Paper ID #44690

Empowering computer-supported collaborative learning with ChatGPT: investigating effects on student interactions

Mr. Han Kyul Kim, University of Southern California

Han Kyul Kim is a Ph.D. student in the Department of Industrial and Systems Engineering at the University of Southern California. He earned his M.S. in Industrial Engineering from Seoul National University and holds B.S. degrees in Industrial & Systems Engineering and Business and Technology Management from KAIST. Previously, he worked as a data scientist and software engineer at Deloitte Consulting, Samsung Electronics, and Seoul National University Hospital. His research interests include machine learning, natural language processing, and learning analytics.

Aleyeh Roknaldin, USC Viterbi School of Engineering

Aleyeh Roknaldin is earning her Ph.D. in industrial and systems engineering from the Daniel J. Epstein Department of Industrial and Systems Engineering at the University of Southern California (USC). Aleyeh earned an M.S. in Engineering Management in 2022 from USC and her B.S. in Chemical Engineering from the University of California, Davis in 2020. Her research interests include learning analytics to investigate how students learn from one another in computer-supported collaborative learning environments and how students interact with generative artificial intelligence in learning contexts.

Shriniwas Prakash Nayak, University of Southern California
Mr. Xiaoci Zhang, University of Southern California

received his Bachelor's in Industrial Engineering and a minor in Mathematics from the Pennsylvania State University in 2021. Currently, he is pursuing his Master's in Industrial and System Engineering at the University of Southern California. His research interests include machine learning, data analysis, and mathematical optimizations.

Marlon Twyman, University of Southern California
Angel Hsing-Chi Hwang, Cornell University

Human-AI Interaction Research

Dr. Stephen Lu, University of Southern California

Stephen Lu is the David Packard Chair in Manufacturing Engineering at University of Southern California. His current professional interests include design thinking, collaborative engineering, technological innovation, and education reform. He has over 330

Empowering Computer-Supported Collaborative Learning With ChatGPT: Investigating Effects on Student Interactions

Han Kyul Kim[†*], Aleyeh Roknaldin[†], Shriniwas Nayak[†], Xiaoci Zhang[†],
Marlon Twyman[†], Angel Hsing-Chi Hwang[‡], Stephen Lu[†]
hankyulk@usc.edu, roknaldi@usc.edu, spnayak@usc.edu,
xiaocizh@usc.edu, marlontw@usc.edu, hh695@cornell.edu, sclu@usc.edu
University of Southern California[†], Cornell University[‡]

Abstract

The surging popularity of generative AI, especially ChatGPT, has evoked both enthusiasm and caution within the education community. To effectively harness the full potential of ChatGPT in educational contexts, it is crucial to thoroughly analyze its impact and suitability for different educational purposes. The current paper aims to contribute to this understanding by investigating the impact of ChatGPT on student interactions in a computer-supported collaborative learning (CSCL) environment when deployed as a conversation agent. Furthermore, we present a concise overview of our iBot system, which seamlessly incorporates ChatGPT's capabilities into CSCL, providing technical guidelines for potential future adoptions in similar educational contexts. Through extensive statistical analysis, we numerically examine how including ChatGPT affects the dynamics of student interactions. Our findings reveal a significant shift in the patterns of student interactions, with a notable increase in active learning interactions occurring predominantly between individual students and ChatGPT, even in the presence of other students. While recognizing that the observed changes may not be universal in every educational context, our quantitative analysis emphasizes the need to consider the potential impact of integrating ChatGPT into collaborative learning scenarios. By shedding light on the effects of ChatGPT in CSCL, this study offers valuable insights for educators and researchers to better understand its role as a conversation agent in collaborative learning settings. It encourages thoughtful consideration of the appropriate use case of ChatGPT and emphasizes the importance of further exploring its potential to enhance collaborative learning experiences.

1. Introduction

Rooted in Vygotsky's zone of proximal development [1], computer-supported collaborative learning (CSCL) posits that students' cognitive development is promoted through shared understanding and meaning-making from their interactions during group activities [2]. Given this inherent nature, active engagement and discussion within student groups play a pivotal role in the success of CSCL [3]. With recent advances in machine learning, a plethora of data-driven methods have emerged to facilitate this essential aspect [4]. One particularly intriguing application of these AI techniques is the use of conversation agents in the form of chatbots that serve as interactive tools to facilitate student engagement [5, 6]. Previous studies applying chatbots in various educational contexts have consistently reported improved learning outcomes and increased interest among students [7].

However, existing research exploring the adoption of chatbots in CSCL faces two critical practical issues. Firstly, generating responses from chatbots based on students' requests involves pre-specifying response patterns, requiring significant manual configuration prior to their actual deployment. Secondly, chatbots generate responses from a knowledge base that must be populated before deployment. These challenges demand additional effort from instructors, adding to the already instructor-dependent nature of CSCL activities [8].

Generative AI such as ChatGPT brings a paradigm shift in chatbot design for CSCL. By leveraging its language generation capabilities, it mitigates the need for manual response scripting or extensive configuration prior to deployment. Furthermore, it empowers students to access a vast knowledge base that was originally utilized to pre-train its large language model. However, concerns related to hallucination [9, 10] and academic integrity [11, 12] have been raised, casting doubt on the applicability of ChatGPT in education. Given the divided views on the application of ChatGPT in education, this paper seeks to introduce and explore a potential use case within the CSCL context. Specifically, we showcase its role as a conversation agent during group brainstorming sessions. Through analyzing students' chat data from these sessions, we conducted a statistical analysis to understand the impact of ChatGPT on student collaborative interactions that are crucial for the success of CSCL.

2. Background

2-1. Conversation agents in brainstorming

Previously, various forms of conversation agents have been introduced or tested to facilitate brainstorming sessions [13, 14]. For instance, [15] validated the feasibility of using a conversation agent as a live facilitator in brainstorming sessions by comparing its performance against human facilitators. However, their study was conducted in a controlled experimental setting, where researchers observed test subjects through a one-sided mirror and manually inserted data for the conversation agent. This experimental setup limits the practical deployment of the conversation agent in real brainstorming sessions within the context of CSCL.

On the other hand, conversation agents in online brainstorming sessions have shown promise in enhancing student learning [16] and guiding divergent brainstorming sessions [13]. These agents generate appropriate responses for student inputs using machine learning-based information retrieval algorithms. A key common limitation of these agents is that they rely on retrieving responses from pre-defined ontologies, requiring substantial manual effort from instructors to either generate or populate these ontologies before deploying the agents. The intensity of technical preparation underscores the need for conversation agents, such as ChatGPT, with language-generation capabilities to alleviate the burden on instructors and enhance their applicability in educational settings. However, before fully adopting ChatGPT in actual classrooms, its implications on students need to be carefully examined. Our paper aims to be among the first to investigate the effectiveness and practicality of ChatGPT as a brainstorming peer agent in a CSCL environment, shedding light on its potential as an innovative tool for collaborative learning.

2-2. Conversation agents in CSCL

The existing research on conversation agents in CSCL has primarily relied on two different approaches. One approach involves rule-based systems, where the conversation agents generate responses based on pre-defined dialogue specifications or concepts corresponding to each student's exercise or behavior. Such examples include agents that require instructors to manually configure the agents to generate or map appropriate responses for different scenarios [17, 18, 19]. To reduce the amount of effort, an authoring toolkit was developed to streamline the pre-configuration process when deploying chatbots in CSCL [20]. Conversely, conversation agents can also rely on pre-populated knowledge bases or ontologies to facilitate correct response generation [21, 22]. These agents use traditional natural language processing techniques, such as Latent Dirichlet Allocation (LDA) [23], to automatically detect topics from students' questions and match suitable answers. However, LDA, by itself, lacks language generation capabilities, limiting its role to that of an automated text pre-processing module within chatbots.

Despite the active interest of the existing CSCL community in deploying conversation agents, the role of a large language model with language generation capabilities as a conversation agent remains largely unexplored. While significant technical advancements have been made in natural language processing, their impact on student engagement during CSCL activities requires thorough investigation to demystify their applicability and establish guidelines for leveraging their massive potential in the educational context.

3. Methods

3-1. CSCL context

To assess the impact of ChatGPT on student interactions, we analyzed student message data collected from a graduate-level product engineering course offered at a university in California during the Spring of 2023. During the final two weeks of the semester, we formed 12 cohorts, each consisting of 4 to 5 students. Out of these 12 cohorts, ChatGPT was integrated as a conversation agent in 6 cohorts, which we will refer to as `Type A` throughout this paper. Conversely, the remaining 6 cohorts that did not include ChatGPT were designated as `Type B`. Each week, the cohorts were tasked with applying the product engineering knowledge they had acquired throughout the semester to engage in discussions and develop multiple creative product design ideas that addressed specific customer requirements. These discussions took place asynchronously online within the group chat interface in Microsoft Teams. At the end of each week, each cohort was required to submit one document listing their product design ideas for review.

To facilitate student access to ChatGPT and streamline data collection for further analysis, we developed the iBot system, which functions as a direct interface to ChatGPT and serves as a

Table 1: Summary of student messages in our CSCL context

	`Type A`				`Type B`		
Cohort number	**Total number of students**	**Total number of student messages**	**Avg number of words per student message**	**Total number of ChatGPT messages**	**Total number of students**	**Total number of student messages**	**Avg number of words per student message**
Cohort #1	5	35	20.371	13	4	25	17.56
Cohort #2	4	83	13.41	28	5	85	14.729
Cohort #3	5	92	31.011	10	4	30	11.767
Cohort #4	5	115	23.191	34	5	61	15.18
Cohort #5	5	38	26.789	18	5	21	29.691
Cohort #6	5	46	24.0	28	5	71	19.239

conversation agent exclusively in `Type A` cohorts. Any student within `Type A` could interact with ChatGPT simply by using the `@iBot` tag within their message, thereby invoking a response from ChatGPT. Instead of directly requesting product design ideas from ChatGPT that satisfy each week's customer requirement, students were instructed to discuss and explore different aspects of their designs with ChatGPT, thereby encouraging creative thinking beyond conventional boundaries. However, no additional specific guidelines on prompting strategies were provided. Conversely, students in `Type B` were assumed to have followed the provided instructions and refrained from accessing ChatGPT, even outside our system. The comprehensive statistics, including the number of student messages and their average length per cohort, are summarized in Table 1.

3-2. System overview

As a reference for those interested in utilizing ChatGPT as a conversation agent in CSCL settings, we provide a brief summary of how our iBot system is set up to enable seamless interactions between students and ChatGPT (Figure 1). When a student's message contains the `@iBot` tag, the iBot would collect all the previous chats in the group chat to serve as contextual information for generating a prompt for ChatGPT. Next, the chats are transformed into an API request. Within this API request, our iBot system distinguishes between the previous messages in the group chat from students and those responses generated by ChatGPT. Subsequently, this API request from the iBot system is sent to OpenAI API, which enables ChatGPT (gpt-3.5-turbo with the temperature parameter set to 0.5) to generate its responses. The obtained response from ChatGPT is posted as a new message in the group chat for students to proceed with their brainstorming sessions.

3-3. Experiment setting

3-3-1. Collaborative Learning Conversation Skills Taxonomy

To conduct a detailed analysis of student interactions in our dataset, we employ the Collaborative Learning Model proposed by [24]. This model introduces the Collaborative Learning Conversation Skills Taxonomy (CST), which classifies the types of interaction-related skills commonly used by students during CSCL [25, 26]. The taxonomy categorizes student

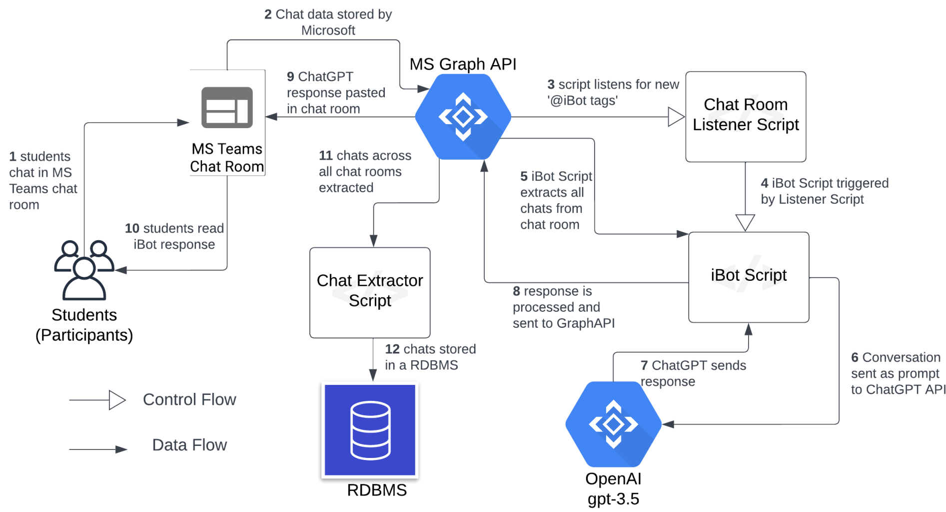

Figure 1: Summary of iBot system architecture

interactions into three main skills: Conversation, Active Learning, and Creative Conflict, each comprising its own distinct set of subskills.

CST was employed to annotate the student messages in our dataset. However, due to a limited number of messages exhibiting creative conflict, we labeled messages solely based on whether the student demonstrated conversation or active learning skills. In Table 2, we present a comprehensive breakdown of the main skills of CST used in this study, along with their respective subskills. Conversation and its subskills primarily focus on planning or executing ideas that have already been discussed. On the contrary, active learning represents student interactions wherein they discuss new ideas, provide feedback to each other, encourage diverse opinions, and ask questions. Existing literature has highlighted that the presence of predominantly active learning messages is a key factor in student learning during collaborative activities [27, 28].

Table 2: Collaborative Learning Conversation Skills Taxonomy as defined in [24]

Main Skill	Sub Skill	Definition
Active Learning	Request	Ask for help/advice in solving the problem or in understanding a team-mates comment.
	Inform	Direct or advance the conversation by providing information or advice.
	Motivate	Provide positive feedback and reinforcement.
Conversation	Task	Shift the current focus of the group to a new subtask or tool.
	Maintenance	Support group cohesion and peer involvement.
	Acknowledge	Inform peers that you read and/or appreciate their comments. Answer yes/no questions.

3-3-2. Data annotation

To measure the impact of ChatGPT on student interactions in CSCL, two experts familiar with CST manually annotated each message according to CST. As used in previous CSCL research [29, 30], we consider the complete message posted by a student as the unit of analysis, defining it as a single student interaction during a brainstorming session. Each message was categorized with one of the main skills in CST, and the messages exhibiting active learning in students were further annotated with subskills. To ensure the consistency and reliability of the annotations, we calculated Cohen's kappa [31] between the two annotators. The Cohen's kappa between the annotators for the main skills of CST was calculated at 93.2%, while the Cohen's kappa for the subskills within active learning was 94.55%. As indicated by [32], these levels of Cohen's kappa show almost perfect agreement between the two annotators, establishing the robustness of the annotated labels used for our subsequent analyses.

4. Results

To investigate the impact of ChatGPT on different types of student interactions, our statistical analysis explores the differences in the level of active learning and conversational interactions between cohorts with and without ChatGPT. To account for potential variations in the number of students participating in each cohort, we measure the impact of ChatGPT on the ratio of each interaction type relative to the total number of interactions within each cohort. This approach allows us to mitigate the potential correlation between cohort size and the number of interactions, enabling us to evaluate the effect of ChatGPT at a group level in a more robust manner.

For all our statistical analyses, we employed repeated ANOVA. In this model, the inclusion of ChatGPT in a cohort and the week of the brainstorming session were set as two independent covariates. Note that the week was included to ensure that we control for variations that may occur within a cohort across the two weeks. Once the F-test yielded statistical significance, we proceeded with the Tukey method [33] for subsequent multiple pairwise comparisons.

4-1. Overall interaction

When comparing the overall interactions between `Type A` and `Type B` cohorts, we did not observe any statistically significant differences in the level of active learning or conversation in student interactions. While the impact of ChatGPT was not evident across all student interactions, this result does not preclude the possibility of statistical significance within different types of student interactions. As ChatGPT was exclusively included in `Type A`, we can further classify the student interactions within `Type A` into two distinct types: student-student interactions and student-ChatGPT interactions. Therefore, we further analyze the level of active learning and conversation within these two types of student interactions.

4-2. Student-student interaction

In order to trigger a response from ChatGPT, students were required to include a `@iBot` tag in their messages. Consequently, any message lacking this tag was considered an interaction between students. For `Type B`, where ChatGPT was not available as a conversation agent, all interactions occurring in them were categorized as student-student interactions. To evaluate the impact of ChatGPT, we compared the levels of active learning and conversation in student-student interactions within these two types of cohorts. We observed a statistically significant variation only in the level of active learning during student-student interactions across different cohorts (Table 3, p-value of 0.038). Further investigation using the Tukey method (Table 4) revealed that this difference is primarily attributed to the presence of ChatGPT as a conversation agent. Specifically, `Type B` exhibited a 22.6 percentage point higher ratio of active learning in student-student interactions. Combining the results from the previous subsection, our findings suggest that although the overall ratio of active learning interactions did not significantly differ between the two types of cohorts, the level of active learning within student-student interactions was notably higher in the cohorts without ChatGPT. This outcome implies that even though there was no major change in the overall student interaction, the internal dynamics of student interactions, particularly the level of active learning within this specific type of student interaction, were significantly altered by introducing ChatGPT as a conversation agent.

Table 3: ANOVA results on the level of active learning in student-student interactions

Source of Variation	SS	df	F
ChatGPT	0.307	1	**4.909***
Week	0.031	1	0.501
ChatGPT × Week	0.080	1	1.285
Residual	1.253	20	

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4: Tukey's method for measuring the impact of ChatGPT on the level of active learning in student-student interaction

Comparison	Difference	Adj. p-value
With ChatGPT vs. Without ChatGPT	0.226	**0.037**

4-3. Student-ChatGPT interaction

When analyzing the student-ChatGPT interactions, we exclusively focus on cohorts with ChatGPT, comparing the levels of active learning and conversation between messages directed towards ChatGPT and those addressed to other students within `Type A`. To account for variations in the number of active learning or conversational messages across these cohorts, we adopt a different approach for calculating the ratios compared to the previous subsections. Instead of comparing the ratio of active learning (or conversational) interactions directed towards students or ChatGPT to the total number of interactions in a cohort, we measure it relative to the total number of active learning (or conversational) messages within that cohort.

Table 5: ANOVA on the level of active learning in student-ChatGPT interactions

Source of Variation	SS	df	F
To ChatGPT	1.722	1	**24.765*****
Week	0.000	1	0.000
ChatGPT $\times$ Week	0.008	1	0.11
Residual	1.391	20	

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: ANOVA on the level of conversation in student-chatGPT interactions

Source of Variation	SS	df	F
To ChatGPT	0.660	1	**4.754***
Week	0.000	1	0.000
ChatGPT $\times$ Week	0.069	1	0.498
Residual	2.775	20	

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The statistical results presented in Tables 5 to 7 reveal significant differences in the levels of both active learning and conversational interactions between student interactions directed towards ChatGPT and those addressed to other students. The results from Tukey's methods in Table 7 provide further insight, indicating that incorporating ChatGPT as a conversation agent in a cohort leads to a significant increase in the level of active learning during student-ChatGPT interactions. Conversely, most of the student-student interactions in this case are characterized by conversational exchanges.

Upon comparing the sheer number of student interactions directed toward ChatGPT and those directed toward other students within `Type A`, we observed a significantly higher number of messages directed toward students. Throughout the two-week period, `Type A` exhibited an average of 18.83 student-student interactions per cohort, while there were 7 student-ChatGPT interactions per cohort. However, our statistical analysis of student-ChatGPT interactions indicates that this significant difference does not necessarily arise from collaborative learning occurring between the students. Rather, it suggests that the actual active learning process within the cohort is becoming more bilateral between students and ChatGPT, while more task or non-learning-related interactions tend to dominate among student-student interactions. These

Table 7: Tukey's method for measuring the impact of ChatGPT on the level of active learning and conversation in student-ChatGPT interaction

Comparison	CST type	Difference	Adj. p-value
To ChatGPT vs.	active learning	-0.5357	**0.000**
To students	conversation	0.3316	**0.034**

findings highlight the distinctive nature of interactions in the presence of ChatGPT and underscore its potential impact on the dynamics of CSCL settings.

5. Discussion

As indicated in [34], the success of CSCL relies heavily on instructors' continuous supervision and feedback to encourage student interactions. Given this context, it is essential to recognize that incorporating ChatGPT alone does not replace instructors' active involvement in orchestrating and monitoring CSCL. Furthermore, our statistical analyses investigating the impact of ChatGPT on different types of student interactions revealed noteworthy behavioral changes among students based on whether ChatGPT is included as a conversation agent in a cohort or not. Specifically, we observed that including ChatGPT diminishes the level of active learning interactions between students, which are pivotal for successful collaborative learning. This decrease was primarily attributed to a relatively higher proportion of learning-related interactions occurring predominantly between individual students and ChatGPT.

When comparing the subskills within active learning present in student-student interactions in `Type A` and `Type B` to those in student-ChatGPT interactions in `Type A`, a significant difference becomes apparent (Figure 2). The majority of active learning interactions with ChatGPT exhibit the sub-skill of "Request." This result suggests that students primarily use ChatGPT as a "search engine" to directly request information related to the given task during each week, rather than engaging in a deeper exploration of different ideas with ChatGPT's assistance. On the other hand, most student-student interactions fall into the sub-skill of "Inform," highlighting the aspect of information exchange and sharing between students during brainstorming sessions.

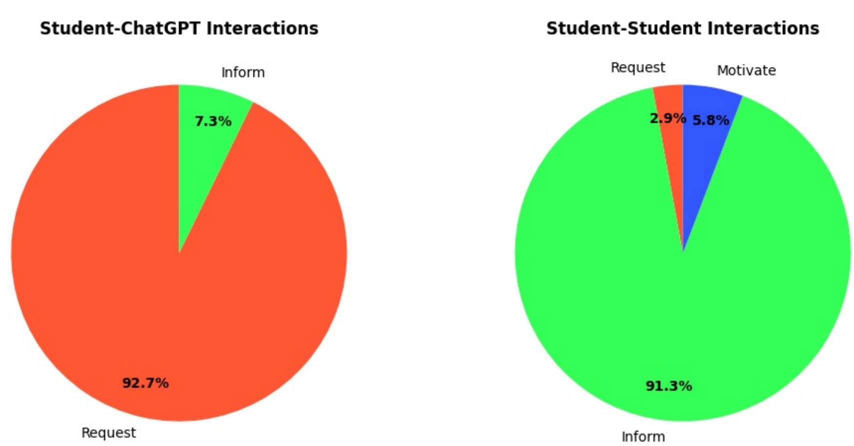

Figure 2: Comparison of active learning subskills in each type of student interaction

Therefore, ChatGPT, acting as a conversation agent in our CSCL context, transforms the intended knowledge co-construction through multilateral interactions between students within a cohort into multiple bilateral interactions. As prior research has emphasized the significance of an equal level of interactions or participation between all members for successful collaborative groups [35], this change in the dynamics of student interactions by incorporating ChatGPT should not be overlooked. This change in student interactions highlights the importance of conducting further systematic research to thoroughly explore the impact of ChatGPT in various educational contexts and identifying its most beneficial applications. As expressed through artifacts at the end of the brainstorming session, we leave the implications of this transformation on students' learning outcomes and creativity to our future work.

6. Conclusion

In this paper, we investigated the impact of ChatGPT in the context of CSCL. Through extensive statistical analyses, we explored how including ChatGPT affects various types of student interactions during CSCL. Our findings reveal that the presence of ChatGPT does not uniformly influence all aspects of student interactions. However, a notable difference was observed in the level of active learning exhibited in student-student interactions between cohorts with and without ChatGPT. This behavioral shift indicates that most of the learning aspects of interaction within cohorts predominantly occurred between individual students and ChatGPT, even in the presence of other students. To foster more effective active learning during CSCL, it is crucial to encourage engagement and knowledge exchange in both student-student and student-ChatGPT interactions. Therefore, gaining a deeper understanding of ChatGPT's role as a conversation agent in CSCL becomes increasingly critical. By elucidating the specific contributions of ChatGPT in collaborative learning settings, we can better harness its potential to enhance student engagement, knowledge sharing, and learning outcomes.

References

[1] L. S. Vygotsky and M. Cole, *Mind in society: Development of higher psychological processes*. Harvard university press, 1978.

[2] D. D. Suthers, "Technology affordances for intersubjective meaning making: A research agenda for cscl," *International Journal of Computer-supported collaborative learning*, vol. 1, pp. 315–337, 2006.

[3] H. Jeong, C. E. Hmelo-Silver, and K. Jo, "Ten years of computer-supported collaborative learning: A meta-analysis of cscl in stem education during 2005–2014," *Educational research review*, vol. 28, p. 100284, 2019.

[4] J. Tang, X. Zhou, X. Wan, and F. Ouyang, "A systematic review of ai applications in computer-supported collaborative learning in stem education," *Artificial Intelligence in STEM Education: The Paradigmatic Shifts in Research, Education, and Technology*, p. 333, 2022.

[5] J. Q. Pérez, T. Daradoumis, and J. M. M. Puig, "Rediscovering the use of chatbots in education: A systematic literature review," *Computer Applications in Engineering Education*, vol. 28, no. 6, pp. 1549–1565, 2020.

[6] C. W. Okonkwo and A. Ade-Ibijola, "Chatbots applications in education: A systematic review," *Computers and Education: Artificial Intelligence*, vol. 2, p. 100033, 2021.

[7] M. A. Kuhail, N. Alturki, S. Alramlawi, and K. Alhejori, "Interacting with educational chatbots: A systematic review," *Education and Information Technologies*, vol. 28, no. 1, pp. 973–1018, 2023.

[8] C. Kaendler, M. Wiedmann, N. Rummel, and H. Spada, “Teacher competencies for the implementation of collaborative learning in the classroom: A framework and research review,” *Educational Psychology Review*, vol. 27, pp. 505–536, 2015.

[9] Z. Ji, N. Lee, R. Frieske, T. Yu, D. Su, Y. Xu, E. Ishii, Y. J. Bang, A. Madotto, and P. Fung, “Survey of hallucination in natural language generation,” *ACM Computing Surveys*, vol. 55, no. 12, pp. 1–38, 2023.

[10] H. Alkaissi and S. I. McFarlane, “Artificial hallucinations in chatgpt: implications in scientific writing,” *Cureus*, vol. 15, no. 2, 2023.

[11] B. McMurtrie, “Teaching: Will chatgpt change the way you teach,” *The Chronicle of Higher Education*, 2023.

[12] J. Rudolph, S. Tan, and S. Tan, “Chatgpt: Bullshit spewer or the end of traditional assessments in higher education?” *Journal of Applied Learning and Teaching*, vol. 6, no. 1, 2023.

[13] C.-H. Wang and T.-Y. Li, “Design of an intelligent agent for stimulating brainstorming,” in *Proceedings of the 2018 10th International Conference on Machine Learning and Computing*, 2018, pp. 305–311.

[14] B. Wieland, J. de Wit, and A. de Rooij, “Electronic brainstorming with a chatbot partner: A good idea due to increased productivity and idea diversity,” *Frontiers in Artificial Intelligence*, vol. 5, p. 880673, 2022.

[15] T. Maier, N. F. Soria Zurita, E. Starkey, D. Spillane, C. McComb, and J. Menold, “Comparing human and cognitive assistant facilitated brainstorming sessions,” *Journal of Engineering Design*, vol. 33, no. 4, pp. 259–283, 2022.

[16] H.-C. Wang, C. P. Rosé, and C.-Y. Chang, “Agent-based dynamic support for learning from collaborative brainstorming in scientific inquiry,” *International Journal of Computer-Supported Collaborative Learning*, vol. 6, pp. 371–395, 2011.

[17] E. Walker, N. Rummel, and K. R. Koedinger, “Designing automated adaptive support to improve student helping behaviors in a peer tutoring activity,” *International Journal of Computer-Supported Collaborative Learning*, vol. 6, pp. 279–306, 2011.

[18] S. Tegos, S. Demetriadis, and A. Karakostas, “Promoting academically productive talk with conversational agent interventions in collaborative learning settings,” *Computers & Education*, vol. 87, pp. 309–325, 2015.

[19] S. Tegos, G. Psathas, T. Tsiatsos, C. Katsanos, A. Karakostas, C. Tsibanis, and S. Demetriadis, “Enriching synchronous collaboration in online courses with configurable conversational agents,” in *Intelligent Tutoring Systems: 16th International Conference, ITS 2020, Athens, Greece, June 8–12, 2020, Proceedings 16*. Springer, 2020, pp. 284–294.

[20] R. Kumar and C. P. Rose, “Architecture for building conversational agents that support collaborative learning,” *IEEE Transactions on Learning Technologies*, vol. 4, no. 1, pp. 21–34, 2010.

[21] F. Clarizia, F. Colace, M. Lombardi, F. Pascale, and D. Santaniello, “Chatbot: An education support system for student,” in *Cyberspace Safety and Security: 10th International Symposium, CSS 2018, Amalfi, Italy, October 29–31, 2018, Proceedings 10*. Springer, 2018, pp. 291–302.

[22] M. Verleger and J. Pembridge, “A pilot study integrating an ai-driven chatbot in an introductory programming course,” in *2018 IEEE frontiers in education conference (FIE)*. IEEE, 2018, pp. 1–4.

[23] D. M. Blei, A. Y. Ng, and M. I. Jordan, “Latent dirichlet allocation,” *Journal of machine Learning research*, vol. 3, no. Jan, pp. 993–1022, 2003.

[24] A. Soller, “Supporting social interaction in an intelligent collaborative learning system,” *International journal of artificial intelligence in education*, vol. 12, no. 1, pp. 40–62, 2001.

[25] U. Rio *et al.*, “Text message categorization of collaborative learning skills in online discussion using support vector machine,” in *2013 International Conference on Computer, Control, Informatics and Its Applications (IC3INA)*. IEEE, 2013, pp. 295–300.

[26] D. W. H. Tan, S. Gottipati, K. J. Shim, and V. Shakararaman, “Integrated discourse analysis & learning skills framework for class conversations,” in *2021 IEEE Frontiers in Education Conference (FIE)*. IEEE, 2021, pp. 1–9.

[27] A. Khan, O. Egbue, B. Palkie, and J. Madden, “Active learning: Engaging students to maximize learning in an online course,” *Electronic Journal of e-learning*, vol. 15, no. 2, pp. pp107–115, 2017.

[28] K. Børte, K. Nesje, and S. Lillejord, “Barriers to student active learning in higher education,” *Teaching in Higher Education*, vol. 28, no. 3, pp. 597–615, 2023.

[29] L. Rourke, T. Anderson, D. R. Garrison, and W. Archer, “Methodological issues in the content analysis of computer conference transcripts,” *International Journal of Artificial Intelligence in Education (IJAIED)*, vol. 12, pp. 8–22, 2001.

[30] B. De Wever, T. Schellens, M. Valcke, and H. Van Keer, “Content analysis schemes to analyze transcripts of online asynchronous discussion groups: A review,” *Computers & education*, vol. 46, no. 1, pp. 6–28, 2006.

[31] J. Cohen, “A coefficient of agreement for nominal scales,” *Educational and psychological measurement*, vol. 20, no. 1, pp. 37–46, 1960.

[32] M. L. McHugh, “Interrater reliability: the kappa statistic,” *Biochemia medica*, vol. 22, no. 3, pp. 276–282, 2012.

[33] J. W. Tukey, “Comparing individual means in the analysis of variance,” *Biometrics*, pp. 99–114, 1949.

[34] A. Van Leeuwen and J. Janssen, “A systematic review of teacher guidance during collaborative learning in primary and secondary education,” *Educational Research Review*, vol. 27, pp. 71–89, 2019.

[35] S. Strauß and N. Rummel, “Promoting regulation of equal participation in online collaboration by combining a group awareness tool and adaptive prompts. but does it even matter?” *International Journal of Computer-Supported Collaborative Learning*, vol. 16, pp. 67–104, 2021.