Educational Research Review 44 (2024) 100616

Contents lists available at ScienceDirect

Educational Research Review

journal homepage: www.elsevier.com/locate/edurev

Review

AI-driven learning analytics applications and tools in computer-supported collaborative learning: A systematic review

Fan Ouyang *, Liyin Zhang **

College of Education, Zhejiang University, #866, Yuhangtang Rd, Hangzhou, Zhejiang, 310058, China

A R T I C L E I N F O

Keywords:
Computer-supported collaborative learning (CSCL)
Learning analytics
AI-Driven learning analytics
Systematic review
Automatic feedback
Artificial intelligence

A B S T R A C T

Artificial intelligence (AI) has brought new ways for implementing learning analytics in computer-supported collaborative learning (CSCL). However, there is a lack of literature reviews that focus on AI-driven learning analytics applications and tools in CSCL contexts. To fill the gap, this systematic review provides an overview of the goals, characteristics, and effects of existing AI-driven learning analytics applications and tools in CSCL. According to the screening criteria, out of the 2607 initially identified articles between 2004 and 2023, 26 articles are included for final synthesis. Our results show that existing tools primarily focus on students' cognitive engagement. Existing tools primarily utilize communicative discourse, behavioral, and evaluation data to present results and visualizations. Despite various formats of feedback are provided in existing tools, there is a lack of design principles to guide the tool design and development process. Moreover, although AI techniques have been applied for presenting statistical information, there is a lack of providing alert or suggestive information in existing tools or applications. Compared with the positive impacts on collaborative learning, our results indicate a lack of support for instructional interventions in existing tools. This systematic review proposes the following theoretical, technological, and practical implications: (1) the integration of educational and learning theories into AI-driven learning analytics applications and tools; (2) the adoption of advanced AI technologies to collect, analyze, and interpret multi-source and multimodal data; and (3) the support for instructors with actionable suggestions and instructional interventions. Based on our findings, we provide further directions on how to design, analyze, and implement AI-driven learning analytics applications and tools within CSCL contexts.

1. Introduction

Computer-supported collaborative learning (CSCL), as a teaching and learning mode, has been widely used to promote knowledge inquiry and construction and enhance students' complex problem-solving skills through peer interactions and communications at the group level (Jeong & Hmelo-Silver, 2016; Wise et al., 2021). However, students usually encounter challenges when it comes to effectively engaging in CSCL activities, due to the lack of appropriate instructional guidance, group awareness, and stimulating mediums that motivate interaction (Chen, Tan, & Pi, 2021; Järvenoja et al., 2020; Shin et al., 2018). To address this problem, learning analytics (LA) can facilitate collaborative learning through the analysis of students' learning process and performance data, in order to

* Corresponding author.
** Corresponding author.
E-mail addresses: fanouyang@zju.edu.cn (F. Ouyang), liyinzhang_@zju.edu.cn (L. Zhang).

https://doi.org/10.1016/j.edurev.2024.100616
Received 4 December 2023; Received in revised form 15 April 2024; Accepted 6 July 2024
Available online 11 July 2024

capture a broad range of learners' behaviors and interactions throughout collaborative learning activities, and provide supportive information for learners (Martinez-Maldonado et al., 2022; Melzner et al., 2022). Moreover, advanced Artificial Intelligence (AI) techniques have been incorporated into learning analytics applications and tools with the advantages of providing real-time, adaptive, and personalized feedback (Chen et al., 2020; Gedrimiene et al., 2024). Existing literature reviews have primarily focused on learning analytics implementations in various educational contexts. However, there remains a gap in review work that specifically examines the integration of learning analytics and AI-driven techniques in CSCL, which plays an important role in providing in-time, adjustable information for collaboration. To address this gap, this systematic literature review reveals the attributes and effects of incorporating AI techniques in learning analytics applications and tools in CSCL, in order to identify future directions for the design, implementation, and evaluation of AI-driven LA applications and tools in CSCL.

2. Literature review

2.1. Computer-supported collaborative learning and collaborative learning analytics

Computer-supported collaborative learning (CSCL) is defined as a means of learning and instruction that can foster the social nature of learning by using a variety of technological and pedagogical strategies (Stahl et al., 2006). CSCL can be implemented in diverse ways, such as online collaborative writing, project-based learning, and technology-mediated discussions, which require negotiation, sharing of meanings, and collaborative knowledge construction in group work (Järvelä & Rosé, 2021; Jeong et al., 2019; Stahl et al., 2006). However, during the CSCL process, it is possible for students to become distracted, feel confused about the collaborative process, and end up with individual work instead of constructing collective knowledge (Cress et al., 2019; Kang et al., 2024). To address this problem, collaborative learning analytics have been proposed to understand, intervene, and optimize the quality of CSCL through analyzing and reporting data derived from learners in various collaborative learning contexts (Martinez et al., 2021; Wise et al., 2021). Collaborative learning analytics tools can reflect the progress during collaboration, evaluate the effectiveness of collaborative learning, and inform learners about particular aspects of collaboration (Jeong et al., 2019; Melzner et al., 2022; Wise et al., 2021). The implementation of collaborative learning analytics tools aims to facilitate students' understanding of teamwork dynamics, provoke reflections for further collaborative work, and provide actionable suggestions (Hmelo-Silver & Jeong, 2021; Lyons et al., 2021; Ouyang et al., 2024).

However, despite these affordances of collaborative learning analytics tools in supporting learners and instructors, studies have reported challenges emerging with the implementation of these tools. First, learners and instructors have difficulties in making real-time adjustments during the process of collaborative activities, due to the lack of intime, personalized feedback provided by learning analytics tools (Wong & Li, 2020). Second, most learning analytics-based feedback requires a large amount of time and labor, which limits the implementation of collaborative learning analytics tools in various teaching and learning contexts (Zhang et al., 2023). There is a call for supporting CSCL with real-time and dynamic feedback information, in order to facilitate self-regulation during the collaborative processes, enhance group awareness and learning adjustments, and ultimately foster collective knowledge construction (Gedrimiene et al., 2024; Pelletier et al., 2022). For example, Zheng et al. (2023) designed and implemented an automatic collaborative learning analytics tool in collaborative learning. This tool effectively enhanced students' collaborative engagement and group performance through the delivered real-time feedback encompassing students' cognitive, social, emotional, behavioral, and metacognitive aspects (Zheng et al., 2023). It is worth clarifying the definitions of feedback in this literature review. Generally speaking, feedback is regarded as corrective information that informs learners about their deviations from their learning objectives when any issues occur, in order to help them improve learning outcomes and achieve learning goals (Wisniewski et al., 2020). In the learning analytics field, feedback can be viewed from a broad perspective, primarily utilizing data analytics and visualizations to present timely and informative messages based on learners' learning process and performance (Banihashem et al., 2022; Vieira et al., 2018). Learning analytics feedback (i.e., the broad definition of feedback used in this work) aims to give insights into students' learning behavior, cognition, or performance to improve their self-regulation, awareness, and reflection (Sedrakyan et al., 2020; Tsai et al., 2021). Taken together, despite the potential of implementing collaborative learning analytics in CSCL, it is necessary to provide real-time and dynamic feedback for CSCL practice and development with automatic AI techniques.

2.2. AI-driven collaborative learning analytics applications and tools

There is a trend of developing AI-driven learning analytics applications and tools to create automatic feedback and adaptive support in CSCL, with the goal to trigger teaching and learning adaptations based on data-driven evidence for CSCL (Wise et al., 2021). On the one hand, based on statistical data, machine learning (ML) algorithms (e.g., Bayesian network, decision trees, support vector machines), deep learning (DL) algorithms, and evolutionary computation (EC) algorithms (e.g., Genetic programming, etc.) have been incorporated in collaborative learning analytics tools to providing categorization, grouping, predicting, and modeling (Fung et al., 2004; Jiao et al., 2022; Susnjak, T., 2023). For example, Ouyang et al. (2023) established a model to automatically predict student's performance scores during collaborative learning tasks. This research provided feedback and relevant learning suggestions for students, which led to improved writing performance in CSCL (Ouyang et al., 2023). On the other hand, based on the analysis of textual data from collaborative learning activities, natural language processing, and text mining techniques are incorporated with collaborative learning analytics tools. For example, Bidirectional Encoder Representations from Transformers (BERT), a natural language processing model, was integrated into a collaborative learning analytics tool to automatically classify students' collaborative discussions, which successfully promoted students' collaborative engagement and performance (Zheng et al., 2023). In summary, various

AI technologies have been integrated with collaborative learning analytics for aggregating, analyzing, and generating real-time and dynamic insights (Aydoğdu, 2021; Gedrimiene et al., 2024; Sandoval et al., 2018).

It is promising to integrate AI technologies into learning analytics tools to better analyze and understand the complexity of CSCL, such as the multiple relationships (e.g., supporting or conflicting perspectives) between students' collaborative discourses (Cress et al., 2021). The AI-driven approach in CSCL environment can improve the efficiency of data analytics, provide personalized feedback for students and instructors, and empower real-time feedback on complex CSCL contexts (Conijn et al., 2020; Holstein et al., 2019; Xing et al., 2023). Taken together, AI-driven learning analytics applications can effectively improve the efficiency of data analysis and representation, support students with real-time, adaptive peer-level or group-level information, and provide dynamic instructional guidance and scaffolding for CSCL.

2.3. *Review of previous systematic reviews*

Existing review works have been focused on the context of higher education, but few reviews specifically focus on LA applications and tools in the context of CSCL environment. For example, Banihashem et al. (2022) identified the data types, analytic methods, and stakeholders of LA tools in the context of higher education. Wong and Li (2020) conducted a review of LA interventions in higher education, summarizing intervention methods, their effects, and associated challenges. In one example of LA tools in CSCL, Hu and Chen (2021) reviewed visual representations used for analyzing collaborative discourse. Their review work unveiled the goals, data sources, visualization designs, and analytical techniques featured in LA applications (Hu & Chen, 2021). However, this review did not explicitly focus on providing AI-driven learning analytics applications within CSCL environment, which provide automatic feedback for students and instructors to understand the complexity of CSCL. Taken together, there is a lack of reviews investigating AI-driven learning analytics applications and tools within CSCL environment to provide real-time and adaptive feedback to support CSCL.

Second, recent reviews of AI-driven applications and tools in educational contexts have mainly focused on describing the general applications, techniques, and effects of AI on education, whereas there is a lack of reviews summarizing how to provide AI-driven, automatic feedback in educational contexts. For example, existing reviews have summarized the general applications of AI techniques (such as adaptive systems, and tutorial robots), categorized specific AI techniques (such as machine learning techniques and deep learning techniques), and concluded probable effects of AI technologies and applications in different educational contexts (e.g., Ouyang, Zheng, & Jiao, 2022; Salas-Pilco & Yang, 2022; Zhan et al., 2022). The use of AI techniques has been found to be useful for facilitating knowledge gain, skill acquisition, group task performance, and social interaction in collaborative learning contexts (Chen et al., 2018). However, we found there is an inadequacy of review work examining the specific steps for implementing AI techniques in CSCL contexts, including data collection, data analysis, data visualization, and the design of AI-driven learning analytics tools.

2.4. *Conceptual framework*

We propose a conceptual framework of AI-driven learning analytics applications and tools in CSCL (see Fig. 1). Drawing upon an existing learning analytics framework (Greller & Drachsler, 2012), we identify six critical dimensions of learning analytics applications and tools in CSCL, including stakeholders (students and instructors), objectives (monitoring, reflection, and prediction), data (data collected from the students or instructors), instruments (techniques and theories), design and visualizations (presentation and interpretation of data analytics results), and effects (effects and ethical considerations) (Greller & Drachsler, 2012). This learning analytics framework serves as a checklist for designing learning analytics applications and tools, comparing context parameters across various learning environments, and applying educational data to support teaching and learning (Greller & Drachsler, 2012).

Furthermore, this learning analytics framework can be developed to reflect the multi-level, process-oriented characteristics of CSCL

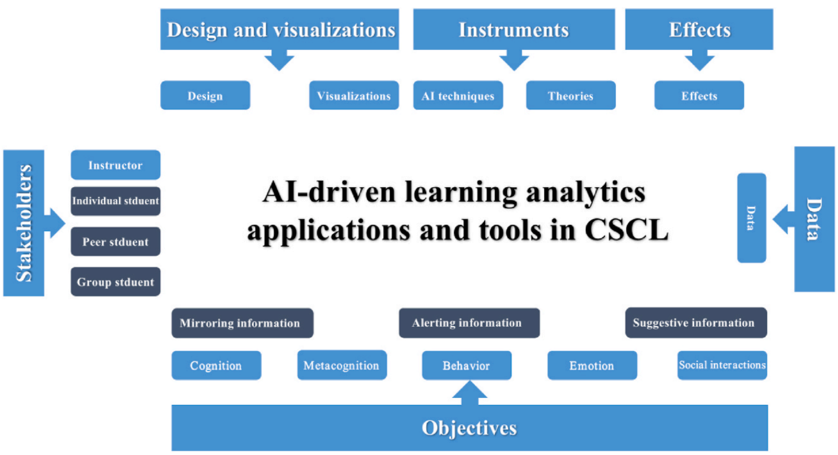


Fig. 1. The conceptual framework.

with the use of AI-driven learning analytics applications and tools in teaching and learning. First, the scope of stakeholders should be expanded in CSCL, because the high-quality collaboration emphasizes the interconnected relationships between individual students, their peers and the group when they share information and knowledge, make peer interactions, and construct group knowledge (Cress et al., 2021; Järvelä et al., 2016). Second, the objectives are extended to include process-oriented analysis of student cognition, metacognition, behavior, emotions, and social networking relationships, which are all crucial factors in CSCL contexts (Zheng et al., 2023). Third, the instruments are expanded to include AI-driven learning analytics applications, such as artificial intelligence, machine learning, and deep learning techniques (Chen et al., 2020). The dimensions of data, design and visualizations, and effects followed the conceptual framework proposed by Greller and Drachsler (2012), representing the fundamental features of AI-driven learning analytics applications and tools used in the CSCL context.

3. Methods

3.1. Research questions

This systematic literature review aims to (1) investigate the implementation of AI-driven learning analytics applications and tools in CSCL and (2) provide a deep understanding by revealing the goals, data sources, design attributes, AI techniques, and effects of AI-driven learning analytics applications and tools in the CSCL contexts. To achieve these goals, this review focuses on the following five research questions:

RQ1: What are the goals of AI-driven learning analytics applications and tools in the CSCL field?
RQ2: What data sources do previous studies use as indicators of AI-driven learning analytics applications and tools in the CSCL field?
RQ3: What are the design attributes of AI-driven learning analytics applications and tools in the CSCL field?
RQ4: What approaches are used to analyze and interpret the learning data in AI-driven learning analytics applications and tools in the CSCL field?
RQ5: What are the effects of AI-driven learning analytics applications and tools on learning and instruction in the CSCL field?

3.2. Research methods and procedures

The research methodology was a systematic literature review approach, that employed reliable and auditable techniques to analyze and understand all relevant research on the researched topic area. Specifically, this research was conducted according to the preferred reporting of items for systematic reviews and meta-analyses (PRISMA) protocol (Moher et al., 2009) to locate papers about AI-driven learning analytics applications and tools in computer-supported collaborative learning. The systematic procedures followed five stages, namely database search, identification of search terms and time period, searching criteria, screening process, and analysis and categorization.

3.3. Database search

To locate the relevant articles, this research was conducted on the following six publisher databases: Web of Science (https://www.webofscience.com/), ACM (https://dl.acm.org/), EBSCO (https://search.ebscohost.com), Scopus (https://www.scopus.com), Wiley Online Library (https://onlinelibrary.wiley.com/), Taylor & Francis (https://www.tandfonline.com/). All searches were conducted in August 2023. Filters were used for the peer-reviewed articles published in English and the time period was limited from January 2004 to December 2023. Although peer-reviewed conferences included substantial research on the design and initial development of AI-driven learning analytics applications and tools, there is an absence of empirical implementation of these tools in CSCL contexts. Consequently, we excluded peer-reviewed conferences from our consideration. Finally, after completing the screening of the full articles, snowballing was performed based on the guidelines described by Wohlin (2014) to find the articles that were not extracted using the search strings.

Table 1
Search string used in searching for articles from the database.

Topic	Search terms
AI-driven techniques for automatic feedback	"artificial intelligence" OR "AI" OR "decision tree" OR "machine learning" OR "neural network" OR "deep learning" OR "k-means" OR "random forest" OR "support vector machines" OR "logistic regression" OR "fuzzy-logic" OR "Bayesian network" OR "latent Dirichlet allocation" OR "natural language processing" OR "genetic algorithm" OR "genetic programming" OR "data mining" OR "text mining" OR "automatic" OR "adaptive" OR "instant" OR "predict"
Learning analytics applications and tools	"learning analytics" OR "educational analytics" OR "teaching analytics" OR "dashboard" OR "system" OR "software" OR "platform" OR "workspace" OR "visualization" OR "feedback" OR "intervention" OR "remediation" OR "hint" OR "recommend"
Computer-supported collaborative learning	("collaborative" OR "cooperative" OR "collaboration" OR "group" OR "peer") AND ("online" OR "computer" OR "technology" OR "software" OR "virtual")

3.4. Identification of search terms

In the search process, we performed a title and abstract screening process by using a set of keywords related to AI-driven techniques for automatic feedback, learning analytics applications and tools, and computer-supported collaborative learning (see Table 1).

3.5. Searching criteria

The searching criteria were proposed to locate the articles that focused on AI-driven learning analytics applications and tools in CSCL environments. A set of inclusion and exclusion criteria were adopted (see Table 2).

3.6. Screening process

The screening process involved the following procedures: 1) removing the duplicated articles; 2) removing the studies that unsatisfied the inclusion rules based on the titles and abstracts; 3) reading the full articles again and removing the studies that unsatisfied the inclusion rules (see Fig. 2). Furthermore, according to Wohlin (2014)'s snowballing method, we broadened our search by checking the reference lists of the articles that were already included. Any articles that failed to meet our set criteria were discarded following the preceding steps. When multiple articles described the same tool (e.g., one author submitted multiple types of work describing the design of one tool), only the most detailed and/or recent article was retained as the representative publication (see Fig. 3).

Initially, one research assistant carried out the preliminary search independently, locating 3017 articles. Second, among the 3, 017 articles, 189 duplicates were identified and removed. Third, through an examination of titles and abstracts, the number of articles was reduced to 655 following the inclusion and exclusion criteria. Fourth, two research assistants independently carried out a detailed review of 20% of the 655 selected articles based on the inclusion and exclusion criteria, achieving an inter-rater agreement of 81.7%. Any disagreement that occurred was resolved through discussion until consensus was attained between the authors and two research assistants. Finally, the second author checked the entire content of the remaining 80% of the 655 chosen articles to determine whether the articles meet with the inclusion and exclusion criteria. Consequently, a total of 26 papers were included in this systematic review and confirmed by the first author.

3.7. Data coding and analysis

This review adopted the deductive and inductive content analysis methodology proposed by Elo and Kyngäs (2008). Based on our proposed conceptual framework, the included publications were analyzed by employing deductive content analysis based on the five components, namely *Goal, Data, Design, Technique,* and *Implementation* of existing AI-driven learning analytics applications and tools (see Table 3). We scrutinized the research findings based on the five components and utilized inductive content analysis to extract and categorize the results under a broader concept. Moreover, basic information on the article itself was captured (e.g., the year of publication, the type of publication), which allowed us to conduct an overview of features of existing studies of AI-driven learning analytics applications and tools in CSCL. The first author independently performed all of the coding. The research assistant randomly chose 23% ($N = 6$) of the papers to confirm their reliability. The inter-rater agreement was initially 77.5% and then brought to 90.7% after discussions.

4. Results

Among the 26 empirical articles, 62% ($N = 16$) of the articles were published within or after 2019. The major countries for the 26 studies were also identified. The most prolific country or area was China which had 10 publications (38%), followed by USA ($N = 3$, 12%) and Spanish ($N = 3$, 12%). The included articles were derived from a variety of journal sources, with the highest frequency of publication from *International Journal of Computer*-Supported *Collaborative Learning* ($N = 7$, 27%), followed by *Computers & Education* ($N = 5$, 19%), *IEEE Transactions on Learning Technologies* ($N = 2$, 8%), and *The Internet and Higher Education* ($N = 2$, 8%).

Table 2
The exclusion and inclusion criteria.

Inclusion criteria	Exclusion criteria
The study should propose a learning analytics tool to support teaching and learning	Studies that are not relevant to the research question.
The proposed tool in a study should provide automatic feedback for users.	Uncompleted research is excluded, for example, research that only reports a tool, but do not report empirical results.
The study is contextualized in computer-supported collaborative learning environment.	Posters or conference proceedings are excluded.
The studies should be published publicly with full-text available.	
The studies should be reported in English.	
The studies should be published from 2004 to 2023.	

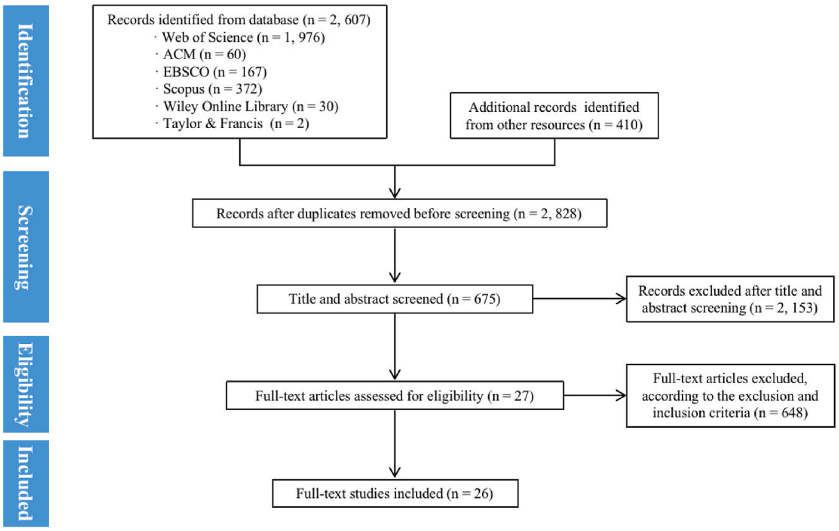


Fig. 2. Selection flowchart used in this research based on PRISMA.

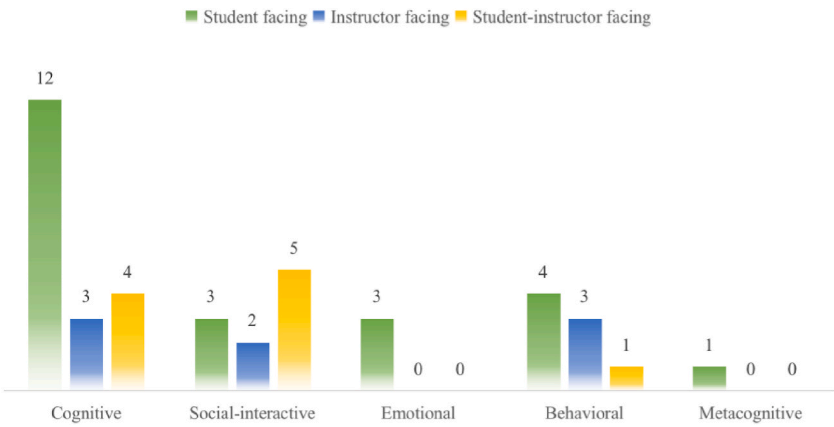


Fig. 3. Distribution of target underlying learning process. Note. Some tools contained multiple target underlying learning processes.

4.1. RQ1: what are the goals of AI-driven learning analytics applications and tools in the CSCL field?

The goals of AI-driven learning analytics applications and tools in the CSCL field were examined from the learning environments, target users, and target learning process. Regarding the learning environment, results revealed that 88% ($N = 23$) of the applications and tools were contextualized in the online learning environment, whereas only 3 of the reviewed articles were tailored to the face-to-face learning environment. First, a majority of tools ($N = 16$, 62%) were designed to assist learners (e.g., Chen, Li, et al., 2021; Echeverría Rodríguez & Cobos Pérez, 2021; Ouyang et al., 2023). These student-facing learning analytics applications and tools primarily highlight the cognitive dimension ($N = 12$) of collaboration, followed by the behavioral ($N = 4$), social-interactive ($N = 3$), emotional ($N = 3$), and metacognitive ($N = 1$) dimensions. Second, 23% ($N = 6$) of the articles were specifically designed for instructors, intending to facilitate the supervision of collaborative processes and provide instructional intervention throughout the courses (e.g., Amarasinghe et al., 2020; Tissenbaum & Slotta, 2019). These instructor-facing tools shed light on cognitive ($N = 3$), behavioral ($N = 3$), and social-interactive ($N = 2$) dimensions. Third, 23% ($N = 6$) of the reviewed articles provided integrated feedback for both students and instructors within a single tool (e.g., Han et al., 2021; Lonchamp, 2010). Among these student- and instructor-facing learning analytics applications and tools, research focused on the social-interactive ($N = 5$) and cognitive ($N = 4$) dimensions of collaboration, whereas lacking attention on the emotional, metacognitive, and behavioral dimensions of collaboration. Among the 26 reviewed articles, only 10 articles (38%) included more than two dimensions in their proposed tools (e.g., Han et al., 2021; Martinez-Maldonado, 2019). Overall, the cognitive dimension was the primary focus of feedback in existing AI-driven learning analytics applications and tools, with a lack of representing multiple aspects of students' collaborative learning.

Table 3
The coding framework for included studies.

Code	Interpretation
Goal	
1. Target user	Learner, Instructor
2. Learning format	Online, Face-to-face
3. Target underlying learning process	*The focused learning process of feedback information in AI-driven learning analytics applications and tools when providing feedback.* Cognitive, Social-interactive, Behavioral, Metacognitive, Emotional
Data	
4. Data sources	Discourse: discussion transcript, discourse feature, interaction; Behavior: frequency of participation, interactions with the platform, progress of activities, time spent; Evaluation: performance scores, tagging
5. Level of data sources	Individual-level, Peer-level, Group-level
Design	
6. Design principle	*Any design principles mentioned in AI-driven learning analytics applications and tools.*
7. Visualization format	Network, Line, Bar, Pie, Radar, Raw text, Timeline, Word cloud, Signal light
Technique	
8. Technique	*Any techniques for data analysis mentioned in the article.*
9. Visualization information	Mirroring information *(reflecting students' status of the collaboration)* Alerting information *(alerting the user to specific features of the collaboration)* Suggestive information *(providing actionable suggestions, prompts, or resources for users to make adjustments)*
Implementation	
10. Effect	Effects on learning: cognitive engagement; behavioral engagement; social engagement; awareness of collaboration; group performance; motivation and learning interest; skill; group regulation; Effects on instruction: awareness of collaboration; instructional intervention; Not specific: not specific

4.2. RQ2: what data sources do previous studies use as indicators of AI-driven learning analytics applications and tools in the CSCL field?

Our results revealed the distinctions among the individual, peer, and group levels across data sources of discourse, behavior, and evaluation data. Individual-level data sources were the most frequently mentioned ($N = 27$), followed by group-level data ($N = 6$), and peer-level data ($N = 22$) (see Fig. 4).

From the individual-level perspective, behavioral data was the most frequently-occurred data category ($N = 12$), including the frequency of participation ($N = 4$), interaction with the platform ($N = 4$), time spent ($N = 2$), and progress of activities ($N = 2$). The second most frequently used data source was the discourse data ($N = 9$), including discourse feature ($N = 6$) and discussion transcript ($N = 3$). The discussion transcripts were collected as log data from the online platform (e.g., Li et al., 2021). Compared with discussion transcripts, discourse features went further by extracting specific characteristics from the transcripts, such as the score of discussion depth (Ouyang et al., 2023), keywords (Peng et al., 2022), and discourse style (e.g., Lonchamp, 2010). Third, the evaluation data, namely tagging ($N = 4$) and performance score (N = 2) were also used as raw data in existing tools. Specifically, the tagging data was obtained by having students assign labels to their discourses or behaviors, which included collecting data related to their learning styles (e.g., Casamayor et al., 2009) and students' sentence classifiers (e.g., Walker et al., 2011).

From the peer-level perspective, interaction ($N = 4$) was the most frequently-occurred data category of the peer-level data. For example, students' interactions with peers were collected by Han et al. (2021) to explain how group members exchanged questions to promote each other's thoughts. Other peer-level data sources, such as frequency of participation ($N = 1$) and performance score ($N = 1$) were rarely used as indicators of AI-driven learning analytics applications and tools.

From the group-level perspective, first, discourse data was the most frequently-occurred group-level data category ($N = 15$), including discussion transcript ($N = 7$), discourse feature ($N = 5$), and interaction ($N = 3$). Specifically, discussion transcripts collected from the online platforms were the most popular group-level data source (e.g., Chen, Li, et al., 2021; Li et al., 2021; Walker et al., 2011). In addition, group-level discourse features included the total number of individual answers (e.g., Amarasinghe et al., 2020), keywords of the collaborative discussion (e.g., Peng et al., 2022), and the delivery of materials from the knowledge base based on emergent semantic connections (e.g., Tissenbaum & Slotta, 2019). In addition, interactions that occur at the group level were also utilized as an indicator in existing applications and tools (e.g., Tissenbaum & Slotta, 2019). Second, behavioral data was also widely

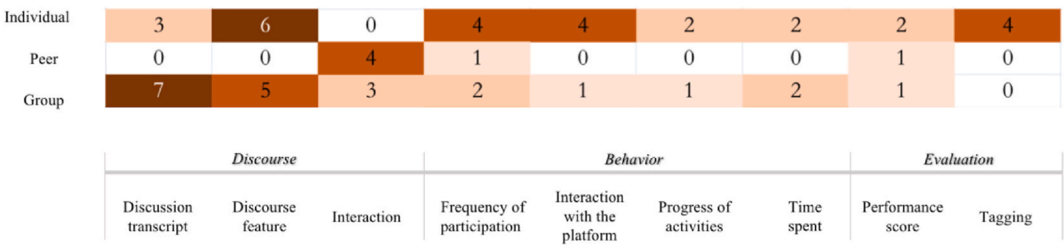


Fig. 4. A heatmap of data sources at individual, peer, and group levels. Note. Some tools included multiple data sources.

involved in existing AI-driven learning analytics applications and tools ($N = 6$), including frequency of participation ($N = 2$), interaction with the platform ($N = 2$), progress of activities ($N = 1$), and time spent ($N = 1$). For example, the frequency of participation included the total number of students who participated in collaborative activities (e.g., Amarasinghe et al., 2020). In addition, the time-spent indicators included average reply time (e.g., Bravo et al., 2008) and total working time (e.g., Bravo et al., 2008; Martinez-Maldonado, 2019) in empirical research. Third, group-level evaluation data was only mentioned once by Ouyang et al. (2023), which collected students' performance scores for writing tasks and further provided performance predictions.

Overall, individual- and group-level data was frequently used as indicators of AI-driven learning analytics applications and tools, whereas peer-level data received less attention. Moreover, considering the data sources across the three levels, various types of discourse data and behavior data were frequently used as data-driven evidence in existing applications and tools.

4.3. RQ3: what are the design attributes of AI-driven learning analytics applications and tools in the CSCL field?

4.3.1. Visualization format

Our findings showed that raw text ($N = 17$, 65%) was the most frequently-used format in AI-driven learning analytics applications and tools, followed by bar plot ($N = 9$, 35%), pie plot ($N = 5$, 19%), and network ($N = 4$, 15%). Timeline graphs, radar plots, word clouds, signals, and line graphs (all smaller than 15%) were seldom used as visualization formats in existing tools. First, raw text was found to be commonly used in the existing tools to illustrate the analysis results of collaborative learning. Various types of texts have been applied in existing tools, including prompts and suggestions (e.g., Strauß & Rummel, 2021; Tissenbaum & Slotta, 2019), scores and numbers (e.g., Ouyang et al., 2023), and conversational agent (e.g., Xie et al., 2021). Second, existing tools used bar plots or pie plots to show the frequency and percentage of students' collaborative discourse, behavior, and performance (e.g., Peng et al., 2022; Van Leeuwen, 2015). Third, the results also suggested that network was used to illustrate the connections between knowledge, discourses, and behaviors during collaborative processes, such as social networks (e.g., Han et al., 2021; Li et al., 2021) and co-word networks (e.g., Chen et al., 2023). Taken together, a variety of visualization formats were employed to provide statistical and cognitive information in existing AI-driven learning analytics applications and tools in CSCL.

4.3.2. Design principle

We found that 15% ($N = 4$) of the reviewed articles have mentioned design principles in AI-driven learning analytics applications and tools. Among the four articles, the design principles served as guidance for how to provide group-level feedback based on educational and learning theories. For example, Chen et al. (2023) followed the social navigation principle for designing the LA tool in CSCL. The principle promoted understandable information on students' social interactions presented in the tool, which enabled students to reflect on their social engagement and adjust their collaboration (Chen et al., 2023). However, among the other tools without design principles or educational and learning theory ($N = 22$, 85%), LA tools might not prioritize providing group-level and peer-level visualizations and feedback, which were essential to reveal the complexity of CSCL. For example, Xie et al. (2021) only utilized individual students' interactions with the platform and performance scores as indicators of students' collaborative performance, neglecting aspects such as collaborative interactions, knowledge sharing, and collective knowledge construction among learners. In summary, existing research lacked the connection between design principles with AI-driven learning analytics applications and tools to guide CSCL.

4.4. RQ4: what approaches are used to analyze and interpret the learning data in AI-driven learning analytics applications and tools in the CSCL field?

4.4.1. AI-driven techniques

Our results showed that only 14 articles had discussed specific techniques. Among these articles, natural language processing (NLP)

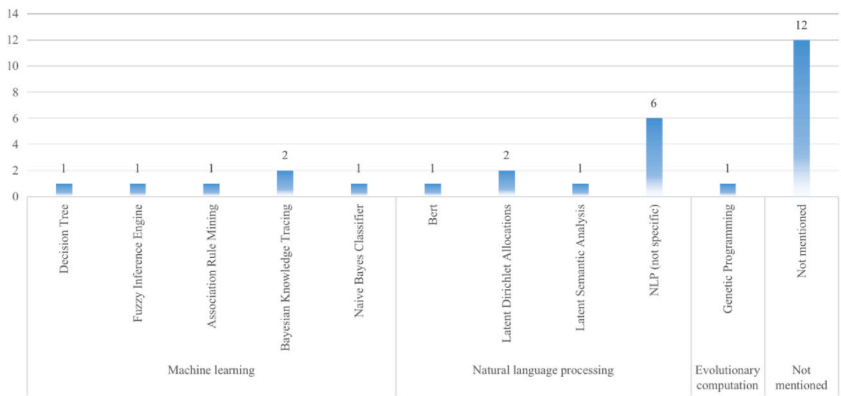
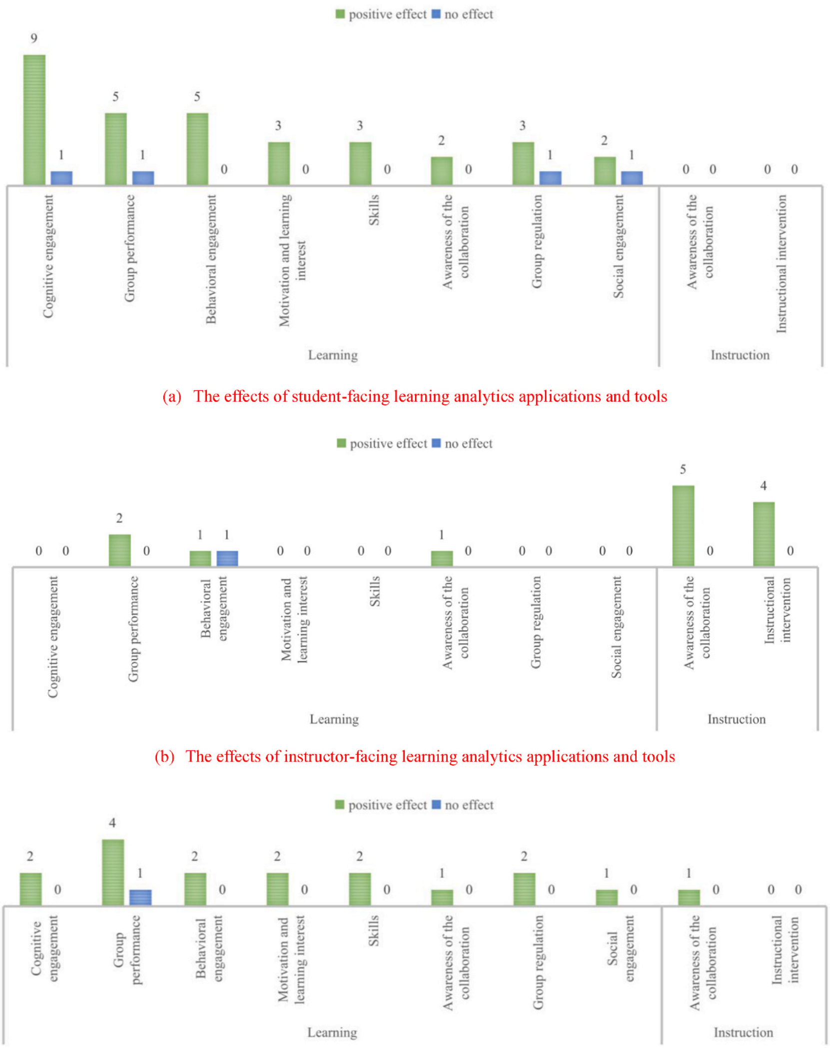


Fig. 5. Techniques used by AI-driven learning analytics tools in CSCL. Note. Some tools utilized multiple techniques.

($N = 10$, 38%), machine learning (ML) techniques ($N = 6$, 23%), and evolutionary computation (EC) algorithms ($N = 1$, 4%) were mentioned in existing AI-driven learning analytics applications and tools (see Fig. 5). First, NLP techniques were frequently employed in existing applications and tools for the analysis of text-based data in CSCL environments. However, most of these tools utilizing NLP ($N = 6$) did not specify the particular models and techniques for pre-training, feature extraction, and prediction or analysis. Among those specified particular models and techniques, three models were mentioned in existing applications and tools, namely Bidirectional Encoder Representations from Transformers (BERT, $N = 1$), Latent Dirichlet Allocations (LDA, $N = 2$), and Latent Semantic Analysis (LSA, $N = 1$). These models played key roles in analyzing the semantic relationships between different words (e.g., Trausan-Matu et al., 2014), encoding and categorizing students' collaborative words (e.g., Zheng et al., 2023), and extracting primary topics from students' collaborative discussions (e.g., Chen, Li, et al., 2021; Erkens et al., 2016). Second, ML techniques were used to generate categorization, modeling, and prediction based on statistical and numerical educational data. For example, the decision tree algorithm ($N = 1$) was implemented in learning analytics tools to assess the degree of relevance between posted online transcripts (e.g., Huang et al., 2011). Fuzzy inference engine ($N = 1$) and association rule mining ($N = 1$) were utilized to calculate the probability of students' learning behaviors. Third, EC algorithms were only mentioned once among the reviewed articles. Specifically, Ouyang et al. (2023) employed Genetic Programming (GC) to create a prediction model using students' performance data, delivering predictive results as automatic feedback to support CSCL. In summary, the most commonly used techniques for providing automatic feedback in CSCL were NLP algorithms, followed by ML techniques and EC algorithms.

4.4.2. Visualization information

The reviewed AI-driven learning analytics applications and tools primarily included three types of visualization information,



(a) The effects of student-facing learning analytics applications and tools

(b) The effects of instructor-facing learning analytics applications and tools

(c) The effects of students- and instructor-facing learning analytics applications and tools

Fig. 6. The effects of existing AI-driven learning analytics applications and tools on CSCL.

including mirroring information ($N = 18$, 69%), suggestive information ($N = 11$, 42%), and alerting information ($N = 10$, 38%). First, mirroring features was the most frequently-occurred information in existing AI-driven learning analytics applications and tools, which reflected students' current status of collaboration. These mirroring features were designed to increase individual students' awareness of the collaborative performance (e.g., Martinez-Maldonado, 2019; Strauß & Rummel, 2021) as well as to enhance group awareness (e.g., Peng et al., 2022; Zheng et al., 2023). Second, the reviewed applications and tools provided actionable suggestions, prompts, or resources to users, in order to assist students and instructors in selecting suitable strategies for further collaboration (e.g., Walker et al., 2011; Zheng et al., 2019; Zheng et al., 2023). Third, 10 of the included tools provided alerting messages for students and instructors, which directed users' attention to particular aspects of collaboration through modeling and data mining (e.g., Echeverría Rodríguez & Cobos Pérez, 2021; Martinez-Maldonado, 2019; Ouyang et al., 2023). Taken together, mirroring information of collaboration was most frequently-utilized for interpreting the learning data from CSCL contexts, whereas alerting information and suggestive information received less attention in existing applications and tools.

4.5. RQ5: what are the effects of AI-driven learning analytics applications and tools on learning and instruction in the CSCL field?

Results showed that AI-driven learning analytics applications and tools had various effects on learning and teaching improvement (see Fig. 6). For student-facing learning analytics applications and tools, results showed positive influence of tools on students' learning processes (such as cognitive engagement, behavioral engagement, and social engagement) and outcomes (such as group performance and skills). Other improvements included increased motivation and learning interest (N = 3), group regulation (N = 3), and awareness of collaboration (N = 2). For instructor-facing learning analytics applications and tools, the most frequently reported impact was enhancing instructors' awareness of students' collaborative processes (e.g., Amarasinghe et al., 2020), followed by providing suggestions for instructional intervention (e.g., Martinez-Maldonado, 2019). There were also improvements in students' learning outcomes, such as facilitating students' group performance, awareness of collaboration, and behavioral engagement, as a result of using these tools. Regarding students- and instructor-facing tools, positive impacts were verified in learning dimensions such as student learning processes (cognitive, behavioral, and social engagement) and outcomes (group performance and skills), as well as instructional dimensions like improving instructors' understanding of students' collaborative processes and providing advice for instructional intervention design. Overall, our review showed more improvements in learning outcomes than instructional enhancements with the usage of AI-driven learning analytics applications and tools.

(a) The effects of student-facing learning analytics applications and tools
(b) The effects of instructor-facing learning analytics applications and tools
(c) The effects of students- and instructor-facing learning analytics applications and tools

5. Discussions and implications

5.1. Addressing the research questions

AI techniques have introduced promising methods for implementing learning analytics in CSCL, which offers the potential to improve both instruction and learning through real-time, personalized feedback. Given the limited literature review focusing on AI-driven learning analytics applications and tools in CSCL, it is necessary to explore research and practice directions of integrating AI techniques with learning analytics tools and to investigate how these tools influence CSCL (Kabudi et al., 2021). This research conducted an overview of AI-driven learning analytics applications and tools in CSCL environment, including goals, data, design attributes, techniques, and effects of existing tools in empirical research. First, our results indicated a focus on the cognitive dimension of collaborative processes in current AI-driven learning analytics tools, but there was a lack of incorporating other learning processes (such as metacognitive, social-interactive, behavioral, and emotional dimensions) to comprehensively reflect students' collaboration. Second, existing research had utilized various data sources (i.e., discourse data, behavioral data, and evaluation data) from the individual- and group-level, but there was a lack of peer-level evidence and multimodal data in existing applications and tools. Third, regarding the design attributes, existing applications and tools had used various visualizations (e.g., raw text, bar plot, pie plot) to present feedback within CSCL contexts, but fell short in connecting design principles during the design and implementation of AI-driven learning analytics applications and tools. Fourth, regarding how to analyze and interpret the learning data, three major categories of AI techniques (i.e., NLP, ML, and EC) were used in these tools for data analysis, data mining, and providing automatic feedback in existing tools. The most frequent interpretation of data was to provide mirroring statistical information on students' behavior and engagement, followed by alerting information and actionable suggestions based on learning data. Fifth, most of the reviewed studies demonstrated positive effects on learning, compared to the instructional improvement observed in existing tools. Based on the results, we propose theoretical, technological, and practical implications for further AI-driven learning analytics applications and tools in CSCL.

5.2. Theoretical implications

Based on our conceptual framework, future work should make efforts to adopt educational and learning theories in AI-driven learning analytics applications and tools in CSCL context, aiming to guide the integration of AI techniques with student needs, ways of interaction and problem-solving within LA tools, and the use of predictive learning analytics. Similar to previous work (e.g., Guzmán-Valenzuela et al., 2021; Khalil et al., 2022), our results showed that few studies have focused on incorporating design principles and educational and learning theories in AI-driven learning analytics applications and tools. In addition, the results indicated that most of the existing applications and tools emphasized students' individual-level engagement, behavior, and performance in CSCL. Given that CSCL emphasizes connections across levels of individual engagement, peer interactions, and group practices (Stahl, 2016), it is essential to emphasize the importance of peer- and group-level evidence that reveals the collaborative learning process and supports socially shared regulation among groups of learners (Grand et al., 2016; Stahl, 2013). To realize this goal, educational and learning theories can be incorporated into the design and implementation of AI-driven learning analytics applications and tools, in order to provide guidance for understanding CSCL from the combination of individual, peer, and group perspectives (Wise et al., 2021). For example, following the principle of reference framework (Jivet et al., 2020), researchers supported students with comparative information from peers (e.g., keywords contributed by another student working on the same topic), which turned out comprehensible for students and aided them in understanding, processing, and reflecting on the analytics and visualizations (Chen, Li, et al., 2021). Overall, it is necessary to connect the design principle of AI-driven learning analytics applications with educational and learning theories, with the goal to focus on the interactive characteristics of CSCL.

5.3. Analytical implications

From the data analysis and modeling perspective, multimodal data is expected to be collected, analyzed, and interpreted to generate a comprehensive understanding of students' collaborative processes (Noroozi et al., 2019; Ouyang, Dai, & Chen, 2022; Panadero & Lipnevich, 2022). Our findings showed that single-modal data (i.e., discourse data or behavioral data) was predominantly utilized as indicators of students' performance, whereas multimodal data (e.g., discourse data as text data, online logs as behavioral data, physical space) was rarely collected and analyzed in existing applications and tools. This limitation might stem from the underutilization of AI-driven learning analytics applications and tools in the face-to-face educational environments, because it is more feasible to collect multimodal data in the face-to-face context compared with the online context. Since emerging wearable devices (e.g., wearable cameras, sensor devices, and smartwatches) are becoming available to collect and analyze multimodal data with low intrusiveness, future learning analytics applications and tools can offer insights into the dynamic change and development of students' learning processes (Ba & Hu, 2023; Chen et al., 2020; Giannakos et al., 2019). For example, Martinez-Maldonado et al. (2022) employed an automated system to visualize social, physical, affective, and epistemic evidence in a clinical simulation context, thereby sparking students' cognitive engagement and reflective ideas during collaboration. In addition, advanced analytics techniques and educational data mining methods can be integrated into future AI-driven learning analytics applications and tools for analyzing, visualizing, and modeling the multimodal data (Dierdorff & Ellington, 2012; Kim et al., 2013; Çini et al., 2023). For example, Noroozi et al. (2019) used MATLAB's image processing, statistics, and machine learning toolboxes to design an integrated analysis tool, that was customized for the visualization and analysis of diverse data sources and provided a navigable view of students' learning situations. These advanced modeling techniques in AI-driven learning analytics hold promise in investigating a range of cognitive and non-cognitive learning processes during collaboration, inspecting students' hidden physiological reactions, and revealing the complexity and reciprocity of learning processes during collaborative learning (Noroozi et al., 2020; Ouhaichi et al., 2023). Taken together, future AI-driven learning applications and tools should extend the use of multimodal learning analytics, in order to understand the complex, synergistic, and dynamic CSCL process and generate in-depth interpretations of learning analytics results on students' status.

5.4. Pedagogical implications

From a practical perspective, future studies should consider integrating student-facing feedback and instructor-facing feedback in AI-driven learning analytics applications and tools to provide comprehensive support for CSCL. First, our results indicated that there were fewer studies demonstrating positive effects on instruction, compared to the various types of learning improvement observed in existing tools. While these tools enhance students' awareness and reflections on collaboration, it is critical for instructors to provide guidance, scaffolding, and intervention through instructor-facing feedback during CSCL (van Leeuwen et al., 2014). Future tools can integrate student- and instructor-facing feedback, which can facilitate not only students' awareness but also instructors' understanding of CSCL processes and dynamical monitoring of collaboration (e.g., Han et al., 2021; Janssen et al., 2007; Zhang et al., 2023).

Second, we recommend the incorporation of actionable suggestions for instructors within AI-driven learning analytics applications and tools. While existing tools positively impacted instructors' awareness of collaboration, our findings revealed that there were fewer tools effectively enhancing instructional interventions during CSCL. As shown in previous research, instructors may have difficulty making instructional adjustments merely based on learning analytics feedback, due to the lack of background information, the complexity of collaboration, and the limited time within a course (Cai et al., 2023; Mangaroska & Giannakos, 2019). To address this problem, it is necessary to empower instructors with real-time, suggestive information through learning analytics applications and tools, prioritize their role in decision-making during the learning analytics feedback process, and involve them in the selection of actionable data features within AI-driven learning analytics applications and tools (Rummel et al., 2016; Wong & Li, 2020). By translating the sophisticated models and metrics to understandable forms, instructors can better adapt the course orchestration, intervene in groups' progress, and ultimately improve the teaching and learning efficiency in CSCL (Grammens et al., 2022; Holstein et al., 2019). Moreover, it is worth noting that generative AI has significantly impacted instruction and learning in higher education, prompting a reevaluation by creating personalized content, transforming learning environments, and offering potential solutions for teachers and students (Lan & Chen, 2024; Thompson et al., 2023). Taken together, there is a need to combine actionable, instructor-facing feedback with student-facing feedback in AI-driven learning analytics applications and tools, in order to promote collaboration by engaging both students and instructors in improving CSCL.

6. Conclusions, limitations, and future directions

AI-driven learning analytics has great potential to advance capabilities for supporting collaboration by offering real-time, adaptive, and detailed feedback that reveals the complexity of dynamic collaborative processes and group interactions (Wise & Schwarz, 2017). Compared with traditional methods that manually analyze group learning engagement after CSCL activities, AI-driven learning analytics can better analyze the large volume of learning data within CSCL contexts, promote students' self-, co-, and socially shared regulation, and support instructors with real-time diagnostic aids for collaboration (Conijn et al., 2020). Moreover, with the development of physiological devices and technologies, AI-driven learning analytics applications and tools can simultaneously trace students' cognitive and non-cognitive processes based on multimodal data (e.g., eye tracking, emotional data, etc.) to support collaborative learning processes and outcomes (Çini et al., 2023; Cukurova et al., 2020). Given the potential of AI techniques in learning analytics tools, this systematic review provides an overview of empirical research by examining the goals, data sources, design attributes, techniques, and effects of AI-driven learning analytics applications and tools within CSCL. The primary limitation of this study is related to the search process. Although we used the keyword list, the selection criteria and the database we employed to capture relevant papers might not guarantee full completeness. Since AI-driven learning analytics is an interdisciplinary field of computer science and education, the latest studies published as conference papers were not included. Moreover, the current study only provided a systematic overview of AI-driven learning analytics applications and tools in CSCL, a formal meta-analysis would be beneficial to report the effect sizes of selected empirical research to gain a deeper understanding of the field. Future design, research, and practice of AI-driven learning analytics applications and tools should shed light on educational and learning theories, use advanced algorithms to collect and analyze multi-source and multimodal data, and provide actionable, instructor- and student-facing feedback for CSCL.

Availability of data and materials

The data will be available on request from the corresponding author.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by National Natural Science Foundation of China (62177041); Zhejiang Province educational science and planning research project (2022SCG256); and Zhejiang University graduate education research project (20220310).

Data availability

Data will be made available on request.

References

Amarasinghe, I., Hernández-Leo, D., Michos, K., & Vujovic, M. (2020). An actionable orchestration dashboard to enhance collaboration in the classroom. *IEEE Transactions on Learning Technologies, 13*(4), 662–675. https://doi.org/10.1109/TLT.2020.3028597

Aydoğdu, S. (2021). A new student modeling technique with convolutional neural networks: Learnerprints. *Journal of Educational Computing Research, 59*(4), 603–619. https://doi.org/10.1177/0735633120969216

Ba, S., & Hu, X. (2023). Measuring emotions in education using wearable devices: A systematic review. *Computers & Education, 200*(1), Article 104797. https://doi.org/10.1016/j.compedu.2023.104797

Banihashem, S. K., Noroozi, O., van Ginkel, S., Macfadyen, L. P., & Biemans, H. J. (2022). A systematic review of the role of learning analytics in enhancing feedback practices in higher education. *Educational Research Review, 37*(1), Article 100489. https://doi.org/10.1016/j.edurev.2022.100489

Bravo, C., Redondo, M. A., Verdejo, M. F., & Ortega, M. (2008). A framework for process–solution analysis in collaborative learning environments. *International Journal of Human-Computer Studies, 66*(11), 812–832. https://doi.org/10.1016/j.ijhcs.2008.08.003

Cai, Z., Gui, Y., Mao, P., Wang, Z., Hao, X., Fan, X., & Tai, R. H. (2023). The effect of feedback on academic achievement in technology-rich learning environments (TREs): A meta-analytic review. *Educational Research Review, 39*(1), Article 100521. https://doi.org/10.1016/j.edurev.2023.100521

Casamayor, A., Amandi, A., & Campo, M. (2009). Intelligent assistance for teachers in collaborative e-learning environments. *Computers & Education, 53*(4), 1147–1154. https://doi.org/10.1016/j.compedu.2009.05.025

Chen, C. M., Li, M. C., Chang, W. C., & Chen, X. X. (2021). Developing a topic analysis instant feedback system to facilitate asynchronous online discussion effectiveness. *Computers & Education, 163*(1), Article 104095. https://doi.org/10.1016/j.compedu.2020.104095

Chen, C. M., Li, M. C., & Huang, Y. L. (2023). Developing an instant semantic analysis and feedback system to facilitate learning performance of online discussion. *Interactive Learning Environments, 31*(3), 1402–1420. https://doi.org/10.1080/10494820.2020.1839505

Chen, W., Tan, J. S., & Pi, Z. (2021). The spiral model of collaborative knowledge improvement: An exploratory study of a networked collaborative classroom. *International Journal of Computer-Supported Collaborative Learning, 16*(1), 7–35. https://doi.org/10.1007/s11412-021-09338-6

Chen, J., Wang, M., Kirschner, P. A., & Tsai, C. C. (2018). The role of collaboration, computer use, learning environments, and supporting strategies in CSCL: A meta-analysis. *Review of Educational Research, 88*(6), 799–843. https://doi.org/10.3102/0034654318791584

Chen, X., Xie, H., Zou, D., & Hwang, G. J. (2020). Application and theory gaps during the rise of artificial intelligence in education. *Computers & Education: Artificial Intelligence, 1*(1), Article 100002. https://doi.org/10.1016/j.caeai.2020.100002

Çini, A., Järvelä, S., Dindar, M., & Malmberg, J. (2023). How multiple levels of metacognitive awareness operate in collaborative problem solving. *Metacognition and Learning, 1*(1), 1–32. https://doi.org/10.1007/s11409-023-09358-7

Conijn, R., Martinez-Maldonado, R., Knight, S., Buckingham Shum, S., Van Waes, L., & van Zaanen, M. (2020). How to provide automated feedback on the writing process? A participatory approach to design writing analytics tools. *Computer Assisted Language Learning, 35*(8), 1–31. https://doi.org/10.1080/09588221.2020.1839503

Cress, U., Rosé, C. P., Law, N., & Ludvigsen, S. (2019). Investigating the complexity of computer-supported collaborative learning in action. *International Journal of Computer-Supported Collaborative Learning, 14*(2), 137–142. https://doi.org/10.1007/s11412-019-09305-2

Cress, U., Rosé, C., Wise, A., & Oshima, J. (2021). *International handbook of computer-supported collaborative learning.* Springer. https://doi.org/10.1007/978-3-030-65291-3

Cukurova, M., Giannakos, M., & Martinez-Maldonado, R. (2020). The promise and challenges of multimodal learning analytics. *British Journal of Educational Technology, 51*(5), 1441–1449. https://doi.org/10.1111/bjet.13015

Dierdorff, E. C., & Ellington, J. K. (2012). Members matter in team training: Multilevel and longitudinal relationships between goal orientation, self-regulation, and team outcomes. *Personnel Psychology, 65*(3), 661–703. https://doi.org/10.1111/j.1744-6570.2012.01255.x

Echeverría Rodríguez, L., & Cobos Pérez, R. (2021). The impact of a motivation booster in blended learning experiences in computer science education. *Revista EIA, 18*(35), 235–246. https://doi.org/10.24050/reia.v18i35.1320

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing, 62*(1), 107–115. https://doi.org/10.1111/j.1365-2648.2007.04569.x

Erkens, M., Bodemer, D., & Hoppe, H. U. (2016). Improving collaborative learning in the classroom: Text mining based grouping and representing. *International Journal of Computer-Supported Collaborative Learning, 11*(1), 387–415. https://doi.org/10.1007/s11412-016-9243-5

Fung, C. C., Li, J. B., Wong, K. W., & Wong, K. P. (2004). A java-based parallel platform for the implementation of evolutionary computation for engineering applications. *International Journal of Systems Science, 35*(1), 741–750. https://doi.org/10.1080/00207720412331303651

Gedrimiene, E., Celik, I., Kaasila, A., Mäkitalo, K., & Muukkonen, H. (2024). Artificial intelligence (AI)-enhanced learning analytics (LA) for supporting career decisions: Advantages and challenges from user perspective. *Education and Information Technologies, 29*(1), 297–322.

Giannakos, M. N., Sharma, K., Pappas, I. O., Kostakos, V., & Velloso, E. (2019). Multimodal data as a means to understand the learning experience. *International Journal of Information Management, 48*(1), 108–119. https://doi.org/10.1016/j.ijinfomgt.2019.02.003

Grammens, M., Voet, M., Vanderlinde, R., Declercq, L., & De Wever, B. (2022). A systematic review of teacher roles and competences for teaching synchronously online through videoconferencing technology. *Educational Research Review, 37*(1), Article 100461. https://doi.org/10.1016/j.edurev.2022.100461

Grand, J. A., Braun, M. T., Kuljanin, G., Kozlowski, S. W. J., & Chao, G. T. (2016). The dynamics of team cognition: A process-oriented theory of knowledge emergence in teams. *Journal of Applied Psychology, 101*(10), 1353–1385. https://doi.org/10.1037/apl0000136

Greller, W., & Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. *Journal of Educational Technology & Society, 15*(3), 42–57. https://www.jstor.org/stable/10.2307/jeductechsoci.15.3.42.

Guzmán-Valenzuela, C., Gómez-González, C., Tagle, A. R. M., & Lorca-Vyhmeister, A. (2021). Learning analytics in higher education: A preponderance of analytics but very little learning? *International Journal of Educational Technology in Higher Education, 18*(1), 1–19. https://doi.org/10.1186/s41239-021-00258-x

Han, J., Kim, K. H., Rhee, W., & Cho, Y. H. (2021). Learning analytics dashboards for adaptive support in face-to-face collaborative argumentation. *Computers & Education, 163*(1), Article 104041. https://doi.org/10.1016/j.compedu.2020.104041

Hmelo-Silver, C. E., & Jeong, H. (2021). An overview of CSCL methods. In U. Cress, C. Rosé, A. F. Wise, & J. Oshima (Eds.), *International handbook of computer-supported collaborative learning* (pp. 65–83). Springer. https://doi.org/10.1007/978-3-030-65291-3_4.

Holstein, K., McLaren, B. M., & Aleven, V. (2019). Co-designing a real-time classroom orchestration tool to support teacher-AI complementarity. *Journal of Learning Analytics, 6*(2), 27–52. https://doi.org/10.18608/jla.2019.62.3

Hu, L., & Chen, G. (2021). A systematic review of visual representations for analyzing collaborative discourse. *Educational Research Review, 34*(1), Article 100403. https://doi.org/10.1016/j.edurev.2021.100403

Huang, C. J., Wang, Y. W., Huang, T. H., Chen, Y. C., Chen, H. M., & Chang, S. C. (2011). Performance evaluation of an online argumentation learning assistance agent. *Computers & Education, 57*(1), 1270–1280. https://doi.org/10.1016/j.compedu.2011.01.013

Janssen, J., Erkens, G., Kanselaar, G., & Jaspers, J. G. M. (2007). Visualization of participation: Does it contribute to successful computer-supported collaborative learning? *Computers & Education, 49*(4), 1037–1065. https://doi.org/10.1016/j.compedu.2006.01.004

Järvelä, S., Kirschner, P. A., Hadwin, A., Järvenoja, H., Malmberg, J., Miller, M., & Laru, J. (2016). Socially shared regulation of learning in CSCL: Understanding and prompting individual-and group-level shared regulatory activities. *International Journal of Computer-Supported Collaborative Learning, 11*(3), 263–280. https://doi.org/10.1007/s11412-016-9238-2

Järvelä, S., & Rosé, C. (2021). Forms of collaboration matters: CSCL across the contexts. *International Journal of Computer-Supported Collaborative Learning, 16*(2), 145–149. https://doi.org/10.1007/s11412-021-09348-4

Järvenoja, H., Järvelä, S., & Malmberg, J. (2020). Supporting groups' emotion and motivation regulation during collaborative learning. *Learning and Instruction, 70*, Article 101090. https://doi.org/10.1016/j.learninstruc.2017.11.004

Jeong, H., & Hmelo-Silver, C. E. (2016). Seven affordances of computer-supported collaborative learning: How to support collaborative learning? How can technologies help? *Educational Psychologist, 51*(2), 247–265. https://doi.org/10.1080/00461520.2016.1158654

Jeong, H., Hmelo-Silver, C. E., & Jo, K. (2019). Ten years of computer-supported collaborative learning: A meta-analysis of CSCL in STEM education during 2005–2014. *Educational Research Review, 28*(1), Article 100284. https://doi.org/10.1016/j.edurev.2019.100284

Jiao, P., Ouyang, F., Zhang, Q., & Alavi, A. H. (2022). Artificial intelligence-enabled prediction model of student academic performance in online engineering education. *Artificial Intelligence Review, 55*(8), 6321–6344. https://doi.org/10.1007/s10462-022-10155-y

Jivet, I., Scheffel, M., Schmitz, M., Robbers, S., Specht, M., & Drachsler, H. (2020). From students with love: An empirical study on learner goals, self-regulated learning and sense-making of learning analytics in higher education. *Internet and Higher Education, 47*(1), Article 100758. https://doi.org/10.1016/j.iheduc.2020.100758

Kabudi, T., Pappas, I., & Olsen, D. H. (2021). AI-enabled adaptive learning systems: A systematic mapping of the literature. *Computers & Education: Artificial Intelligence, 2*(1), Article 100017. https://doi.org/10.1016/j.caeai.2021.100017

Kang, J., Zhou, Y., Rajarathinam, R. J., Tan, Y., & Shaffer, D. W. (2024). Unveiling joint attention dynamics: Examining multimodal engagement in an immersive collaborative astronomy simulation. *Computers & Education, 213*, Article 105002. https://doi.org/10.1016/j.compedu.2024.105002

Khalil, M., Prinsloo, P., & Slade, S. (2022). The use and application of learning theory in learning analytics: A scoping review. *Journal of Computing in Higher Education*, 1–22. https://doi.org/10.1007/s12528-022-09340-3

Kim, Y. R., Park, M. S., Moore, T. J., & Varma, S. (2013). Multiple levels of metacognition and their elicitation through complex problem-solving tasks. *The Journal of Mathematical Behavior, 32*(3), 377–396. https://doi.org/10.1016/j.jmathb.2013.04.002

Lan, Y. J., & Chen, N. S. (2024). Teachers' agency in the era of LLM and generative AI. *Educational Technology & Society, 27*(1), I–18. https://www.jstor.org/stable/48754837.

Li, Y., Li, X., Zhang, Y., & Li, X. (2021). The effects of a group awareness tool on knowledge construction in computer-supported collaborative learning. *British Journal of Educational Technology, 52*(3), 1178–1196. https://doi.org/10.1111/bjet.13066

Lonchamp, J. (2010). Customizable computer-based interaction analysis for coaching and self-regulation in synchronous CSCL systems. *Journal of Educational Technology & Society, 13*(2), 193–205.

Lyons, K. M., Lobczowski, N. G., Greene, J. A., Whitley, J., & McLaughlin, J. E. (2021). Using a design-based research approach to develop and study a web-based tool to support collaborative learning. *Computers & Education, 161*(1), Article 104064. https://doi.org/10.1016/j.compedu.2020.104064

Maldonado, R., Gašević, D., Echeverria, V., Nieto, G. F., Swiecki, Z., & Shum, S. B. (2021). What do you mean by collaboration analytics? A conceptual model. *Journal of Learning Analytics, 8*(1), 126–153. https://doi.org/10.18608/jla.2021.7227
Mangaroska, K., & Giannakos, M. (2019). Learning analytics for learning design: A systematic literature review of analytics-driven design to enhance learning. *IEEE Transactions on Learning Technologies, 12*(4), 516–534. https://doi.org/10.1109/TLT.2018.2868673
Martinez-Maldonado, R. (2019). A handheld classroom dashboard: Teachers' perspectives on the use of real-time collaborative learning analytics. *International Journal of Computer-Supported Collaborative Learning, 14*(1), 383–411. https://doi.org/10.1007/s11412-019-09308-z
Martinez-Maldonado, R., Elliott, D., Axisa, C., Power, T., Echeverria, V., & Buckingham Shum, S. (2022). Designing translucent learning analytics with teachers: An elicitation process. *Interactive Learning Environments, 30*(6), 1077–1091. https://doi.org/10.1080/10494820.2019.1710541
Melzner, N., Dresel, M., & Kollar, I. (2022). Examining the regulation of motivational and comprehension-related problems during collaborative learning. *Metacognition and Learning, 17*(3), 813–836. https://doi.org/10.1007/s11409-022-09316-9
Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma, G. (2009). Preferred reporting items for systematic reviews and metaanalyses: The PRISMA statement. *PLoS Medicine, 6*(7), Article e1000097. https://doi.org/10.1371/journal.pmed.1000097.t001
Noroozi, O., Alikhani, I., Järvelä, S., Kirschner, P. A., Juuso, I., & Seppänen, T. (2019). Multimodal data to design visual learning analytics for understanding regulation of learning. *Computers in Human Behavior, 100*, 298–304. https://doi.org/10.1016/j.chb.2018.12.019
Noroozi, O., Pijeira-Díaz, H. J., Sobocinski, M., Dindar, M., Järvelä, S., & Kirschner, P. A. (2020). Multimodal data indicators for capturing cognitive, motivational, and emotional learning processes: A systematic literature review. *Education and Information Technologies, 25*, 5499–5547. https://doi.org/10.1007/s10639-020-10229-w
Ouhaichi, H., Spikol, D., & Vogel, B. (2023). Research trends in multimodal learning analytics: A systematic mapping study. *Computers & Education: Artificial Intelligence, 4*, Article 100136. https://doi.org/10.1016/j.caeai.2023.100136
Ouyang, F., Dai, X., & Chen, S. (2022). Applying multimodal learning analytics to examine the immediate and delayed effects of instructor scaffoldings on small groups' collaborative programming. *International Journal of STEM Education, 9*(1), 1–21. https://doi.org/10.1186/s40594-022-00361-z
Ouyang, F., Wu, M., Zheng, L., Zhang, L., & Jiao, P. (2023). Integration of artificial intelligence performance prediction and learning analytics to improve student learning in online engineering course. *International Journal of Educational Technology in Higher Education, 20*(1), 1–23. https://doi.org/10.1186/s41239-022-00372-4
Ouyang, F., Zhang, L., Wu, M., & Jiao, P. (2024). Empowering collaborative knowledge construction through the implementation of a collaborative argument map tool. *The Internet and Higher Education, 62*(1), Article 100946. https://doi.org/10.1016/j.iheduc.2024.100946
Ouyang, F., Zheng, L., & Jiao, P. (2022). Artificial intelligence in online higher education: A systematic review of empirical research from 2011 to 2020. *Education and Information Technologies, 27*(6), 7893–7925. https://doi.org/10.1007/s10639-022-10925-9
Panadero, E., & Lipnevich, A. A. (2022). A review of feedback models and typologies: Towards an integrative model of feedback elements. *Educational Research Review, 35*(1), Article 100416. https://doi.org/10.1016/j.edurev.2021.100416
Pelletier, K., McCormack, M., Reeves, J., Robert, J., & Arbino, N. (2022). 2022 Educause horizon report: Teaching and learning edition. https://library.educause.edu/resources/2022/4/2022-educause-horizon-report-teaching-and-learning-edition. (Accessed 1 January 2023).
Peng, Y., Li, Y., Su, Y., Chen, K., & Jiang, S. (2022). Effects of group awareness tools on students' engagement, performance, and perceptions in online collaborative writing: Intergroup information matters. *The Internet and Higher Education, 53*(1), Article 100845. https://doi.org/10.1016/j.iheduc.2022.100845
Rummel, N., Walker, E., & Aleven, V. (2016). Different futures of adaptive collaborative learning support. *International Journal of Artificial Intelligence in Education, 26*(2), 784–795. https://doi.org/10.1007/s40593-016-0102-3
Salas-Pilco, S. Z., & Yang, Y. (2022). Artificial intelligence applications in Latin American higher education: A systematic review. *International Journal of Educational Technology in Higher Education, 19*(1), 1–20. https://doi.org/10.1186/s41239-022-00326-w
Sandoval, A., Gonzalez, C., Alarcon, R., Pichara, K., & Montenegro, M. (2018). Centralized student performance prediction in large courses based on low-cost variables in an institutional context. *The Internet and Higher Education, 37*(1), 76–89. https://doi.org/10.1016/j.iheduc.2018.02.002
Sedrakyan, G., Malmberg, J., Verbert, K., Järvelä, S., & Kirschner, P. A. (2020). Linking learning behavior analytics and learning science concepts: Designing a learning analytics dashboard for feedback to support learning regulation. *Computers in Human Behavior, 107*, Article 105512. https://doi.org/10.1016/j.chb.2018.05.004
Shin, Y., Kim, D., & Jung, J. (2018). The effects of representation tool (visible-annotation) types to support knowledge building in computer-supported collaborative learning. *Educational Technology & Society, 21*(2), 98–110. https://www.jstor.org/stable/26388383.
Stahl, G. (2013). Learning across levels. *International Journal of Computer-Supported Collaborative Learning, 8*(1), 1–12. https://doi.org/10.1007/s11412-013-9169-0
Stahl, G. (2016). The group as paradigmatic unit of analysis: The contested relationship of computer-supported collaborative learning to the learning sciences. In M. A. Evans, M. J. Packer, & R. K. Sawyer (Eds.), *Reflections on the learning sciences* (pp. 76–102). Cambridge University Press.
Stahl, G., Koschmann, T., & Suthers, D. (2006). Computer-supported collaborative learning: An historical perspective. In R. K. Sawyer (Ed.), *Cambridge handbook of the learning sciences* (pp. 409–426). Cambridge University Press.
Strauß, S., & Rummel, N. (2021). 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, 16*(1), 67–104. https://doi.org/10.1007/s11412-021-09340-y
Susnjak, T. (2023). Beyond predictive learning analytics modelling and onto explainable artificial intelligence with prescriptive analytics and ChatGPT. *International Journal of Artificial Intelligence in Education, 1*(1), 1–31. https://doi.org/10.1007/s40593-023-00336-3
Thompson, K., Corrin, L., & Lodge, J. M. (2023). AI in tertiary education: Progress on research and practice. *Australasian Journal of Educational Technology, 39*(5), 1–7. https://doi.org/10.14742/ajet.9251
Tissenbaum, M., & Slotta, J. (2019). Supporting classroom orchestration with real-time feedback: A role for teacher dashboards and real-time agents. *International Journal of Computer-Supported Collaborative Learning, 14*(1), 325–351. https://doi.org/10.1007/s11412-019-09306-1
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*(1), 127–156. https://doi.org/10.1007/s11412-014-9190-y
Tsai, Y. S., Mello, R. F., Jovanović, J., & Gašević, D. (2021). Student appreciation of data-driven feedback: A pilot study on OnTask. In *LAK21: 11th international learning analytics and knowledge conference* (pp. 511–517). New York, NY: Association for Computing Machinery. https://doi.org/10.1145/3448139.3448212.
Van Leeuwen, A. (2015). Learning analytics to support teachers during synchronous CSCL: Balancing between overview and overload. *Journal of Learning Analytics, 2*(2), 138–162. https://doi.org/10.18608/jla.2015.22.11
van Leeuwen, A., Janssen, J., Erkens, G., & Brekelmans, M. (2014). Supporting teachers in guiding collaborating students: Effects of learning analytics in CSCL. *Computers & Education, 79*(1), 28–39. https://doi.org/10.1016/j.compedu.2014.07.007
Vieira, C., Parsons, P., & Byrd, V. (2018). Visual learning analytics of educational data: A systematic literature review and research agenda. *Computers & Education, 122*, 119–135. https://doi.org/10.1016/j.compedu.2018.03.018
Walker, E., Rummel, N., & Koedinger, K. R. (2011). Designing automated adaptive support to improve student helping behaviors in a peer tutoring activity. *International Journal of Computer-Supported Collaborative Learning, 6*(1), 279–306. https://doi.org/10.1007/s11412-011-9111-2
Wise, A. F., Knight, S., & Shum, S. B. (2021). Collaborative learning analytics. In U. Cress, C. Rosé, A. F. Wise, & J. Oshima (Eds.), *International handbook of computer-supported collaborative learning* (pp. 425–443). Springer. https://doi.org/10.1007/978-3-030-65291-3_23.
Wise, A. F., & Schwarz, B. B. (2017). Visions of CSCL: Eight provocations for the future of the field. *International Journal of Computer-Supported Collaborative Learning, 12*(1), 423–467. https://doi.org/10.1007/s11412-017-9267-5
Wisniewski, B., Zierer, K., & Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology, 10*(1), Article 487662. https://doi.org/10.3389/fpsyg.2019.03087
Wohlin, C. (2014). Guidelines for snowballing in systematic literature studies and a replication in software engineering. In *Proceedings of the 18th international conference on evaluation and assessment in software engineering* (pp. 1–10). https://doi.org/10.1145/2601248.2601268

Wong, B. T. M., & Li, K. C. (2020). A review of learning analytics intervention in higher education (2011–2018). *Journal of Computers in Education, 7*(1), 7–28. https://doi.org/10.1007/s40692-019-00143-7

Xie, T., Liu, R., Chen, Y., & Liu, G. (2021). MOCA: A motivational online conversational agent for improving student engagement in collaborative learning. *IEEE Transactions on Learning Technologies, 14*(5), 653–664. https://doi.org/10.1109/TLT.2021.3129800

Xing, W., Zhu, G., Arslan, O., Shim, J., & Popov, V. (2023). Using learning analytics to explore the multifaceted engagement in collaborative learning. *Journal of Computing in Higher Education, 35*(3), 633–662. https://doi.org/10.1007/s12528-022-09343-0

Zhan, Z., Tong, Y., Lan, X., & Zhong, B. (2022). A systematic literature review of game-based learning in Artificial Intelligence education. *Interactive Learning Environments*, 1–22. https://doi.org/10.1080/10494820.2022.2115077

Zhang, L., Wu, M., & Ouyang, F. (2023). The design and implementation of a teaching and learning analytics tool in a face-to-face, small-sized course in China's higher education. *Education and Information Technologies, 1*(1), 1–24. https://doi.org/10.1007/s10639-023-11940-0

Zheng, L., Li, X., Zhang, X., & Sun, W. (2019). The effects of group metacognitive scaffolding on group metacognitive behaviors, group performance, and cognitive load in computer-supported collaborative learning. *The Internet and Higher Education, 42*(1), 13–24. https://doi.org/10.1016/j.iheduc.2019.03.002

Zheng, L., Long, M., Niu, J., & Zhong, L. (2023). An automated group learning engagement analysis and feedback approach to promoting collaborative knowledge building, group performance, and socially shared regulation in CSCL. *International Journal of Computer-Supported Collaborative Learning, 18*(1), 101–133. https://doi.org/10.1007/s11412-023-09386-0

Fan Ouyang is a research professor at College of Education, Zhejiang University. Her research interests are collaborative knowledge construction, learning analytics and educational data mining, and artificial intelligence in education.

Liyin Zhang is a Ph.D. student under Fan Ouyang's supervision. Her research interests are learning analytics and artificial intelligence in education.