planes fly?” for the content problem space.

Previous research of argumentation and knowledge building may provide a basis for identifying optimal prompts that support human-AI communication. Conceptual and empirical research on prompting processes in human-AI communication is urgently needed. In particular, a better understanding is required of how prompts for content and prompts for rhetorical issues influence the AI outcomes and the resulting knowledge construction. The CSCL community needs to develop procedures to convert the prevalent knowledge telling activities to knowledge transforming, in which both partners take up information from each other and reciprocally link and integrate information. For this purpose, ethnomethodological approaches (Garfinkel, 1967), as they have been applied for observing knowledge construction in human groups (see Stahl, 2012), could be further developed, and adjusted to human-AI interaction. This approach could look at individual or multiple users in their actual interactions with ChatGPT, embedded in their everyday lives, to understand how they accomplish these interactions. The assessment of the situational practices observed in this process could be complemented, for example, by interviews and questionnaires to provide a more complete understanding of these practices.

For optimal interactions, humans need awareness of their AI partner: What was the text basis used to train the tool? What kind of data and information does the AI rely on, how does it create its output? Research has described the relevance of group awareness (the perception of certain characteristics of a group and its members) for interaction among humans (e.g., Engelmann et al., 2009). In this new format of human-AI interactions, the awareness concept needs to be refined to allow for more asymmetrical relationships. Humans may know about an AI’s competencies and shortcomings in advance as part of their digital literacy. But during collaboration there may also be indicators that both partners must identify and interpret. It is an open question whether just the human partner or also the AI partner can create awareness of the respective partner—depending in part on whether one assumes that awareness requires consciousness. Research needs to consider how communication should unfold to allow the AI partner to develop a kind of *theory of mind* (Holterman & van Deemter, 2023; see also Trott et al., 2023), meaning that it would be able to recognize the mental state of its interaction partners (including thoughts, feelings, expectations, and motivation), allowing for perspective taking. This would enable communication that is more than knowledge telling and exchange of information to move in the direction of knowledge transforming.

As the AI partner simply relies on associations and does not derive meaning, the question arises whether knowledge construction solely occurs with respect to human agents while the AI just provides new information and induces irritations, or whether knowledge construction happens *between* both agents. While previous research used ethnographical methods to describe moments of shared understanding among human agents, future research should adapt these methods to human-AI collaborations. Can instances be identified where an AI, with the help of human partners, comes to new insights? Can shared knowledge construction be induced? Or is human-AI collaboration still a situation where both partners stimulate and irritate each other in such a way that knowledge development happens only within but not between agents?

Addressing these far-reaching questions requires the full methodological breadth that CSCL and the Learning Sciences have to offer. Beyond ethnographic and ethnomethodological research and discourse processing, experimental studies with human-AI dyads or groups could compare the efficacy of particular prompts and communication strategies for human learning and knowledge construction. In doing so, concrete borrowings can also be made from human-computer interaction research. Moreover, cognitive modeling can simulate people's thinking and their mental information processing while interacting with generative AI in increasingly sophisticated computerized models. Large-scale studies with simulated human participants that induce particular prompts could also test how AI partners react, and how and to what extent they are able to adapt to people's knowledge.

This vast array of potentially helpful methods can be cast into concrete research programs. As a first step, people (individuals but also small groups of people) should be systematically and extensively observed in their actual engagement with language-based AI tools. In the context of such exploratory studies, it can be observed and analyzed how people interact with ChatGPT and other tools, what they believe these tools can and cannot do. For these examinations, conversation analysis can be used as well as targeted analysis of uptake events (Suthers et al., 2010). In addition, participants can be asked to think aloud while collaboratively writing with the tools. In-depth interviews with participants could be a good way to get them to take a stand on their behavior and reflect on it. Finally, users could be examined to determine whether they actually developed an understanding (e.g., on how airplanes can fly) or whether completely new ideas were developed in the interaction.

Subsequently, experimental research programs are needed. Here, the variation of a large number of independent variables is conceivable. For example, it could be varied what people know or think they know about the content being addressed. In addition, their knowledge about and expectations of generative AI tools could be manipulated. Varying the (perceived) trustworthiness is also feasible and it can then be recorded to what extent persons recognize this trustworthiness. Moreover, experimental studies could vary the alleged capabilities (e.g., in terms of supposed consciousness) of the tools and capture in what way humans would assign an assumed theory of mind or other typically human abilities to the AI tool and how this in turn affects human-AI collaboration and knowledge construction.

Overall, there are many unanswered questions, but they would certainly be worth answering. CSCL research has the theoretical equipment and the methodological instruments to pursue these important questions. Only then can CSCL research and practice succeed in keeping pace with the currently rapid technological developments and be prepared for a future in which such tools are an integral part of everyday life.

Funding Open Access funding enabled and organized by Projekt DEAL.

Declarations

Competing interests There are no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Andriessen, J., Baker, M., & Suthers, D. D. (Eds.) (2013). *Arguing to learn: Confronting cognitions in computer-supported collaborative learning environments* (Vol. 1). Springer Science & Business Media.
- Bakhtin, M. (1986). *Speech genres and other late essays*. University of Texas Press.
- Chi, M. T., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, 49(4), 219–243.
- Cooper, G. (2023). Examining science education in chatgpt: An exploratory study of generative artificial intelligence. *Journal of Science Education and Technology*, 32(3), 444–452.
- Cress, U., & Kimmerle, J. (2008). A systemic and cognitive view on collaborative knowledge building with wikis. *International Journal of Computer-Supported Collaborative Learning*, 3(2), 105–122.
- Dellermann, D., Ebel, P., Söllner, M., & Leimeister, J. M. (2019). Hybrid intelligence. *Business & Information Systems Engineering*, 61, 637–643.
- Engelmann, T., Dehler, J., Bodemer, D., & Buder, J. (2009). Knowledge awareness in CSCL: A psychological perspective. *Computers in Human Behavior*, 25(4), 949–960.
- Garfinkel, H. (1967). *Studies in ethnomethodology*. Prentice-Hall.
- Hollan, J., Hutchins, E., & Kirsh, D. (2000). Distributed cognition: Toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction*, 7(2), 174–196.
- Holterman, B., & van Deemter, K. (2023). Does ChatGPT have Theory of Mind? arXiv preprint arXiv:2305.14020.
- Holtz, P., Kimmerle, J., & Cress, U. (2018). Using big data techniques for measuring productive friction in mass collaboration online environments. *International Journal of Computer-Supported Collaborative Learning*, 13, 439–456.
- Hutchins, E. (1995). *Cognition in the wild*. MIT Press.
- Jeong, H., Cress, U., Moskaliuk, J., & Kimmerle, J. (2017). Joint interactions in large online knowledge communities: The A3C framework. *International Journal of Computer-Supported Collaborative Learning*, 12, 133–151.
- Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274.
- Kimmerle, J., Fischer, F., & Cress, U. (2021). Argumentation and knowledge construction. In U. Cress, C. Rosé, A. F. Wise, & J. Oshima (Eds.), *International handbook of computer-supported collaborative learning* (pp. 183–198). Springer International Publishing.
- Kimmerle, J., Moskaliuk, J., Oeberst, A., & Cress, U. (2015). Learning and collective knowledge construction with social media: A process-oriented perspective. *Educational Psychologist*, 50, 120–137.
- Koschmann, T. D. (1999). Toward a dialogic theory of learning: Bakhtin's contribution to understanding learning in settings of collaboration. In C. M. Hoadley, & J. Roschelle (Eds.), *Proceedings of the computer support for collaborative learning (CSCL) 1999 conference*. International Society of the Learning Sciences.
- Luhmann, N. (1986). The autopoiesis of social systems. In F. Geyer & J. Van der Zouwen (Eds.), *Sociocybernetic paradoxes* (pp. 172–192). Sage.

- Maturana, H. R., & Varela, F. J. (1991). *Autopoiesis and cognition: The realization of the living* (Vol. 42). Springer Science & Business Media.
- Mugny, G., & Doise, W. (1978). Socio-cognitive conflict and structure of individual and collective performances. *European Journal of Social Psychology*, 8, 181–192.
- Oeberst, A., Kimmerle, J., & Cress, U. (2016). What is knowledge? who creates it? who possesses it? the need for novel answers to old questions. In U. Cress, J. Moskaliuk, & H. Jeong (Eds.), *Mass collaboration and education* (pp. 105–124). Springer International Publishing.
- Piaget, J. (1977). *The development of thought: Equilibration of cognitive structures*. Viking.
- Roschelle, J., & Teasley, S. D. (1995). The construction of shared knowledge in collaborative problem solving. In C. O'Malley (Ed.), *Computer supported collaborative learning* (pp. 69–97). Springer Berlin Heidelberg.
- Scardamalia, M., & Bereiter, C. (1987). Knowledge telling and knowledge transforming in written composition. In S. Rosenberg (Ed.), *Advances in applied psycholinguistics, Vol. 1. Disorders of first-language development; Vol. 2. Reading, writing, and language learning* (pp. 142–175). Cambridge University Press.
- Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. In K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences* (pp. 97–118). Cambridge University Press.
- Schworm, S., & Renkl, A. (2007). Learning argumentation skills through the use of prompts for self-explaining examples. *Journal of Educational Psychology*, 99(2), 285–296.
- Stahl, G. (2006). *Group cognition: Computer support for building collaborative knowledge*. MIT Press.
- Stahl, G. (2012). Ethnomethodologically informed. *International Journal of Computer-Supported Collaborative Learning*, 7, 1–10.
- Stegmann, K., Weinberger, A., & Fischer, F. (2007). Facilitating argumentative knowledge construction with computer-supported collaboration scripts. *International Journal of Computer-Supported Collaborative Learning*, 2, 421–447.
- Suthers, D. D., Dwyer, N., Medina, R., & Vatrpu, R. (2010). A framework for conceptualizing, representing, and analyzing distributed interaction. *International Journal of Computer-Supported Collaborative Learning*, 5, 5–42.
- Toulmin, S. (1958). *The uses of argument*. Cambridge University Press.
- Trausan-Matu, S. (2009). The polyphonic model of hybrid and collaborative learning. In F. L. Wang, J. Fong, & R. Kwan (Eds.), *Handbook of research on hybrid learning models: Advanced tools, technologies, and applications* (pp. 466–486). Information Science Reference.
- Trausan-Matu, S., Dascalu, M., & Rebedea, T. (2014). PolyCAFe—automatic support for the polyphonic analysis of CSCL chats. *International Journal of Computer-Supported Collaborative Learning*, 9(2), 127–156.
- Trott, S., Jones, C., Chang, T., Michaelov, J., & Bergen, B. (2023). Do large language models know what humans know? *Cognitive Science*, 47, e13309.
- Varela, F. J., Maturana, H. R., & Uribe, R. (1974). Autopoiesis: The organization of living systems, its characterization and a model. *Bio Systems*, 5, 187–196.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.