International Society of the Learning Sciences

A Clustering Approach to Collaborative Problem Solving: Utilizing Discourse and Trace Log Data

Daeun Hong, Chen Feng, Xiaotian Zou, Sijia Huang, Cindy E. Hmelo-Silver
dh37@iu.edu, carrfeng@iu.edu, xz107@iu.edu, sijhuang@iu.edu, chmelosi@iu.edu
Indiana University
Tianshu Wang, Halim Acosta, Krista Glazewski, James Lester
Twang43@ncsu.edu, hacosta@ncsu.edu, kdglazew@ncsu.edu, lester@ncsu.edu
North Carolina State University

Abstract: The paper proposes a collaborative problem solving (CPS) clustering approach using a set of CPS indicators derived from process and discourse data to capture students' CPS practices within a CSCL context. We utilized a dataset generated from 68 middle schoolers in 17 groups interacting with a game-based CSCL environment as the clustering model input. We identify four patterns of CPS engagement and examine whether the model effectively conveys information regarding students' CPS engagement in relation to their learning performance. Lastly, we discuss the implications of the approach and its potential for further development.

Introduction

Collaborative problem solving (CPS) is identified as one of the critical skills for achieving success in the 21st century (OECD, 2017). CPS activities can positively influence student learning, engagement, and the quality of solutions to problems (Graesser et al., 2018; Jeong et al., 2019). CPS consists of two broad dimensions: social and cognitive, each involving diverse subskills (Graesser et al., 2018). The cognitive dimension pertains to problem-solving activities, while the social dimension is concerned with collaboration. CPS further includes a wide range of subskills within these two dimensions, although there are subtle differences at more detailed levels across different theoretical models and subject areas. During CPS activities, individuals must engage in multifaceted processes such as sharing and negotiating ideas and experiences, jointly coordinating behaviors, regulating learning processes, and applying social strategies to sustain positive communication (Hesse, 2015; Liu et al., 2016). Therefore, CPS is a highly complex process that necessitates that learners employ a range of skills to succeed. As a result, students may encounter various challenges while collaboratively navigating their learning as they confront authentic, ill-structured problems (Savery, 2015). Given that successfully addressing complex problems hinges on these diverse skills, substantial endeavors have been made to monitor and evaluate students' CPS engagement, with the goal of providing appropriate scaffolding to support CPS processes in response to their challenges and needs (Belland, 2017).

In collaborative game-based inquiry environments, a form of computer-supported collaborative learning (CSCL), students learn content and enhance their CPS skills as they engage with complex problems (Jeong et al., 2019; Saleh et al., 2019). Specifically, students participate in computer-mediated learning activities, interacting with the game environment and their peers to solve these problems. As technology has advanced, various types of process and discursive data have been used to monitor and assess students' CPS engagement (von Davier et al., 2017). Previous studies have measured and examined students' CPS performance by using process data, such as trace log, as evidence of CPS practices, mapping this evidence to the dimensions of CPS frameworks (Andrew-Todd & Forsyth, 2020; Andrew-Todd et al., 2023; Chang et al., 2017; Sun et al., 2022). However, due to the inherent complexity of CPS, assessing students' performance and accounting for their learning outcomes in relation to CPS practices remains challenging, especially in CSCL classrooms. Moreover, detecting lower-level dimensions of CPS, rather than focusing on higher-level aspects (i.e., social and cognitive) of CPS (Andrew-Todd & Forsyth, 2020), has been challenging, limiting the ability to understand their specific needs and provide contingent support during CPS activities. Additionally, few studies have examined the effectiveness of a CPS model that represents and monitors students' CPS engagement using both process data and in-person discourse. Therefore, to address such methodological challenges, this study aims to examine a clustering approach with a set of CPS indicators to effectively track and diagnose students' CPS behaviors in a middle school science context within a CSCL environment (von Davier et al., 2017; Andrew-Todd & Forsyth, 2020). In the methods paper, we first propose a CPS clustering model—a statistical technique that identifies subgroups based on similarities—using a set of CPS indicators derived from both process and discourse data to capture students' lower-level CPS practices within a CSCL context, providing a foundation for adaptive scaffolding. After presenting the clustering model, we examine what information it can effectively convey or misrepresent regarding students' CPS practices by addressing the following research questions:

RQ1: What types of CPS patterns does the clustering model produce as a result?
RQ2: What insights does the model potentially offer regarding student learning and their CPS practices?

1) Do the CPS patterns generated by the model provide insights into students' learning performance in relation to their CPS engagement?
2) What information might the model be able to provide, misrepresent, or overlook regarding students' CPS practices and learning performance?

Theoretical frameworks of collaborative problem solving in CSCL

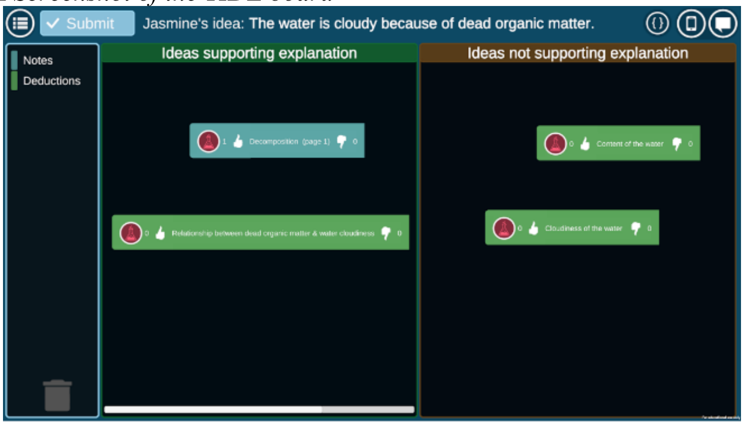
As an essential skill for the 21st century, several CPS theoretical frameworks have been proposed and applied in science domains within CSCL environments, where more than two learners collaborate to achieve a shared goal (Andrew-Todd & Forsyth, 2020; Chai et al., 2023). Specifically, Liu et al. (2016) introduced a CPS framework focused on practices from a discursive perspective, where CPS activities can be observed through face-to-face conversations and text-mediated communication. Their framework includes four major categories: (a) sharing ideas, (b) negotiating ideas, (c) regulating problem-solving, and (d) maintaining communication. Sharing ideas captures how individuals exchange task-related information and present different perspectives to others. Negotiating ideas refers to the process of reaching a consensus and building shared understanding through negotiation. Regulating problem-solving involves collectively monitoring, reflecting, planning, and regulating team discussions for effective problem-solving. The final category, maintaining positive communication, pertains to strategies for social interaction that are unrelated to the core tasks, such as fostering team rapport. Under each category, Liu et al. (2016) specified detailed observable discursive practices. Building on a review of CPS literature in CSCL and related fields, Andrew-Todd and Kerr (2019) proposed a CPS framework known as the CPS ontology. In their model, the social dimension includes: (a) maintaining communication, (b) sharing information, (c) establishing shared understanding, and (d) negotiating. Maintaining communication involves socially oriented practices that are not directly tied to tasks, such as using chat emojis, greeting teammates, and apologizing. Sharing information refers to providing task-related information and ideas. Establishing shared understanding involves integrating and consolidating knowledge that has been agreed upon. Negotiation focuses on resolving conflicts and reaching agreements. The cognitive dimension of this framework includes: (a) exploring and understanding, (b) representing and formulating, (c) planning, (d) executing, and (e) monitoring. Exploring and understanding refers to making sense of the problem, relevant content knowledge, and the task at hand. Representing and formulating involve expressing knowledge through actions and communication. Planning involves developing problem-solving strategies, while executing refers to carrying out these plans. Monitoring includes actions and communication aimed at tracking progress and assessing teammates' status. Unlike Liu et al.'s (2016) framework, Andrew-Todd and Kerr (2019) included linking CPS constructs to non-verbal indicators, such as in-game actions, which can serve as evidence of CPS performance. Despite these differences, the two frameworks share many similarities, with some subcategories being interchangeable. Indeed, Sun et al. (2020) showed promise towards the integration of various CPS frameworks into a single model, applying it within CSCL contexts to assess students' CPS competencies.

CRYSTAL ISLAND: ECOJOURNEYS

CRYSTAL ISLAND: ECOJOURNEYS was designed to help middle school students learn life science and support collaborative problem solving (CPS) within an inquiry-based game learning environment (Saleh et al., 2019). In this game, teams of three to four students who collaboratively investigate the cause of illness in tilapia fish at a local farm. The game consists of three quests, preceded by a tutorial. Initially, in each quest, students individually explore information about water quality and aquatic ecosystems by collecting notes in their in-game notebook, interacting with non-player characters, and watching videos while navigating the game's environment. Following this, they collaboratively engage in the TIDE (*Talk*, *Investigate*, *Deduce*, and *Explain*) inquiry cycle, where students prominently apply their CPS skills across the three quests.

In TIDE activities, students are required to determine whether a given claim (e.g., the water is cloudy because of dead organic matter) is supported by the information students learned during their individual investigations. While working on the task, students utilize a collaborative whiteboard, the TIDE board, which enables them to see each other placing notes on the board in real time (see Figure 1). They share their arguments about the claim by placing relevant notes as evidence into one of the columns on the board. They also use a thumbs-up or thumbs-down button to express their agreement and disagreement, respectively. Additionally, students can revisit learning materials such as their notebooks and videos to review the science content during the activities. After completing all the activities, individual students are required to submit the final solution, regarding the issues in the local aquatic ecosystem and the reasons for the tilapia fish becoming sick. Throughout the gameplay, including the CPS activities, learners communicate through in-game chat and face-to-face conversations.

Figure 1
A Screenshot of the TIDE board



Methods

Participants

In this study, 17 focal groups comprising 68 consenting middle school students in grades 6-8 participated in gameplay, averaging one hour per session. Depending on the session length, they completed all the learning activities in the game over four to six days. The classroom implementations occurred from Fall 2022 to Spring 2024 in seven science classes across four middle schools located in the Midwest and Southeastern regions of the U.S. Each focal group, consisting of four students with mixed academic performance levels, was formed by the teachers based on their observations. The students used individual laptops to collaboratively play the game. The unit of analysis focuses on the collaborative activity periods across the three quests per group, excluding the tutorial. After removing any missing data, the 17 focal groups produced a dataset from a total of 46 periods of TIDE activities for the data analysis of the study.

Data sources and analytic methods

In terms of data sources, the two primary pipelines for clustering CPS patterns include each group's discourse data and in-game actions. Specifically, for the discourse data, we used video recordings of the focal groups' interactions, including their in-person conversations and in-game chat, extracted from the game trace log data. After transcribing the in-person discussions during the 46 periods of the TIDE activities, we collected a total of 3,002 utterances, with each sentence counted as one utterance. For the in-game actions used in the model, we extracted game trace logs associated with CPS cognitive practices within the game context. These actions included reviewing notes in the notebook, reading notes on the TIDE board, and moving notes during each TIDE activity. In relation to the clustering model results (RQ1) and its examination (RQ2), we used students' pre- and post-assessment scores and their final solutions, to classify different levels of learning performance. Particularly, for RQ2-2, we utilized the focal group's video data—spanning approximately 372 minutes over the 46 periods of TIDE activities—to investigate what insights the CPS clustering model can offer.

For the analysis of RQ1 and 2, we classified the focal groups into higher-, middle-, or lower-performing categories based on their final solutions and pre- and post-assessment scores. Groups where all members demonstrated positive learning gains and explicitly linked the fish kill to aquatic ecosystem components and their interrelationships were categorized as higher performing. In contrast, we classified groups where all members exhibited negative learning gains or had post-assessment scores below 25 out of 39, and provided either superficial descriptions or explanations not covered in the game as lower performing. The remaining groups were labeled as middle-performing. The two coders reached 100% agreement in their classifications, resulting in six groups categorized as higher-performing, four as middle-performing, and seven as lower-performing.

Additionally, as input for the model in addressing RQ1 and RQ2, we converted the text-mediated discourse data and game trace logs into the required format for the CPS clustering model. Initially, two coders adapted the coding scheme by Liu et al. (2016) to fit our learning context. Using the adjusted CPS coding scheme (Table 1), they independently coded 611(20%) utterances, resulting in an inter-rater reliability of Cohen's κ = 0.84. One coder then coded the remaining utterances. We then conducted lag sequential analysis (Bakeman &

Gottman, 1997) to determine the frequency and probability of each combination of transitions from one CPS code to another (e.g., negotiating to negotiating (n_n), sharing to regulating (s_r), and out-of-domain to maintaining (ood_m)) for TIDE activity. An ANOVA was used to identify critical transitions that showed significant differences ($p < .05$) in frequency and probability between higher- and lower-performing groups. Out of 72 possible transition combinations, 8 (i.e., s_s, s_n, s_m, n_s, n_n, n_r, ood_s, and m_s) were identified. To compare and interpret the values of the critical transition indicators, we converted the values of the critical transitions to z-standardized values. In addition to pre-processing for the CPS transitions, we utilized the Linguistic Inquiry and Word Count tool (LIWC, Boyd et al., 2022) to obtain scores for use of cognitive processing words, which represent the quality of the groups' science discussions for TIDE activities. For the trace log data reflecting in-game actions, we calculated average time spent reviewing notes in the notebook, the average time spent reading notes on the TIDE board, and average frequency of moving notes during each TIDE activity.

Table 1
The Adjusted CPS Coding Scheme (With Code Abbreviations in Parentheses)

Code	CPS sub-dimension	Description
Sharing (s)	Sharing information and ideas	(Ask to) share knowledge, resources, ideas, and information related to tasks or game functions.
Negotiating (n)	Negotiating ideas	(Ask to) express agreement or disagreement with an opinion, or provide additional information to clarify, correct, or elaborate on ideas.
Regulating (r)	Regulating problem solving	Set goals and plans, monitor and reflect on the team's work, and manage frustrations and challenges.
Maintaining (m)	Maintaining positive communication	Have and maintain respectful, lively communication while supporting each other.
Out of domain (ood)	Out of domain	Utterances unrelated to the hands-on task.
Other (o)	Other	Inaudible or incomprehensible utterances, (fixing) typos, and mumbling, etc.

To address RQ2, after converting all the data, we input the converted data into the CPS model and performed latent profile analysis (Spurk et al., 2020) to cluster patterns of CPS. The number of clusters was determined based on the values of the Bayesian Information Criterion (Spurk et al., 2020), resulting in the extraction of the CPS clusters. For RQ2-1, while acknowledging the small sample size, we conducted a chi-square analysis to determine whether there was a significant association between certain CPS patterns and learning performance. Additionally, we examined the extent to which each cluster contributed to each learning performance level by computing Pearson residuals. For RQ2-2, based on the results from the previous research questions, we selected and analyzed video segments that displayed contrasting and consistent patterns relevant to learning performance. Through video analysis (Derry et al., 2010), we qualitatively investigated what insights the CPS model might provide, as well as what it might misrepresent or overlook, by comparing the CPS practices in the videos to those informed by the model.

Design of a clustering model for CPS engagement

In the previous study by Hong et al. (2023), we examined students' CPS engagement through text-mediated communication data, including in-game chat and face-to-face conversations. Given the communication-heavy nature of our learning environment, we adopted the CPS framework by Liu et al. (2016), which primarily focuses on CPS from a discursive perspective. However, the study suggested that students' in-game actions might provide insights into the cognitive aspects of their CPS practices, such as making sense of scientific concepts and constructing responses to questions. This finding led us to combine the CPS framework by Andrew-Todd and Kerr (2019), which incorporates both non-verbal and verbal indicators of social and cognitive CPS engagement, with Liu et al.'s (2016) framework. To integrate the frameworks, we first aligned the sub-facets of Liu et al.'s (2016) framework with those under the social dimension from Andrew-Todd and Kerr's (2019) model. 'Sharing information', 'negotiation', and 'maintaining communication' correspond to 'sharing ideas', 'negotiating ideas', and 'maintaining positive communication', respectively. Additionally, 'regulating problem-solving' from Liu et al.'s (2016) framework encompasses planning, monitoring, and executing CPS practices. Thus, each existing CPS code from Liu et al.'s (2016) framework was integrated with its corresponding sub-dimension(s) from Andrew-Todd and Kerr's (2019) framework. We then identified two sub-facets from Andrew-Todd and Kerr's (2019) framework—'exploring and understanding' and 'representing and formulating'—that were not captured in Liu et al.'s (2016) framework. These were added as new CPS sub-facets to the integrated framework. Following this, we determined which indicators could represent each integrated CPS code. While none of the social dimension CPS codes could be mapped to non-verbal indicators, the cognitive dimension CPS codes were mapped to non-verbal indicators, specifically in-game actions. As shown in Table 2, we developed a new integrated CPS framework that measures CPS competencies through both verbal and non-verbal indicators, tailored to the game environment.

Table 2
An Integrated CPS Framework for [GAME]

Dimension	Source	CPS sub-dimension	Integrated CPS code	Measure
Social	Andrew-Todd & Kerr (2019) Liu et al. (2016)	Sharing information; Sharing ideas	Sharing (s)	Discourse
	Andrew-Todd & Kerr (2019) Liu et al. (2016)	Establishing shared understanding, Negotiation; Negotiating ideas	Negotiating (n)	Discourse
	Andrew-Todd & Kerr (2019) Liu et al. (2016)	Planning, Executing, Monitoring; Regulating problem solving	Regulating (r)	Discourse
	Andrew-Todd & Kerr (2019) Liu et al. (2016)	Maintaining communication; Maintaining positive communication	Maintaining (m)	Discourse
	N/A	NA	Out-of-domain (ood)	Discourse
	NA	NA	Other (o)	Discourse
Cognitive	Andrew-Todd & Kerr (2019)	Exploring & Understanding	Understanding	Actions
	Andrew-Todd & Kerr (2019)	Representing & Formulating	Representing	Actions

Furthermore, we determined specific inputs for each code in the integrated model to cluster CPS engagement patterns. A previous study suggested that aggregate counts of each CPS code were insufficient to capture the dynamics of turn-taking, transactive conversations, and their quality, which may have significant implications for understanding CPS engagement (Hong et al., 2023). Therefore, this study focused on analyzing transitions from one CPS code to another within the social dimension of CPS practices using lag sequential analysis (Bakeman & Gottman, 1997), a method that helps understand the dynamic features of communication patterns (Chang et al., 2017). As mentioned earlier, we obtained the frequency and probability of each combination of CPS transitions through lag sequential analysis, then identified critical transitions using ANOVA. Additionally, we identified in-game actions corresponding to the cognitive CPS codes, 'understanding' and 'representing,' and extracted relevant game trace logs: the duration of reviewing notes in the Notebook and reading notes placed by students on the TIDE board (for understanding), and the frequency of moving notes during TIDE activities (for representing). Finally, to capture the quality of students' scientific conversations during the activities, we used scores for cognitive process words obtained from the LIWC tool. As a result, and as shown in Table 3, we designed the clustering model, which requires the following input data to represent and monitor students' CPS engagement.

Table 3
An Overview of the Inputs to the CPS Cluster Model

Aspect	Dimension	CPS code	Data	Input
Quantity	Social	Sharing Negotiating Regulating Maintaining Out-of-domain Other	Coded discourse data	Transitions between CPS codes: · Frequency: s_s, s_n, n_s, n_n, n_r, ood_s · Probability: s_m, m_s
	Cognitive	Understanding	Game trace logs	Duration of reviewing notes in Notebooks; Duration of reading TIDE notes

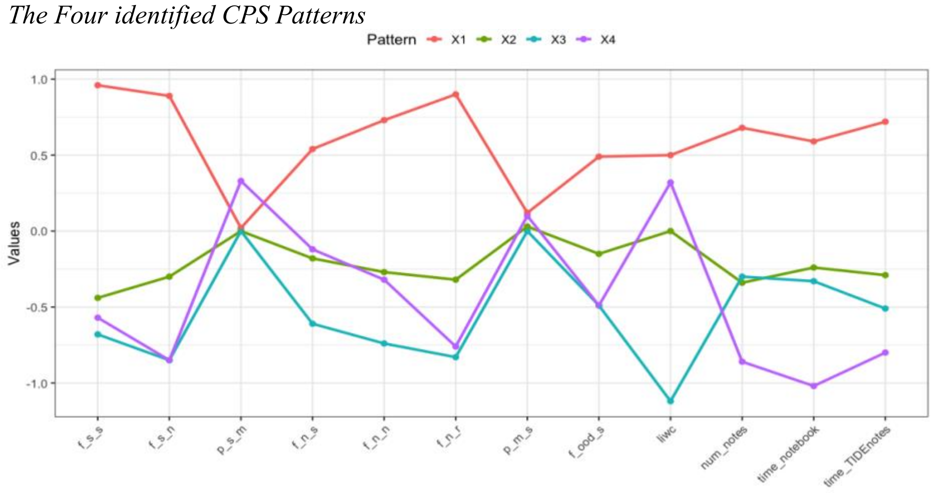
		Representing		Frequency of placing, moving, and removing TIDE notes
Quality	Social	NA	Discourse data	LIWC score of using cognitive process words (Boyd et al., 2022)

Results

RQ1: CPS patterns revealed through the clustering model

Using the indicators proposed earlier, we identified four distinct patterns of student CPS engagement by selecting the best fitting latent profile analysis model (see Figure 2). The first pattern (red line, *active knowledge building pattern*), which was of medium size (n_1 = 16, 35%), exhibited a high frequency of sharing and negotiating transitions, substantial use of learning materials, and high scores for cognitive process words. This pattern likely reflects CPS engagement characterized by active science discussions with extensive use of learning objects. The second pattern (green line, *low-to-average CPS pattern*), the largest (n_2 = 20, 43%), demonstrated a lower frequency of sharing and negotiating transitions, less use of learning materials, but an average quality of discussion. It may represent an inactive CPS pattern, with less engagement in understanding and group discussions. The third pattern (blue line, *minimal CPS pattern*), smaller in size (n_3 = 8, 17%), was characterized by low frequencies of sharing and negotiating transitions, low scores for cognitive processing words, and minimal use of learning materials. This may indicate low-quality science discussions and minimal in-game behaviors for understanding and representation. The fourth pattern (purple line, *quality discussion with low CPS pattern*), the smallest (n_4 = 2, 4%), showed relatively higher probabilities of the s_m transition, moderate-quality discussions, and minimal use of learning materials. Overall, the clustering model revealed CPS patterns that potentially represent diverse verbal and behavioral CPS engagement conditions, each with varying social and cognitive needs.

Figure 2
The Four identified CPS Patterns



RQ2: Potential insights into student learning and CPS engagement

The second research question is addressed by answering two sub questions: one regarding insights into student learning performance (RQ2-1), and the other concerning the ground truth of the clustering model (RQ2-2). For RQ2-1, we analyzed the relationships between the clustered patterns and students' learning performance to assess whether the clustering model for CPS engagement could provide insights into students' learning. While acknowledging the small sample size, we conducted a chi-square test to examine significant differences in CPS patterns, revealing a notable difference between the higher- and lower-performing groups, χ^2 (3, n=36) = 8.276, $p < .05$. We further investigated which patterns were more strongly associated with higher or lower achievement by computing Pearson residuals to assess the extent of contribution of each cluster to higher and lower performance. The findings indicated that the first (*active knowledge building*) pattern, with frequent sharing, negotiating transitions, and use of learning materials (53.8%), and the third (*minimal CPS*) pattern, with low sharing, negotiating transitions, minimal cognitive words, and use of materials (43.2%), were moderately to highly associated with higher and lower performance, respectively. In contrast, the other patterns—limited sharing, negotiating, and material use but average discussion quality; and higher sharing-to-maintaining transitions with moderate discussions and some material use—had only marginal links to learning performance. These results may offer implications for how students should be scaffolded for successful learning. These results suggest that the CPS patterns revealed by the clustering model may provide valuable insights into students' learning performance in CSCL classrooms.

To evaluate the ground truth of the clustering model, we qualitatively compared the CPS patterns generated by the model to what actually happened in the classrooms, using video analysis to examine the effectiveness of the clustering model for CPS engagement. We reviewed 16 episodes from TIDE activities where groups exhibited either consistent or contrasting results between their learning performance and the CPS pattern identified by the model. Specifically, we analyzed eight episodes from four higher- and lower-achieving groups that also displayed the first and third pattern revealed by the clustering model, respectively. Following this, we examined eight additional episodes from three groups that displayed patterns only marginally linked to higher performance yet were classified as higher-achieving, or that showed patterns closely associated with higher performance but were classified as lower-performing. Through this comparison, we identified several recurring patterns that highlighted what information the model might be able to provide, misrepresent, or overlook in relation to students' CPS engagement.

In terms of the model's affordances, the results indicated that the CPS patterns effectively captured students' engagement in CPS, which likely contributes to higher and lower achievement. For instance, during the TIDE activity, students from higher-achieving groups ensured that everyone in their group agreed on the placement of each note by actively inviting their peers into the discussion, sharing and negotiating ideas with further elaboration, and opening and reading each note while checking others' opinions. They also collectively supported a group member in constructing a response and collaboratively corrected answers, leading to active transitions from sharing to negotiating, negotiating to regulating, and greater use of cognitive processing words. These actions characterize the first pattern, marked by a high frequency of activities aimed at building understanding and formulating answers. In other episodes, students in higher-performing groups read the notes on the board and revisited their notebooks to verify whether each note was correctly placed and to find supporting evidence. They then expressed whether they agreed or disagreed with the placement of the note by hitting the '(dis)agree' button or making utterances. Such behaviors resulted in frequent use of learning materials and quality discussions, characterized by active sharing, negotiation, and regulation during the TIDE activity. These actions and utterances are effectively represented in the first CPS pattern, which aligns with their learning performance. In contrast, in cases where a lower-performing group exhibited the third pattern, students worked individually on the TIDE activity without communicating and attempted to move forward quickly due to tension among them. This resulted in minimal contributions to the science discussions and few actions related to CPS engagement, which aligns with the characteristics of the third pattern. Thus, the CPS patterns identified by the model may represent how students engage in CPS activities and inform potential needs based on their specific patterns.

However, some episodes raised concerns about the clustering model's potential to misrepresent student engagement in CPS activities. Specifically, the model could not determine whether students' actions were part of intentional CPS processes or simply random. For example, in a group of lower-performing students, they randomly opened notes on the TIDE board and made utterances like 'What is the content of water?' and 'Just put water temp,' without providing any justification to complete the task quickly. Furthermore, the model was ineffective in distinguishing whether verbal and non-verbal actions constituted collective contributions. In one TIDE episode, a student observed the group's work and made consecutive comments such as 'That's good,' 'No, that's not good,' 'Move this to another column,' and 'I don't agree,' while other students contributed few comments, resulting in a lack of elaboration or transactive conversations. Despite this, those verbal actions were still reflected in the relatively higher use of cognitive words, as well as increased instances of sharing, negotiating, and regulating, all of which characterize the first pattern. Additionally, the clustering model overlooked instances where students were encouraged to use the in-game chat rather than in-person conversations. For example, due to restrictions on in-person discussions and the burden of typing, students from two higher-performing groups communicated much less during the CPS activities. This led to the display of the second, third, and fourth patterns across the TIDE activities, characterized by relatively lower use of cognitive processing words and a reduced frequency of sharing, negotiating, and regulating. Thus, in such cases, the clustering model might inaccurately suggest that a group did not exhibit sufficient verbal and non-verbal actions that are possibly conducive to quality CPS engagement. This suggests that processing data from groups with distinct communication conditions through a single clustering model may not effectively inform students' CPS engagement.

To summarize, in RQ2, we found that the clustering model has the potential to provide insights into students' learning performance and to identify meaningful CPS patterns associated with that performance through the clustered CPS patterns. Furthermore, the video analysis revealed that while the clustering model was effective in providing information about students' meaningful CPS patterns related to their learning performance, there are instances where it may overlook contextual information and misrepresent their CPS engagement.

Discussion and conclusion

The current study proposed a CPS clustering model and evaluated the effectiveness of a set of its indicators to monitor and assess students' CPS. The results suggest that the integrated framework provides deeper insights into students' CPS engagement by employing both verbal and non-verbal measures that adapt to the learning environment. This approach helps address gaps in the current CPS frameworks, which do not fully leverage the affordances of different frameworks due to their separate use (Sun et al., 2020), offering valuable implications for researchers studying CPS. Furthermore, tracking CPS transitions seems to be a useful approach for capturing extended science discussions, particularly those characterized by idea sharing and negotiation, rather than relying solely on aggregate counts of CPS practices (Hong et al., 2023). This transition approach can potentially offer meaningful insights into turn-taking and transactive conversations between students for researchers investigating productive CPS practices in CSCL environments. The study also indicates that the clustering model effectively reflects diverse engagement conditions and informs lower-level CPS practices related to learning performance, helping educators identify students' complex needs based on their CPS patterns during activities. It has the potential to address the challenges of measuring students' CPS, which is inherently complex and difficult to assess, especially in CSCL contexts. Despite the promising implications, the model still requires refinement to better distinguish between intentional contributions and random in-game actions. This suggests that inputting a sequence of in-game actions, rather than individual actions in isolation, could help differentiate whether an action is an intentional contribution. For instance, students wanting to proceed quickly might open notes successively, hit the 'agree' button on all of them, and send a chat saying, 'let's move on,' whereas students properly contributing to CPS tasks are likely to examine the notes and make utterances to express or share ideas. This way, analyzing the sequence of in-game actions may provide clearer insights into their CPS engagement. Additionally, as articulated in the results, clustering may be less reliable when students rely solely on in-game chat for communication during CPS activities, suggesting that incorporating data across multiple modalities may improve the clustering model's ability to better capture students' CPS engagement in CSCL classroom settings. We acknowledge this limitation of the study that the small sample size may violate the statistical assumptions for chi-square tests and the clustering analysis, necessitating cautious interpretation of the results. Further research with larger samples is needed to validate our findings. However, given the inherent challenge of obtaining large samples within the CSCL context, conducting a clustering analysis with ground truth verification through video analysis may be a viable approach to examine students' CPS engagement.

In conclusion, this study contributes a structured methodological approach for identifying patterns of CPS engagement within CSCL contexts. The proposed clustering approach with a set of indicators provides a replicable method for monitoring students' CPS practices, offering insights that researchers can adapt to assess and better understand CPS patterns. Beyond its immediate analytical benefits, this approach enables the diagnosis of students' specific needs during CPS tasks, allowing for context-specific support contingent on their learning conditions. Specifically, the model helps identify whether students engage in in-depth negotiation processes through sharing and negotiating transitions, whether they effectively utilize learning materials for understanding, as indicated by in-game actions, and whether their discussions maintain quality or consist of random talk, measured by quality indicators such as LIWC scores. Based on these insights, scaffolds can be designed to promote extended transactive conversations, encourage knowledge building through use of learning materials, or refocus discussions on hands-on tasks as needed. This research thus could serve as a gateway for designing and delivering appropriate scaffolding tailored to students' diverse needs for successful learning in CSCL contexts.

References

Andrews-Todd, J., & Forsyth, C. M. (2020). Exploring social and cognitive dimensions of collaborative problem solving in an open online simulation-based task. *Computers in human behavior*, *104*, 105759.

Andrews-Todd, J., Jiang, Y., Steinberg, J., Pugh, S. L., & D'Mello, S. K. (2023). Investigating collaborative problem solving skills and outcomes across computer-based tasks. *Computers & Education*, *207*, 104928.

Andrews-Todd, J., & Kerr, D. (2019). Application of ontologies for assessing collaborative problem solving skills. *International Journal of Testing*, *19*(2), 172-187.

Bakeman, R., & Gottman, J. M. (1997). *Observing interaction: An introduction to sequential analysis*. Cambridge university press.

Belland, B. R., Walker, A. E., Kim, N. J., & Lefler, M. (2017). Synthesizing results from empirical research on computer-based scaffolding in STEM education: A meta-analysis. *Review of Educational Research*, *87*(2), 309-344.

Boyd, R. L., Ashokkumar, A., Seraj, S., & Pennebaker, J. W. (2022). The development and psychometric properties of LIWC-22. *Austin, TX: University of Texas at Austin*, *10*.

Chai, H., Hu, T., & Wu, L. (2023). Computer-based assessment of collaborative problem solving skills: A systematic review of empirical research. *Educational Research Review*, 100591.

Chang, C. J., Chang, M. H., Chiu, B. C., Liu, C. C., Chiang, S. H. F., Wen, C. T., ... & Chen, W. (2017). An analysis of student collaborative problem solving activities mediated by collaborative simulations. *Computers & Education*, *114*, 222-235.

von Davier, A. A., Hao, J., Liu, L., & Kyllonen, P. (2017). Interdisciplinary research agenda in support of assessment of collaborative problem solving: Lessons learned from developing a collaborative science assessment prototype. *Computers in Human Behavior*, *76*, 631-640.

Hong, D., Feng, C., Zou, X., Hmelo-Silver, C. E., Glazewski, K., Wang, T., ... & Lester, J. (2024). Examining Coordinated Computer-Based Fixed and Adaptive Scaffolds in Collaborative Problem-Solving Game Environments. In *Proceedings of the 17th International Conference on Computer-Supported Collaborative Learning-CSCL 2024, pp. 43-50*. International Society of the Learning Sciences.

Derry, S. J., Pea, R. D., Barron, B., Engle, R. A., Erickson, F., Goldman, R., ... & Sherin, B. L. (2010). Conducting video research in the learning sciences: Guidance on selection, analysis, technology, and ethics. *The journal of the learning sciences*, *19*(1), 3-53.

Graesser, A., Foltz, P. W., Rosen, Y., Shaffer, D. W., Forsyth, C., & Germany, M. L. (2018). Challenges of assessing collaborative problem solving. *Assessment and teaching of 21st century skills: Research and applications*, 75-91.

Hesse, F., Care, E., Buder, J., Sassenberg, K., & Griffin, P. (2015). A framework for teachable collaborative problem solving skills. *Assessment and teaching of 21st century skills: Methods and approach*, 37-56.

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*, 100284.

Kerr, D., Andrews, J. J., & Mislevy, R. J. (2016). The in-task assessment framework for behavioral data. *The Wiley handbook of cognition and assessment: Frameworks, methodologies, and applications*, 472-507.

Liu, L., Hao, J., von Davier, A. A., Kyllonen, P., & Zapata-Rivera, J. D. (2016). A tough nut to crack: Measuring collaborative problem solving. In *Handbook of research on technology tools for real-world skill development* (pp. 344-359). IGI Global.

OECD. (2017). PISA 2015 collaborative problem solving framework. *PISA 2015 assessment and analytical framework: Science, reading, mathematic, financial literacy and collaborative problem solving*, 131-188.

Saleh, A., Yuxin, C., Hmelo-Silver, C. E., Glazewski, K. D., Mott, B. W., & Lester, J. C. (2020). Coordinating scaffolds for collaborative inquiry in a game-based learning environment. *Journal of research in science teaching*, *57*(9), 1490-1518.

Savery, J. R. (2015). Overview of problem-based learning: Definitions and distinctions. In A. Walker, H. Leary, C. E. Hmelo-Silver, & P. A. Ertmer, P. A. (Eds.), *Essential readings in problem-based learning: Exploring and extending the legacy of Howard S. Barrows* (pp. 5–15). West Lafayette, IN: Purdue University Press.

Spurk, D., Hirschi, A., Wang, M., Valero, D., & Kauffeld, S. (2020). Latent profile analysis: A review and "how to" guide of its application within vocational behavior research. *Journal of vocational behavior*, *120*, 103445.

Sun, C., Shute, V. J., Stewart, A., Yonehiro, J., Duran, N., & D'Mello, S. (2020). Towards a generalized competency model of collaborative problem solving. *Computers & Education*, *143*, 103672.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Award No. DRL-2112635. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.