


RESEARCH ARTICLE

Open Access



Multimodal learning analytics of collaborative patterns during pair programming in higher education

WeiQi Xu¹, Yajuan Wu² and Fan Ouyang^{1*} 

*Correspondence:
fanouyang@zju.edu.cn

¹ College of Education, Zhejiang University, #866, Yuhangtang Rd., Hangzhou 310058, Zhejiang, China

² Teachers College, Columbia University, New York, USA

Abstract

Pair programming (PP), as a mode of collaborative problem solving (CPS) in computer programming education, asks two students work in a pair to co-construct knowledge and solve problems. Considering the complex multimodality of pair programming caused by students' discourses, behaviors, and socio-emotions, it is of critical importance to examine their collaborative patterns from a holistic, multimodal, dynamic perspective. But there is a lack of research investigating the collaborative patterns generated by the multimodality. This research applied multimodal learning analytics (MMLA) to collect 19 undergraduate student pairs' multimodal process and products data to examine different collaborative patterns based on the quantitative, structural, and transitional characteristics. The results revealed four collaborative patterns (i.e., a consensus-achieved pattern, an argumentation-driven pattern, an individual-oriented pattern, and a trial-and-error pattern), associated with different levels of process and summative performances. Theoretical, pedagogical, and analytical implications were provided to guide the future research and practice.

Keywords: Collaborative problem solving, Computer-supported collaborative learning, Pair programming, Computer programming education, Collaborative pattern, Multimodal learning analytics

Introduction

Grounded upon the sociocultural perspective of learning (Vygotsky, 1978), collaborative problem-solving (CPS) focuses on group members' knowledge construction and meaningful practices through continuous interactions and idea improvement with the technological and pedagogical supports (Hmelo-Silver & DeSimone, 2013; Stahl, 2009). Pair programming (PP), as a mode of CPS in computer programming education, asks students work together to solve challenging programming tasks, improve computational thinking, and enhance real-world problem-solving ability (Beck & Chizhik, 2013; Chittum et al., 2017; Sun et al., 2020). However, PP is a complex phenomenon, in which multiple modals (e.g., communication, behavior, socio-emotion, etc.) interact constantly to form different collaborative patterns and finally influence the quality of collaboration (Stahl & Hakkarainen, 2021). Considering the complex factors that may influence PP, it is

necessary to investigate the collaborative patterns of PP as well as their associations with the collaborative quality. Recently, some research has explored students' collaborative patterns in CPS (e.g., Han & Ellis, 2021; Lin et al., 2014; Webb et al., 2021), but results varied regarding the relations between students' collaborative patterns and the quality of collaboration. More importantly, we found that most of the previous works merely analyzed the single dimension (e.g., cognitive process, interactive type) of CPS, but rarely examined the dynamic and temporal characteristics formed through multimodality during collaboration, which might lead to an incomplete understanding of the complexity in collaboration. To fill this gap, this research collected the multimodal process-oriented data (including verbal audios, computer screen recordings, facial expression recordings) and programming products data during students' PP in higher education and utilized multimodal learning analytics (MMLA) to detect and analyze students' collaborative patterns. Specifically, we identified clusters based on the assessment of collaborative processes and final products, and further examined the quantitative, structural, and transitional characteristics of different clusters to reveal the collaborative patterns. Based on the results, we provided theoretical, pedagogical, and analytical implications to promote future practice and research.

Literature review

Grounded upon the social, cultural, and situated perspectives of learning (Vygotsky, 1978), collaborative problem-solving (CPS) emphasizes students collaborate together to solve ill-structured problems, construct knowledge, and achieve shared goals (Damon & Phelps, 1989; Dillenbourg, 1999). Compared to the instructor-centered learning mode, CPS aims to achieve an active and constructive learning process through students' mutual interaction and knowledge co-construction (Brown et al., 1989; O'Donnell & Hmelo-Silver, 2013). Pair programming (PP), as a mode of CPS in computer programming education, requires two students engage in a coordinated way to solve programming problems and complete complex programming tasks (Bryant et al., 2006; Denner et al., 2021). Recently, PP has been widely used as a learning approach in higher education to promote active learning (Hawlitschek et al., 2022). PP emphasizes that knowledge is not considered as predefined and structural information delivered from instructors but is explored and constructed by students during the collaborative process of programming and debugging (Sun et al., 2020). Moreover, empirical studies have indicated that PP has potential in arousing novice learners' motivation and interests in computer science (Chittum et al., 2017), fostering their computational thinking skills (Romero et al., 2017), and improving their problem-solving abilities in reality (Beck & Chizhik, 2013).

However, PP is a complex phenomenon that involves multimodal interaction and coordination between individual student, student group, learning environment, and the knowledge artefact (Stahl & Hakkarainen, 2021). Specifically, the multimodality can be reflected through student pair's communication (Barron, 2000; Ouyang & Xu, 2022), behavior (Stahl, 2017), emotion (Kwon et al., 2014), interaction (Zemel & Koschmann, 2013), etc. Furthermore, the multimodality emerges during the collaboration with complex, multilevel, multilayered characteristics, which may influence the quality of collaborative learning (Byrne & Callaghan, 2014; Hilpert & Marchand, 2018). However, previous empirical research varied about the relations between

students' collaborative patterns and the quality of collaboration (e.g., Han & Ellis, 2021; Lin et al., 2014; Webb et al., 2021). For instance, Lin et al. (2014) detected 45 college students' CPS patterns in online forum based on their cognitive engagement; the manipulation-centered pattern demonstrated a deeper cognition of students in collaboration, while the discussion-centered pattern appeared more off-topic discussions. Webb et al. (2021) identified 45 students' collaborative patterns in the third-grade mathematics course based on their interaction characteristics. There were groups that took turns to initiate a strategy, groups with students that generated their own strategies, and groups where one student took responsibility to generate the strategies. The results indicated that no single pattern was better than other patterns for leading students' success in collaboration. Moreover, these works mostly focused on the single aspect (e.g., cognitive process, interactive type) of CPS without considering the complexity, multimodality, and dynamics of collaboration, which might cause incomprehensive understandings of the collaborative patterns (Borge & Mercier, 2019). Overall, exploring collaborative patterns in PP, especially from a multimodal, dynamic, holistic perspective, is necessary to help researchers, instructors, and students unfold the complex factors that influence the collaborative quality as well as how they influence (Lu & Churchill, 2014; Perera et al., 2009).

From an analytical perspective, due to the complexity and multimodality of CPS, multidimensional, temporal, and fine-grained approaches are called for exploring students' collaborative patterns in computer programming education. Multimodal learning analytics (MMLA), as a new trend of learning analytics, leverage advances in multimodal data (e.g., speech, eye gaze, heart rate, body movement data) to capture and mining learning process and to address the challenges of investigating multiple, complex learning-relevant constructs in learning scenarios (Mu et al., 2020; Ochoa & Worsley, 2016; Wiltshire et al., 2019). Recently, relevant research has applied MMLA to reveal the complex, multimodal, and dynamic characteristics of CPS. For example, Sun et al., (2021) utilized discourses analysis, click stream analysis, and video analysis to analyze 63 junior high school students' discourses, behaviors, and perceptions during collaborative programming. Kawamura et al., (2021) modeled 48 students' wakefulness states on e-learning platforms and further detected drowsy students according to their multimodal data (i.e., face recognition, seat pressure, and heart rate). Wiltshire et al. (2019) collected multimodal data (i.e., gesture, speech, mouse and keyboard movement) from 42 pairs of undergraduate students and used growth curve modelling to investigate how students' multimodal movement coordination dynamically changed during collaboration. Overall, compared to traditional statistical analysis (e.g., questionnaire data, performance assessment data), MMLA has the potential to reveal the complex, multimodal, dynamic collaborative patterns in PP from a multidimensional, temporal, and fine-grained perspective.

To address these gaps, the current study applied MMLA to examine students' collaborative patterns in a face-to-face, computer-supported PP environment in higher education. Specifically, we collected students' multimodal process-oriented data (including verbal audios, computer screen recordings, facial expression recordings) and programming products data. We proposed an analytical framework that integrated MMLA methods to identify students' collaborative clusters in PP and further revealed the characteristics of clusters. Specifically, two main research questions were proposed:

RQ1: *What clusters can be detected based on the process and summative assessment during the PP process?*

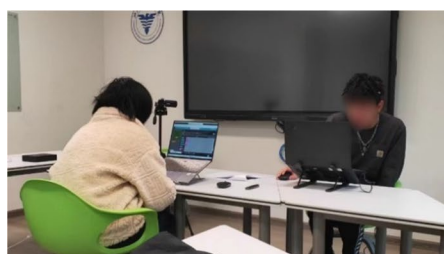
RQ2: *What are the collaborative patterns of different clusters in terms of multimodal learning analytics of process data?*

Methodology

Research context, participants, and programming procedures

The participants were 40 undergraduate students (23 males, 17 females) without prior programming foundation or experience. 20 pairs (2 students/group) were randomly assigned. Specifically, the 20 pairs included 5 male-only pairs, 6 female-only pairs, and 9 mixed pairs. The research dataset consisted of 19 datasets; data from one pair (i.e., a mixed pair) was damaged, which was excluded in this research. The research environment was a computer-supported collaborative problem solving activity. Two students in the same group sat opposite to each other and controlled a computer individually (see Fig. 1a). The computer screens were connected and shared by a remote screen control software. Student groups were asked to collaborate and learn programming on an online programming platform *Minecraft Hour of code* (<https://code.org/minecraft>) (see Fig. 1b). The platform is designed for novice programming learners with gamification and graphical programming.

Two sections were designed to support student pairs' PP process (each section lasted 25 min). In the first section, students watched the instructional videos and learned to use the coding blocks (i.e., loop, if) on the platform by completing a series of programming tasks together. In the second section, group members collaborated to complete a final programming task within 25 min by using the coding skills they had learned. The final programming task included two requirements: (1) creating a five-by-five brick building with at least four bricks over water, and (2) the foundation of the building was first constructed with boulders and then with woods. Pairs were asked to use at least two *loop* blocks, two *if* blocks, one *loop-if* nested block, and less than 30 coding blocks to complete the above task requirements. During the final task, both students had rights to control and operate the platform. All participants signed the consent forms and agreed to participate in the research.



(a) The face-to-face, computer-supported pair programming environment



(b) The *Minecraft's Hour of code* platform

Fig. 1 The research context

Data collection and dataset

The research dataset consisted of 19 datasets collected from 19 pairs of participants. This research collected the multimodal process-oriented data and programming product data of student pairs through two ways. First, video recorders (with audio) were used to capture student pairs' verbal communications and facial expressions. Second, computer screen videos (with audio) were recorded to capture student pairs' behavioral operations on the platform as well as their final programming products. Each dataset included audio recordings of verbal communication data of pairs (about 475 min in total), computer screen recordings of click stream data (about 475 min in total), video recordings of facial expression data (about 475 min in total), and the final products of pair programming task data.

The analytical framework, procedures and methods

An overall analytical framework was proposed to examine the multimodal characteristics of collaborative patterns. The framework included the first step of the assessment and clustering as well as the second step of collaborative pattern analysis. In the first step of assessment and clustering, K-means clustering was conducted to detect the collaborative clusters based on student pairs' process and summative assessment. In the second step of collaborative pattern analysis, Quantitative content analysis (QCA), click stream analysis (CSA), and video analysis (VA) were used to analyze student pairs' *verbal communication*, *operational behavior*, and *facial expression* dimensions. Further, statistical analysis (SA), epistemic network analysis (ENA), and process mining (PM) were used to examine the verbal communication, operational behavior, and facial expression dimensions, in order to reveal the *quantitative*, *structural*, and *transitional* characteristics of different clusters.

Assessment and clustering

First, *process assessment* was conducted based on the video recording of PP processes. Based on a previously validated assessment framework (Meier et al., 2007), *process assessment* was conducted in terms of nine dimensions, including (1) sustaining mutual understanding, (2) dialogue management, (3) information pooling, 4) reaching consensus, 5) task division, 6) time management, (7) technical coordination, 8) reciprocal interaction, and 9) individual task orientation (see Table 1). Specifically, a three-level assessment framework (1=almost not, 3=partially, 5=completely) was used to measure the collaborative quality during students' PP process. Two raters completed student pair's process assessment. Two raters watched the video recordings and rated 30% of the dataset independently, and then discussed to resolve the differences between them. Finally, they rated the other data independently and cross-checked each other's rating results. The inter-rater reliability with the Krippendorff's (2004) alpha reliability was 0.892.

Second, *summative assessment* was conducted to measure the final products of PP. Drawing from the previous relevant literature (Wang et al., 2021; Xu et al., 2022; Zheng et al., 2022), we proposed a three-level summative assessment framework (1=low, 3=medium, 5=high), including two dimensions of problem solving and coding skill

Table 1 The process assessment framework of collaborative quality (Meier et al., 2007)

Dimension	Rating rules		
	Low (1 point)	Medium (3 points)	High (5 points)
1. Sustaining mutual understanding	Students never or rarely sought peer feedback	Students tried to seek peer feedback, but failed to achieve mutual understanding	Students frequently clarified and elicited peer feedback to achieve mutual understanding
2. Dialogue management	Students' turn-taking was always confused due to the overlaps or chaos in dialogues	Students' turn-taking was sometimes fluent, but there was still overlaps or chaos	Students' turn-taking was always fluent by means of questions or explicit handovers in dialogues
3. Information pooling	Students did not gather and share enough information	Students shared enough information, but sometimes it was not task-relevant	Students gathered and shared as much task-relevant information as possible
4. Reaching consensus	Students failed to reach consensus	Students could reach consensus, yet lacked critical discussion and evidence exchange	Students reached consensus based on deep discussions and evidence-based arguments
5. Task division	Students did not divide the task into subtasks	Students tried dividing the task into subtasks, but the goals and plans were unclear	Students divided the task into subtasks appropriately with explicit goals and plans
6. Time management	Students failed to monitor or manage their time	Students managed the time but did not consciously monitor the remaining time	Students continually managed the time and monitored the remaining time based on progress
7. Technical coordination	Students did not master the basic operation to reach technical coordination	Students mastered the technical operations, but did not take turns to coordinately operate the platform	Students coordinated with each other and took turns to operate the online platform
8. Reciprocal interaction	Students failed to form respectful and supportive interaction	Students basically respected each other, yet one-side dominant behaviors still existed	Students respected each other equally and encouraged one another to make contributions
9. Individual task orientation	Both students showed little interests in the task and usually became distracted	One student concentrated on the task, while the other usually became distracted	Both students focused on the task at most of the time and avoided distractions

Table 2 The summative assessment framework of collaborative product

Dimension	Sub-dimension	Rating rules		
		Low (1 point)	Medium (3 points)	High (5 points)
1. Problem solving (Zheng et al., 2022)	1a. Finish time	The students failed to complete the task in 25 min	The students completed the task in 20–25 min	The students completed the task less than 20 min
	1b. Completeness	All task requirements were not completed	Parts of the task requirements were completed	All of the task requirements were completed
2. Coding skill (Wang et al., 2021; Xu et al., 2022)	2a. Coding structure	The students didn't not use the coding blocks correctly	The students used more than 30 coding blocks to finish the task correctly	The students used less than 30 coding blocks to finish the task correctly
	2b. Coding function	The students used functional blocks <i>if</i> and <i>loop</i> incorrectly	The students used functional blocks <i>if</i> and <i>loop</i> correctly but did not use them effectively	The students used functional blocks <i>if</i> and <i>loop</i> correctly and used them effectively

(see Table 2). Specifically, on the dimension of problem solving, two sub-dimensions (i.e., finish time, completeness) were used to assess whether the student pair completed the PP task correctly as required (Zheng et al., 2022). Two requirements of the final task were rated on the completeness dimension, respectively. On the dimension of coding skill, two sub-dimensions (i.e., coding structure, coding complexity) were used to assess whether the student pair could apply the coding skills that they have learned appropriately to solve the task (Wang et al., 2021; Xu et al., 2022). Summative assessment of final programming products was completed by two raters. Rater 1 first rated 25% of the dataset and rater 2 rated again to discuss with Rater 1 and reached an agreed assessment framework. Finally, two raters independently rated the other data and reached an inter-rater reliability with the Krippendorff's (2004) alpha reliability of 0.959.

Then, K-means clustering was used to extract the similar clusters of student groups' PP based on the process and summative assessment. K-means clustering, as an unsupervised algorithm, is designed to partition two-way, two-mode data (i.e., N objects with measurements on P variables) into K classes (MacQueen, 1967; Steinley, 2006). K-means clustering was run through R package *factoextra* (Kassambara & Mundt, 2017). To achieve an alignment, the process and summative assessment of student pairs were transferred into standard scores before K-means clustering. Elbow method was used to select and determine the optimal value of K clusters. This method gives total within sum of squares (TWSS) for each value of K through the iteration; the value of K is optimal when TWSS drops dramatically and reaches an inflection point (i.e., elbow) (Kodinariya & Makwana, 2013).

Collaborative pattern analysis

Quantitative content analysis (QCA), click stream analysis (CSA), and video analysis (VA) were used to analyze the process data of students' PP. The computer screen recording data and video recording data (with audio) were transcribed by two researchers to record students' verbal communications, operational behaviors, facial expressions in the same time scale. During the transcription, the unit of

analysis for audio recording data was the unit of a sentence spoken by a student; the unit of analysis for the operation was a clickstream behavior conducted by a student when a student moved or clicked the mouse on the platform; and the unit of analysis for facial expression was one time of facial expression when a student was speaking or operating the computer. After the transcription, 19 datasets included 10,874 units of data (Mean = 572.32; SD = 48.25). There were a total of 3604 units of verbal data (Mean = 189.68; SD = 40.08), 2,356 units of behavior data (Mean = 124.00; SD = 51.33), and 4914 units of facial data (Mean = 258.63; SD = 38.39).

Based on the previous relevant literature (Díez-Palomar et al., 2021; Pekrun et al., 2002; Rogat & Adams-Wiggins, 2015; Sun et al., 2020, 2021), a coding framework was proposed to analyze the process data of PP on the *verbal communication*, *operational behavior*, and *facial expression* dimensions (see Table 3). The coding procedure were completed by three raters. Rater 1 first coded 30% of the dataset according to the proposed coding scheme. Next, rater 2 coded the data again and discussed with rater 1 to solve discrepancies. At this phase, Krippendorff's (2004) alpha reliability was 0.853

Table 3 The coding framework

Dimension	Code	Description
Verbal communication (Díez-Palomar et al., 2021; Sun et al., 2020)	Self-talk (ST)	A student spoke to himself/herself
	Question proposal (QP)	A student asked some questions
	Simple response (SR)	A student simply replied to others (e.g., Yes, ahh...)
	Opinion expression (OE)	A student expressed new ideas, opinions, or solutions
	Argumentation (AG)	A student argued about other's communication or operation
	Knowledge construction (KC)	A student constructed knowledge or shared explicit opinions based on the previous information or perspectives
	Consensus reaching (CR)	A student reached a consensus to other's opinions (e.g., I agree that...)
	Function Maintenance (FM)	A student maintained effective team function through regulative discourse (e.g., Let's try...)
	Negative response (NR)	A student ignored, avoided, or responded negatively to what others say
Operational behavior (Sun et al., 2021)	Adjusting parameter (AP)	A student adjusted parameters (such as quantity, direction) in a coding block
	Adjusting code (AC)	A student selected, assembled, and adjusted coding block
	Running program (RP)	A student executed the "Run" command to run the coding blocks
	Debugging (DB)	A student debugged and modified the coding blocks based on the existing problems
Facial expression (Pekrun et al., 2002; Rogat & Adams-Wiggins, 2015)	Positive (PO)	A student expressed positive social emotions such as smiling, nodding
	Moderate (MO)	A student showed no explicit facial expression
	Negative (NE)	A student expressed negative social feelings by frowning, curling mouth, squinting eyes, etc

between two raters. Finally, rater 1 coded the rest of dataset, then rater 3 double-checked the coding results to decide if there were any problems.

Next, three analytics methods were used to reveal the *quantitative*, *structural*, and *transitional* characteristics of the collaborative patterns. From a quantitative perspective, statistical analysis (SA) was used to analyze the frequency of verbal communication, operational behavior, and facial expression and then a one-way analysis of variance (ANOVA) was conducted to test the significance of differences among clusters.

From a structural perspective, epistemic network analysis (ENA) was used to demonstrate the structure of connections among the verbal communication, operational behavior, and facial expression dimensions in different clusters. ENA can detect and represent the accumulative connections between elements in coded data in dynamic networks (Csanadi et al., 2018; Shaffer et al., 2016). In this research, ENA was conducted on all codes of three dimensions. ENA Webkit (epistemicnetwork.org) was utilized to conduct ENA analysis and its visualization (Marquart et al., 2018). Referring to threshold value used in previous research (Shaffer et al., 2016), we set the threshold of edge weight as 0.25 in ENA and showed the strong and representative connections rather than all connections, in order to clearly interpret the structural characteristics among different clusters.

From a transitional perspective, process mining (PM) was used to detect and visualize the transitional processes of the verbal communication, operational behavior, and facial expression dimensions among different collaborative clusters. PM is a temporal data mining and analysis method that focuses exclusively on transitions between events or activities (Reimann, 2009; Schoor & Bannert, 2012). The software *Disco* 3.1.4 was used to analyze PM models that examine and visualize the code transitions (Rozinat & Günther, 2012).

Results

After the process and summative assessment of student's PP, the clustering results of K-means generated based on the distribution of corresponding standard scores. With the value of K as suggested by the elbow method ($K=4$) (see Fig. 2), the optimal clustering results revealed four clusters of collaborative types, consisting of 5, 5, 6, and 3 student pairs for *Cluster 1* (i.e., the yellow section), *Cluster 2* (i.e., the green section), *Cluster 3* (i.e., the blue section), and *Cluster 4* (i.e., the orange section), respectively (see Fig. 3).

Among the four clusters, Cluster 1 had the highest score of collaborative processes (Mean = 38.60, SD = 2.33), followed by Cluster 2 (Mean = 31.80, SD = 3.71), Cluster 3 (Mean = 19.67, SD = 2.21), and Cluster 4 (Mean = 12.33, SD = 1.89). Cluster 1 also had the highest score of collaborative products (Mean = 21.80, SD = 2.04), followed by Cluster 4 (Mean = 13.67, SD = 1.89), Cluster 2 (Mean = 11.80, SD = 0.98), and Cluster 3 (mean = 11.67, SD = 1.49). In summary, Cluster 1 had the high performance in both process and summative assessment. Cluster 2, Cluster 3, and Cluster 4 had a low-level performance in summative assessment. Cluster 2 had the relatively high performance in process assessment, while Cluster 4 had a relatively low performance in process assessment.

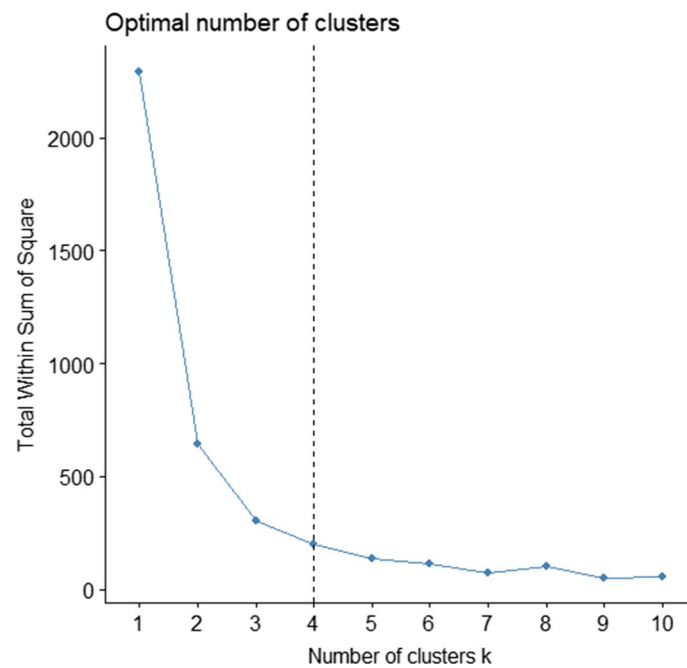


Fig. 2 The optimal clusters of “K” with the elbow method

From a quantitative perspective

From a quantitative perspective, ANOVA with the Bonferroni correction was conducted to test the significant differences between the four collaborative clusters on the three dimensions. Levene tests were conducted before ANOVAs and the results showed the homogeneity of variance. Moreover, post-hoc pairwise comparisons were conducted to further reveal significant differences between clusters (see Table 4). Considering that some codes (i.e., NR, RP, NE) were not normally distributed, a non-parametric test was conducted to cross-check the ANOVA results. The results showed that there were significant differences in the frequency of KC, CR, and PO ($p < 0.05$) with the Bonferroni correction under the Kruskal–Wallis test. Specifically, on the *verbal communication* dimension, there was statistically significant difference on both KC and CR, where Cluster 1 had the highest frequency, followed by Cluster 2, Cluster 3, and Cluster 4. However, there were no statistically significant differences on the other codes (i.e., ST, QP, SR, OE, CR, FM, NR) ($p > 0.05$). In addition, OE and ST appeared frequently while NR appeared infrequently in all the four clusters. On the *operational behavior* dimension, no statistically significant differences were found on the codes (i.e., AP, AC, RP, DB). Moreover, all four clusters had a low level of frequency on AP and a high level of frequency on AC. On the *facial expression* dimension, statistical significances were found on PO (Cluster 1 > Cluster 3 > Cluster 4; Cluster 2 > Cluster 4). Moreover, there were no statistically significant differences on MO (all four clusters had a high level of MO) and NE (all four clusters had a low level of NE).

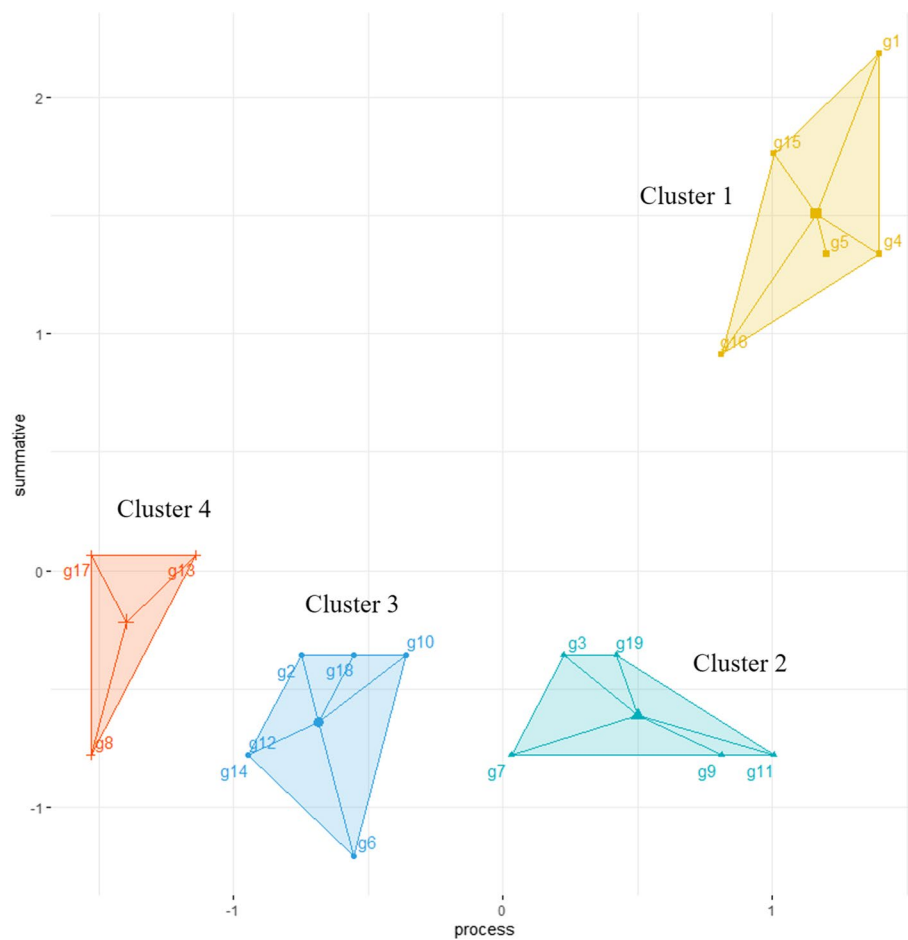


Fig. 3 The K-means clustering results ($K=4$)

From a structural perspective

From a structural perspective, the characteristics among the four clusters were reflected by the connection values and the centroid locations of the ENA plots (see Fig. 4). For all four clusters, most of the codes shared strong connections with MO in epistemic networks. Specifically, regular characteristics among the four clusters were reflected by four pairs of connected codes (connection values >0.40), including OE – MO, ST – MO, AC – MO, and SR – MO. Moreover, OE – MO had the strongest connections (connection values >0.85) among all the pairs in all four clusters. NR, NE, and AP were weakly associated with other codes (connection values <0.25) in the four clusters.

Different characteristics were identified among the four clusters of collaborative types, reflected by the locations of the centroid in epistemic networks (shown as red nodes in Fig. 4). In Cluster 1, the centroid of the epistemic network was located at the upper left corner, mainly focusing on PO, KC, and CR (i.e., connection value of MO-PO=0.55, connection value of MO-KC=0.45, connection value of MO-CR=0.31). In Cluster 2, the centroid of the epistemic network was located at the lower left corner, mainly focusing on AG, SR, FM, and QP (i.e.,

Table 4 Results of code frequencies and one-way ANOVAs of four collaborative cluster types

Code	Cluster 1 (N = 5) mean (SD)	Cluster 2 (N = 5) mean (SD)	Cluster 3 (N = 6) mean (SD)	Cluster 4 (N = 3) mean (SD)	ANOVA		Pairwise comparison
					F	P	
Verbal communication							
ST	35.40 (8.05)	33.60 (7.14)	30.33 (12.01)	31.33 (9.45)	0.28	>0.10	
QP	17.60 (7.70)	20.80 (2.59)	18.00 (10.77)	13.67 (2.08)	0.57	>0.10	
SR	30.00 (4.18)	34.00 (8.63)	24.67 (9.69)	23.67 (5.84)	1.54	>0.10	
OE	50.00 (9.70)	55.00 (9.33)	47.00 (21.62)	44.00 (2.00)	0.45	>0.10	
AG	26.60 (9.45)	38.20 (17.17)	21.67 (5.79)	17.33 (11.59)	2.73	>0.05	
KC	22.60 (9.99)	16.40 (9.63)	12.17 (3.60)	5.67 (3.51)	3.54	<0.05*	Cluster 1 > Cluster 3; Cluster 1 > Cluster 4
CR	14.80 (7.63)	13.40 (4.04)	9.17 (4.54)	4.00 (2.65)	3.23	<0.05*	Cluster 1 > Cluster 4; Cluster 2 > Cluster 4
FM	22.20 (6.54)	23.40 (5.94)	15.00 (6.54)	13.67 (9.07)	2.38	>0.10	
NR	2.60 (1.52)	3.00 (4.64)	4.17 (4.96)	6.33 (3.22)	0.64	>0.10	
Operational behavior							
AP	7.80 (3.11)	5.40 (2.07)	6.67 (4.41)	3.67 (2.08)	1.13	>0.10	
AC	39.80 (7.82)	33.00 (11.60)	30.83 (11.41)	38.67 (21.22)	0.60	>0.10	
RP	17.60 (5.98)	9.80 (8.08)	19.33 (7.09)	22.33 (12.70)	1.94	>0.10	
DB	20.00 (6.36)	23.40 (9.84)	30.00 (15.70)	21.33 (13.58)	0.74	>0.10	
Facial expression							
PO	30.80 (15.45)	21.80 (4.60)	18.17 (5.88)	7.00 (4.58)	4.46	<0.05*	Cluster 1 > Cluster 3 > Cluster 4; Cluster 2 > Cluster 4
MO	231.60 (50.62)	244.20 (28.77)	210.17 (58.84)	196.67 (23.86)	0.89	>0.10	
NE	0.20 (0.45)	0.80 (2.49)	1.33 (2.81)	1.33 (1.53)	0.50	>0.10	

* $p < 0.05$

connection value of MO-AG=0.68, connection value of MO-SR=0.64, connection value of MO-FM=0.42, connection value of MO-QP=0.41, connection value of AG-SR=0.32). In Cluster 3, the centroid of the epistemic network was located at the upper right corner, mainly focusing on RP, ST, and AC (i.e., connection value of MO-RP=0.76, connection value of MO-ST=0.64, connection value of MO-AC=0.50). In Cluster 4, the centroid was located at the lower right corner, mainly focusing on OE, DB, NE, and NR (i.e., connection value of MO-OE=0.99, connection value of MO-DB=0.47, connection value of MO-NR=0.17, connection value of MO-NE=0.04). In summary, Cluster 1 concentrated on positively constructing knowledge and reaching consensus; Cluster 2 concentrated on arguing, asking question, simple replying and maintaining function; Cluster 3 concentrated on self-talking, adjusting code and running program; Cluster 4 concentrated on negatively expressing opinion, responding and debugging.

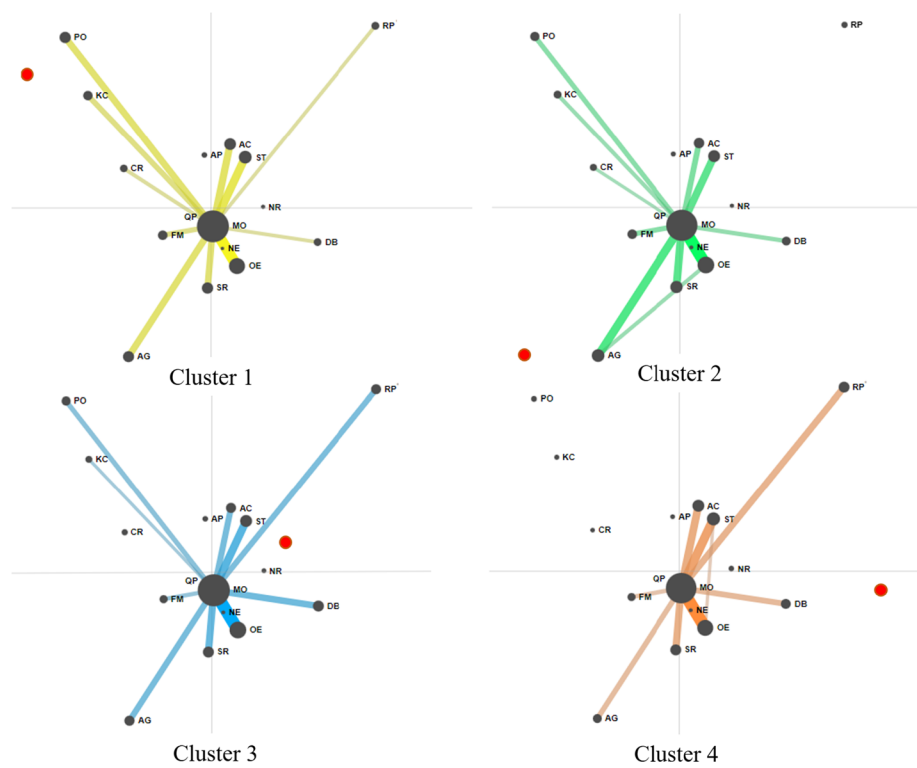


Fig. 4 The epistemic network analysis of four collaborative clusters. The threshold of edge weight in the epistemic networks was set as 0.25 to show the representative connections and structural characteristics (i.e., connection value ≥ 0.25) among the verbal communication, operational behavior, and facial expression dimensions

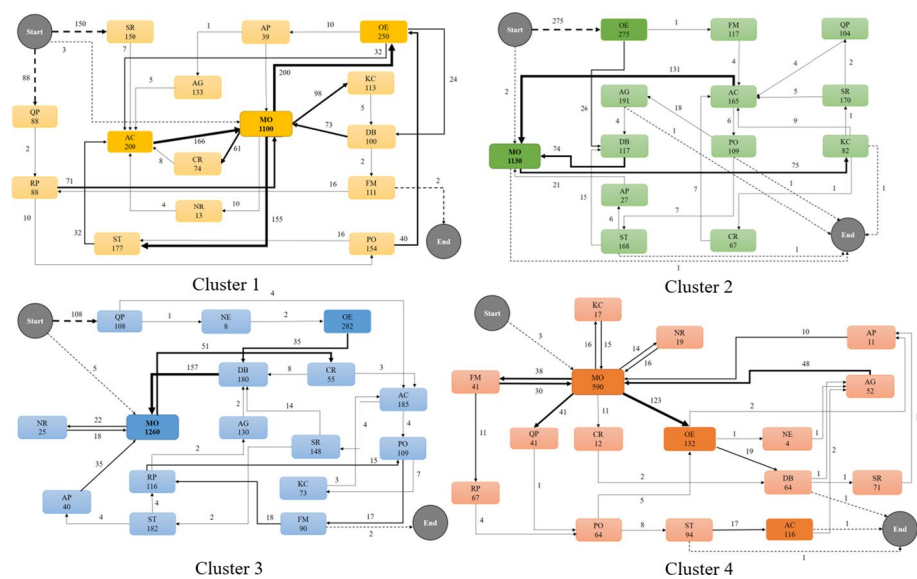


Fig. 5 The process mining results of four collaborative clusters. In the process models, the boxes refer to the absolute frequencies of codes and the arrows refer to the observed directional transitions from code A to code B

From a transitional perspective

From a transitional perspective, the characteristics among the four clusters were reflected by the code transitions in the process models (see Fig. 5). The regular characteristics of four clusters began by *verbal communication* (SR, QP in Cluster 1; OE, FM in Cluster 2; QP in Cluster 3; KC, NR, FM, QP in Cluster 4) and *moderate emotion* (MO in all four clusters), then moved to *operational behavior* (AC, RP in Cluster 1; DB, AC in both Cluster 2 and Cluster 3; RP, DB in Cluster 4), and finally ended with *verbal communication* (FM in Cluster 1 and Cluster 3; AG, KC, ST in Cluster 2; ST in Cluster 4).

Different transitional characteristics were found among four clusters (see Fig. 5). In Cluster 1, student pairs were more likely to start with two paths, including SR AC MO and QP RP MO/PO. Student pairs mainly ended with FM, which indicated that they regulated to maintain the function at the end of PP. Moreover, three loops appeared frequently in Cluster 1, including MO CR AC MO, MO ST AC MO, and MO OE DB MO. These results indicated that students tended to adjust coding blocks through self-talking and reaching consensus, and debug the programs through expressing new opinions. In Cluster 2, students usually started their collaboration with OE and then divided into two paths, namely OE FM AC MO/PO and OE DB MO. Compared to other three clusters, Cluster 2 ended with more codes (i.e., AG, ST, KC, MO, PO). Two loops often appeared during the PP processes, including MO KC (SR QP) AC MO and MO KC AC PO AG DB MO. These results indicated that students were more likely to constructed knowledge to drive the coding behaviors, but usually argued with each other when debugging programs. In Cluster 3, students had high probability to start their collaboration with QP, then moved to the path of NE OE DB or directly moved to AC. They mainly ended with FM, which also indicated that they regulated to maintain the function of pairs in the end. Two loops usually appeared in Cluster 3, including MO NR MO and MO CR DB MO. These results indicated that students sometimes replied to the peer negatively and sometimes reached a consensus to debug programs. In Cluster 4, the code transitions and loops started with MO and ended with ST, AC and DB. Specifically, a loop of MO OE DB AG MO appeared most frequently among all loops, which indicated that they expressed opinions to debug and solve problems but usually argued with each other. In addition, the loops of MO KC MO, MO NR MO and MO FM MO sometimes appeared, which also implied that pairs not only made regulations, but also had negative interactions when constructing knowledge.

Discussions and implications

This research applied MMLA to examine students' collaborative patterns in a face-to-face, computer-supported PP environment in higher education. Specifically, we collected students' multimodal process-oriented data and programming products data, and proposed an analytical framework integrating MMLA methods to detect and examine student pairs' collaborative patterns. Based on the process and summative assessment results, four clusters were detected from 19 pairs through K-means clustering, namely *Cluster 1* (5 pairs), *Cluster 2* (5 pairs), *Cluster 3* (6 pairs), and *Cluster 4* (3 pairs). Cluster 1, with the high performance in both process and summative assessment, was characterized as a *positively-engaged, knowledge-constructed, and consensus-achieved* pattern.

Cluster 2, with a relatively high performance in process assessment but a low performance in summative assessment, was characterized as a *moderately-engaged, argumentation-driven, and opinion-divergent* pattern. Cluster 3, with the low performance in both process assessment and summative assessment, was characterized as a *negatively-engaged, individual-oriented, and problems-unsolved* pattern. Cluster 4, with a low performance in process assessment but a relatively higher performance in summative assessment, was characterized as a *negatively-engaged, opinion-centered, and trial-and-error* pattern. Overall, this research revealed four clusters of student pairs with distinct collaborative patterns and performances, that initially verify the complexity, multimodality, and dynamics of CPS as well as their relations with collaborative quality.

From a theoretical perspective, this research contributed to the extant literature on CPS through revealing how complex connections among multimodality emerged into different collaborative patterns which in turn influenced the collaborative quality of final products. First, regarding the highly-performed collaborative pattern (i.e., Cluster 1), we found that opinion expression after a series of operations and trials could form a foundation for deep-level knowledge construction and group regulation to achieve high-quality collaboration (Ouyang & Chang, 2019; Park et al., 2015). Moreover, compared to negative emotions (i.e., Cluster 3, 4), students' positive emotions might contribute to the high quality of collaboration, like Cluster 1 did (Törmänen et al., 2021). Furthermore, consensus reaching in argumentation is also the key to achieve a high-quality of collaboration (Straus, 2002). Second, previous research verified that argumentation contributed to CPS through cognitive elaboration and knowledge construction (Stegmann et al., 2012), but constant argumentation without peers' consensus might result in divergence of opinions and inefficient collaboration (i.e., Cluster 2). Third, inconsistent with previous research that highlighted the role of self-talk in promoting self-regulation in CPS (DiDonato, 2013), the frequent use of self-talks (i.e., Cluster 3) might result in too much individual-oriented opinion expression and less group negotiation, which may in turn lead to the failure of collaboration. Students in Cluster 3 also spent most of the time on debugging, which somehow indicated that they encountered difficulties without successful programming in collaboration (Klahr & Carver, 1988). Fourth, compared to Cluster 3, students in Cluster 4 tended to express opinions together and appeared more programming running behaviors during debugging to achieve a relatively higher summative performance. Hence, running programming and debugging could together reflect students' persistence and productive struggle in PP that help them learn from failures (Kapur, 2008; Kim et al., 2022).

From a pedagogical perspective, instructors should concentrate on the collaborative process and provide appropriate scaffoldings and interventions to support a high quality of collaborative programming. First, instructors should provide scaffoldings to enhance student pairs' collaboration quality based on the characteristics of collaborative patterns. For example, students in Cluster 3 were more likely to be individual-orientated rather than group-orientated, which led to the low performance in both process and summative assessment; therefore, instructors can regulate their collaboration through some metacognitive scaffoldings (e.g., planning group's goal) and socio-emotional scaffoldings (e.g., encouraging students to collaborate) to achieve

group cohesion within student pairs (Molenaar et al., 2014; Ouyang et al., 2021). In addition, students in Cluster 3 and Cluster 4 had constant debugging and frequent errors, which might indicate that they were not familiar with the programming skills; therefore, cognitive scaffoldings (e.g., task-relevant information or hint) can be provided to help them solve the problems in programming (Ouyang & Xu, 2022; Zhong & Si, 2021). Second, most of the students mainly expressed moderate emotions rather than positive emotions during the PP processes. However, positive social emotion plays an important role to motivate learning interest, lessen tension, and improve social cohesion in collaboration (Rogat & Adams-Wiggins, 2015), such as how Cluster 1 performed in this research. Hence, the engagement of instructors as social supporters during students' collaborative programming, might mobilize the collaborative atmosphere to reach the goal of high-quality PP (Ouyang & Scharber, 2017; Ouyang & Xu, 2022). Third, since constant argumentation and opinion divergence are the critical factors that resulted in low-quality PP (e.g., Cluster 2), instructors are supposed to pay attention to the conflicting moments in argumentation and make appropriate interventions (e.g., easing the atmosphere, providing new ideas) to guide the co-construction of knowledge and problem-solving (Barron, 2000). Overall, instructors should be aware of student pair' collaborative patterns as well as the complex characteristics, and support their work appropriately with varied scaffoldings.

From an analytical perspective, since CPS is a complex and adaptive phenomenon (Stahl & Hakkarainen, 2021), multimodal data collection and learning analytics are suggested for future works to explore the complex problems and phenomena in CPS (Jacobson et al., 2016; Ouyang et al., 2022). Compared to traditional performance evaluation (e.g., test score, product data) and self-report data (e.g., questionnaire, interview), process-based multimodal data and learning analytics methods provides us a holistic, complementary, fine-grained perspective to understand the complex nature of CPS (Hilpert & Marchand, 2018; Kapur, 2011). Recently, many research has used multimodal data (e.g., speed rate, gesture, body movement, eye movement) as well as learning analytics methods to examine the complex, synergistic, and dynamic collaborative patterns and characteristics in CPS (e.g., Mu et al., 2020; Ouyang et al., 2022; Wiltshire et al., 2019). Echoing this trend, this research collected student pairs' multimodal data (i.e., verbal audios, computer screen recordings, facial expression recordings, final products data) and applied multiple learning analytics methods (e.g., content analysis, epistemic network analysis, process mining) to investigate the collaborative patterns in PP as well as their quantitative, structural, and transitional characteristics. Furthermore, advanced and automated artificial intelligence (AI) algorithms (e.g., hidden Markov model, natural language processing, recurrence quantification analysis) are advised to analyze the complexity and dynamics of collaboration in the future research (Gorman et al., 2020; Hoppe et al., 2021). Compared to traditional learning analytics methods, AI-driven methods have potential to analyze multimodal and nonlinear data and extract the complex and dynamic structure of CPS (de Carvalho & Zárate, 2020). Overall, due to the complexity of CPS, it is critical to capture the fine-grained process data and utilize multimodal learning analytics to reveal the collaborative patterns as well as their implicit characteristics (Kapur, 2011; Reimann, 2009).

Conclusions, limitations, and future directions

Since it is challenging for novice programmers to succeed in collaborative programming, it is necessary to investigate how their multimodality can form different collaborative patterns and how different patterns contribute to the quality of collaborative programming. Using MMLA, the current research collected and analyzed multimodal data to understand the collaborative patterns during student pairs' PP in higher education. The results detected four collaborative patterns associated with different levels of process and summative performances. Based on these findings, the current research proposed theoretical, pedagogical, and analytical implications to guide future practice and research. There are two limitations in the current research, which lead to future research directions. First, since the current study aimed to explore collaborative clusters and patterns, the research design may lead to a threat to validity (Drost, 2011; Humphry & Heldsinger, 2014), which should be addressed in future research. For example, regarding the internal validity, we did not control the gender distribution of student pairs, which might partially influence the collaborative processes. In addition, although participants did not have prior programming foundations or experiences, no pre-test was set to measure and control students' prior programming knowledge. Moreover, the difficulty of the programming tasks may also have impacts on student collaboration. Regarding the external validity, the sample size of student pairs had a limited range of demographic backgrounds. Therefore, future CPS research is supposed to strictly control internal validity (e.g., gender, prior knowledge, task) and expand the sample size and pair structure and arrangement to test, validate, or modify the implications. Second, this MMLA research merely collected students' discourse, online behaviors, and facial expression from video data to analyze the CPS processes, and there is a lack of other multimodal data, such as physiological and psychological data. In addition, the facial expressions were coded manually rather than automated identification based on software, which might reduce the data analysis efficiency and accuracy. Therefore, AI-driven data collection and analysis methods as well as more modalities of data (e.g., physiological, eye tracking data) can provide further insights into CPS research. Overall, it is valuable to examine different collaborative patterns of novice programmers through MMLA, in order to tease out fine-grained and complex features, which serves as a data-driven evidence for promoting the quality of computer programming in higher education.

Acknowledgements

The authors would like to thank students who participated in this research.

Author contributions

WX designed and conducted data analysis, and wrote the manuscript draft; YW facilitated research design, collected and coded the data; and FO designed and supervised the research and revised the manuscript. All authors read and approved the final manuscript.

Funding

Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. This work was supported by National Natural Science Foundation of China (62177041); Zhejiang Province educational science and planning research Project (2022SCG256); Zhejiang University graduate education research Project (20220310).

Availability of data and materials

The data was available upon request from the first author.

Declarations

Competing interests

The authors declare that they have no competing interests.

Received: 28 September 2022 Accepted: 13 December 2022

Published online: 08 February 2023

References

- Barron, B. (2000). Achieving coordination in collaborative problem-solving groups. *Journal of the Learning Sciences*, 9(4), 403–436. https://doi.org/10.1207/S15327809JLS0904_2
- Beck, L., & Chizhik, A. (2013). Cooperative learning instructional methods for CS1: Design, implementation, and evaluation. *ACM Transactions on Computing Education*, 13(3), 10–32. <https://doi.org/10.1145/2492686>
- Borge, M., & Mercier, E. (2019). Towards a micro-ecological approach to CSCL. *International Journal of Computer-Supported Collaborative Learning*, 14(2), 219–235. <https://doi.org/10.1007/s11412-019-09301-6>
- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18(1), 32–42. <https://doi.org/10.3102/0013189X018001032>
- Bryant, S., Romero, P., & Du Boulay, B. (2006). The collaborative nature of pair programming. In P. Abrahamsson, M. Marchesi, & G. Succi (Eds.), *Extreme programming and agile processes in software engineering* (pp. 53–64). Springer. https://doi.org/10.1007/11774129_6
- Byrne, D., & Callaghan, G. (2014). *Complexity theory and the social sciences*. Routledge.
- Chittum, J. R., Jones, B. D., Akalin, S., & Schram, Á. B. (2017). The effects of an afterschool STEM program on students' motivation and engagement. *International Journal of STEM Education*, 4(1), 11–26. <https://doi.org/10.1186/s40594-017-0065-4>
- Csanadi, A., Egan, B., Kollar, I., Shaffer, D. W., & Fischer, F. (2018). When coding-and-counting is not enough: Using episodic network analysis (ENA) to analyze verbal data in CSCL research. *International Journal of Computer-Supported Collaborative Learning*, 13(4), 419–438. <https://doi.org/10.1007/s11412-018-9292-z>
- Damon, W., & Phelps, E. (1989). Critical distinctions among three approaches to peer education. *International Journal of Educational Research*, 13, 9–19. [https://doi.org/10.1016/0883-0355\(89\)90013-X](https://doi.org/10.1016/0883-0355(89)90013-X)
- de Carvalho, W. F., & Zárte, L. E. (2020). A new local causal learning algorithm applied in learning analytics. *The International Journal of Information and Learning Technology*, 38(1), 103–115. <https://doi.org/10.1108/IJILT-04-2020-0046>
- Denner, J., Green, E., & Campe, S. (2021). Learning to program in middle school: How pair programming helps and hinders intrepid exploration. *Journal of the Learning Sciences*, 30(4–5), 611–645. <https://doi.org/10.1080/10508406.2021.1939028>
- DiDonato, N. C. (2013). Effective self-and co-regulation in collaborative learning groups: An analysis of how students regulate problem solving of authentic interdisciplinary tasks. *Instructional Science*, 41(1), 25–47. <https://doi.org/10.1007/s11251-012-9206-9>
- Diez-Palomar, J., Chan, M. C. E., Clarke, D., & Padrós, M. (2021). How does dialogical talk promote student learning during small group work? An exploratory study. *Learning, Culture and Social Interaction*, 30, 100540. <https://doi.org/10.1016/J.LCSI.2021.100540>
- Dillenbourg, P. (1999). What do you mean by collaborative learning? In P. Dillenbourg (Ed.), *Collaborative-learning: Cognitive and computational approaches* (pp. 1–19). Elsevier.
- Drost, E. A. (2011). Validity and reliability in social science research. *Education Research and Perspectives*, 38(1), 105–123. <https://doi.org/10.3316/informit.491551710186460>
- Gorman, J. C., Grimm, D. A., Stevens, R. H., Galloway, T., Willemsen-Dunlap, A. M., & Halpin, D. J. (2020). Measuring real-time team cognition during team training. *Human Factors*, 62(5), 825–860. <https://doi.org/10.1177/0018720819852791>
- Han, F., & Ellis, R. A. (2021). Patterns of student collaborative learning in blended course designs based on their learning orientations: A student approaches to learning perspective. *International Journal of Educational Technology in Higher Education*, 18(1), 1–16. <https://doi.org/10.1186/s41239-021-00303-9>
- Hawiltschek, A., Berndt, S., & Schulz, S. (2022). Empirical research on pair programming in higher education: A literature review. *Computer Science Education*. <https://doi.org/10.1080/08993408.2022.2039504>
- Hilpert, J. C., & Marchand, G. C. (2018). Complex systems research in educational psychology: Aligning theory and method. *Educational Psychologist*, 53(3), 185–202. <https://doi.org/10.1080/00461520.2018.1469411>
- Hmelo-Silver, C. E., & DeSimone, C. (2013). Problem-based learning: An instructional model of collaborative learning. In C. Hmelo-Silver, C. A. Chinn, C. Chan, & A. O'Donnell (Eds.), *The international handbook of collaborative learning*. Routledge.
- Hoppe, H. U., Doberstein, D., & Hecking, T. (2021). Using sequence analysis to determine the well-functioning of small groups in large online courses. *International Journal of Artificial Intelligence in Education*, 31, 680–699. <https://doi.org/10.1007/s40593-020-00229-9>
- Humphry, S. M., & Heldsinger, S. A. (2014). Common structural design features of rubrics may represent a threat to validity. *Educational Researcher*, 43(5), 253–263. <https://doi.org/10.3102/0013189X1454215>
- Jacobson, M. J., Kapur, M., & Reimann, P. (2016). Conceptualizing debates in learning and educational research: Toward a complex systems conceptual framework of learning. *Educational Psychologist*, 51(2), 210–218. <https://doi.org/10.1080/00461520.2016.1166963>
- Kapur, M. (2008). Productive failure. *Cognition and Instruction*, 26(3), 379–424. <https://doi.org/10.1080/07370000802212669>
- Kapur, M. (2011). Temporality matters: Advancing a method for analyzing problem-solving processes in a computer-supported collaborative environment. *International Journal of Computer-Supported Collaborative Learning*, 6(1), 39–56. <https://doi.org/10.1007/s11412-011-9109-9>
- Kassambara, A., & Mundt, F. (2017). *Package 'factoextra'. Extract and visualize the results of multivariate data analyses*. [Software]. R Package Version 1.0.7.

- Kawamura, R., Shirai, S., Takemura, N., Alizadeh, M., Cukurova, M., Takemura, H., & Nagahara, H. (2021). Detecting drowsy learners at the wheel of e-learning platforms with multimodal learning analytics. *IEEE Access*, 9, 115165–115174. <https://doi.org/10.1109/ACCESS.2021.3104805>
- Kim, C., Vasconcelos, L., Belland, B. R., Umutlu, D., & Gleasman, C. (2022). Debugging behaviors of early childhood teacher candidates with or without scaffolding. *International Journal of Educational Technology in Higher Education*, 19(1), 1–26. <https://doi.org/10.1186/s41239-022-00319-9>
- Klahr, D., & Carver, S. M. (1988). Cognitive objectives in a LOGO debugging curriculum: Instruction, learning, and transfer. *Cognitive Psychology*, 20(3), 362–404. [https://doi.org/10.1016/0010-0285\(88\)90004-7](https://doi.org/10.1016/0010-0285(88)90004-7)
- Kodinariya, T. M., & Makwana, P. R. (2013). Review on determining number of cluster in K-Means clustering. *International Journal*, 1(6), 90–95.
- Krippendorff, K. (2004). Reliability in content analysis: Some common misconceptions and recommendations. *Human Communication Research*, 30(3), 411–433. <https://doi.org/10.1093/hcr/30.3.411>
- Kwon, K., Liu, Y. H., & Johnson, L. P. (2014). Group regulation and social-emotional interactions observed in computer supported collaborative learning: Comparison between good vs. poor collaborators. *Computers & Education*, 78, 185–200. <https://doi.org/10.1016/j.compedu.2014.06.004>
- Lin, P. C., Hou, H. T., Wu, S. Y., & Chang, K. E. (2014). Exploring college students' cognitive processing patterns during a collaborative problem-solving teaching activity integrating Facebook discussion and simulation tools. *The Internet and Higher Education*, 22, 51–56. <https://doi.org/10.1016/j.iheduc.2014.05.001>
- Lu, J., & Churchill, D. (2014). Using social networking environments to support collaborative learning in a Chinese university class: Interaction pattern and influencing factors. *Australasian Journal of Educational Technology*, 30(4), 1–15. <https://doi.org/10.14742/ajet.655>
- MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In L. M. Le Cam & J. Neyman (Eds.), *Proceedings of the Fifth Berkeley symposium on mathematical statistics and probability* (pp. 281–297). University of California Press.
- Malmberg, J., Järvelä, S., & Järvenoja, H. (2017). Capturing temporal and sequential patterns of self-, co-, and socially shared regulation in the context of collaborative learning. *Contemporary Educational Psychology*, 49, 160–174. <https://doi.org/10.1016/j.cedpsych.2017.01.009>
- Marquart, C. L., Hinojosa, C., Swiecki, Z., Eagan, B., & Shaffer, D. W. (2018). *Epistemic network analysis* [Software]. Version 1.6.0. <https://epistemicnetwork.org>
- Meier, A., Spada, H., & Rummel, N. (2007). A rating scheme for assessing the quality of computer-supported collaboration processes. *International Journal of Computer-Supported Learning*, 2, 63–86. <https://doi.org/10.1007/s11412-006-9005-x>
- Molenaar, I., Sleegers, P., & van Boxtel, C. (2014). Metacognitive scaffolding during collaborative learning: A promising combination. *Metacognition and Learning*, 9(3), 309–332. <https://doi.org/10.1007/s11409-014-9118-y>
- Mu, S., Cui, M., & Huang, X. (2020). Multimodal data fusion in learning analytics: A systematic review. *Sensors*, 20(23), 6856. <https://doi.org/10.3390/s20236856>
- O'Donnell, A. M., & Hmelo-Silver, C. E. (2013). Introduction: What is collaborative learning? An overview. In C. E. Hmelo-Silver, C. A. Chinn, C. K. K. Chan, & A. M. O'Donnell (Eds.), *The international handbook of collaborative learning* (pp. 93–111). Routledge.
- Ochoa, X., & Worsley, M. (2016). Augmenting learning analytics with multimodal sensory data. *Journal of Learning Analytics*, 3(2), 213–219. <https://doi.org/10.18608/jla.2016.32.10>
- Ouyang, F., & Chang, Y. H. (2019). The relationships between social participatory roles and cognitive engagement levels in online discussions. *British Journal of Educational Technology*, 50(3), 1396–1414. <https://doi.org/10.1111/bjet.12647>
- Ouyang, F., Chen, Z., Cheng, M., Tang, Z., & Su, C.-Y. (2021). Exploring the effect of three scaffoldings on the collaborative problem-solving processes in China's higher education. *International Journal of Educational Technology in Higher Education*, 18(35), 1–22. <https://doi.org/10.1186/s41239-021-00273-y>
- 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., & Scharber, C. (2017). The influences of an experienced instructor's discussion design and facilitation on an online learning community development: A social network analysis study. *The Internet and Higher Education*, 35, 34–47. <https://doi.org/10.1016/j.iheduc.2017.07.002>
- Ouyang, F., & Xu, W. (2022). The effects of three instructor participatory roles on a small group's collaborative concept mapping. *Journal of Educational Computing Research*, 60(4), 930–959. <https://doi.org/10.1177/07356331211057283>
- Park, J. B. H., Schallert, D. L., Sanders, A. J. Z., Williams, K. M., Seo, E., Yu, L. T., Vogler, J. S., Song, K., Williamson, Z. H., & Knox, M. C. (2015). Does it matter if the teacher is there? A teacher's contribution to emerging patterns of interactions in online classroom discussions. *Computers and Education*, 82, 315–328. <https://doi.org/10.1016/j.compedu.2014.11.019>
- Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist*, 37(2), 91–106. <https://doi.org/10.1207/S15326985EP3702>
- Perera, D., Kay, J., Koprinska, I., Yacef, K., Zaïane, & O. R. (2009). Clustering and sequential pattern mining of online collaborative learning data. *IEEE Transactions on Knowledge and Data Engineering*, 21(6), 759–772. <https://doi.org/10.1109/TKDE.2008.138>
- Reimann, P. (2009). Time is precious: Variable- and event-centred approaches to process analysis in CSCL research. *International Journal of Computer-Supported Collaborative Learning*, 4(3), 239–257. <https://doi.org/10.1007/S11412-009-9070-Z/FIGURES/4>
- Rogat, T. K., & Adams-Wiggins, K. R. (2015). Interrelation between regulatory and socioemotional processes within collaborative groups characterized by facilitative and directive other-regulation. *Computers in Human Behavior*, 52, 589–600. <https://doi.org/10.1016/j.chb.2015.01.026>

- Romero, M., Lepage, A., & Lille, B. (2017). Computational thinking development through creative programming in higher education. *International Journal of Educational Technology in Higher Education*, 14(42), 1–15. <https://doi.org/10.1186/s41239-017-0080-z>
- Rozinat, A., & Günther, C. W. (2012). Disco [Software]. Version 3.1.4. <https://fluxicon.com/disco/>
- Schoor, C., & Bannert, M. (2012). Exploring regulatory processes during a computer-supported collaborative learning task using process mining. *Computers in Human Behavior*, 28(4), 1321–1331. <https://doi.org/10.1016/J.CHB.2012.02.016>
- Shaffer, D. W., Collier, W., & Ruis, A. R. (2016). A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data. *Journal of Learning Analytics*, 3(3), 9–45. <https://doi.org/10.18608/jla.2016.33.3>
- Stahl, G. (2009). *Studying virtual math teams*. Springer.
- Stahl, G. (2017). Group practices: A new way of viewing CSCL. *International Journal of Computer-Supported Collaborative Learning*, 12(1), 113–126. <https://doi.org/10.1007/s11412-017-9251-0>
- Stahl, G., & Hakkarainen, K. (2021). Theories of CSCL. In U. Cress, C. Rosé, A. F. Wise, & J. Oshima (Eds.), *International handbook of computer-supported collaborative learning* (pp. 23–44). Springer. https://doi.org/10.1007/978-3-030-65291-3_2
- Stegmann, K., Wecker, C., Weinberger, A., & Fischer, F. (2012). Collaborative argumentation and cognitive elaboration in a computer-supported collaborative learning environment. *Instructional Science*, 40(2), 297–323. <https://doi.org/10.1007/s11251-011-9174-5>
- Steinley, D. (2006). K-means clustering: A half-century synthesis. *British Journal of Mathematical and Statistical Psychology*, 59(1), 1–34. <https://doi.org/10.1348/000711005X48266>
- Straus, D. (2002). *How to make collaboration work: Powerful ways to build consensus, solve problems, and make decisions*. Berrett-Koehler Publishers.
- Sun, D., Ouyang, F., Li, Y., & Chen, H. (2020). Three contrasting pairs' collaborative programming processes in China's secondary education. *Journal of Educational Computing Research*, 59(4), 740–762. <https://doi.org/10.1177/0735633120973430>
- Sun, D., Ouyang, F., Li, Y., & Zhu, C. (2021). Comparing learners' knowledge, behaviors, and attitudes between two instructional modes of computer programming in secondary education. *International Journal of STEM Education*, 8, 54. <https://doi.org/10.1186/s40594-021-00311-1>
- Törmänen, T., Järvenoja, H., & Mänty, K. (2021). All for one and one for all—How are students' affective states and group-level emotion regulation interconnected in collaborative learning? *International Journal of Educational Research*, 109, 101861. <https://doi.org/10.1016/j.ijer.2021.101861>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Springer.
- Wang, L., Geng, F., Hao, X., Shi, D., Wang, T., & Li, Y. (2021). Measuring coding ability in young children: relations to computational thinking, creative thinking, and working memory. *Current Psychology*. <https://doi.org/10.1007/s12144-021-02085-9>
- Webb, N. M., Ing, M., Burnheimer, E., Johnson, N. C., Franke, M. L., & Zimmerman, J. (2021). Is there a right way? Productive patterns of interaction during collaborative problem solving. *Education Sciences*, 11(5), 214. <https://doi.org/10.3390/educsci11050214>
- Wiltshire, T. J., Steffensen, S. V., & Fiore, S. M. (2019). Multiscale movement coordination dynamics in collaborative team problem solving. *Applied Ergonomics*, 79, 143–151. <https://doi.org/10.1016/j.apergo.2018.07.007>
- Xu, W., Geng, F., & Wang, L. (2022). Relations of computational thinking to reasoning ability and creative thinking in young children: Mediating role of arithmetic fluency. *Thinking Skills and Creativity*, 44, 101041. <https://doi.org/10.1016/j.tsc.2022.101041>
- Zemel, A., & Koschmann, T. (2013). Recalibrating reference within a dual-space interaction environment. *International Journal of Computer-Supported Collaborative Learning*, 8(1), 65–87. <https://doi.org/10.1007/s11412-013-9164-5>
- Zheng, L., Zhen, Y., Niu, J., & Zhong, L. (2022). An exploratory study on fade-in versus fade-out scaffolding for novice programmers in online collaborative programming settings. *Journal of Computing in Higher Education*, 34, 489–516. <https://doi.org/10.1007/s12528-021-09307-w>
- Zhong, B., & Si, Q. (2021). Troubleshooting to learn via scaffolds: Effect on students' ability and cognitive load in a robotics course. *Journal of Educational Computing Research*, 59(1), 95–118. <https://doi.org/10.1177/0735633120951871>

Publisher's Note

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