Leveraging Speech Data and Text Classification for Identifying Socially Shared Regulation in Collaborative Problem

Wenting Sun[1], Jiangyue Liu[2]

Humboldt-Universität zu Berlin, Germany[1]
Suzhou University, China[2]

Abstract

To investigate how individuals interdependently regulate activities in collaborative learning to achieve shared learning goals, a concept named Socially Shared Regulation (SSR) has emerged [1]. The frequency of SSR appears to be significantly positively related to students' immediate knowledge gains when SSR functions to activate collaborative learning through new activities that further the learning process and by challenging ongoing interactions to find alternative directions [2]. With the rise of ChatGPT, researchers have shown increased interest in exploring text classification techniques in online discussions. However, research on using text classification techniques to classify SSR phases, especially in collaborative problem-solving (CPS) learning, remains limited. To deeply explore essential activities, accumulating research in engineering education has emphasized the value of learners' conversations during CPS [3,4]. Nevertheless, the application of text classification techniques to classify SSR phases in real-world classrooms, particularly in authentic practice courses, is still understudied.
This study collected valid speech recordings from 28 undergraduates in an engineering practice course. Eighteen hours of group dialogues were manually transcribed into text and tagged with five SSR phases: orientation, planning, support strategies, monitoring, and evaluation and reflection.
As a result, 4,258 SSR phases were identified. Seven text classifiers were built, including Logistic Regression (LR), Naïve Bayes (NB), Random Forest (RF), K-Nearest Neighbors (KNN), Support Vector Machine (SVM), XGBoost, and BERT. The SVM classifier outperformed the others with an accuracy of 0.67. This study provides an example of using text classification to identify SSR phases from speech transcriptions in an authentic face-to-face engineering practice course. These insights offer educators and designers a comprehensive guide to promoting effective CPS and SSR dynamics in authentic CPS settings, thereby enhancing the overall success of CPS.

Keywords: *Classroom Discourse, Learning Analytics, Machine Learning, Social Shared Regulation*

1. Introduction

With the shift in Computer-Supported Collaborative Learning (CSCL) from the perspective of individual learners to groups of learners [5], the transition of metacognition from individualistic models to socially situated models has become increasingly prevalent in research [6]. To investigate how individuals interdependently regulate activities in collaborative learning to achieve shared learning goals, a concept named Socially Shared Regulation (SSR) has emerged [1]. The frequency of SSR appears to be significantly positively related to students' immediate knowledge gains when SSR functions to activate collaborative learning through new activities that further the learning process and by challenging ongoing interactions to find alternative directions [2].
The popularity of ChatGPT has significantly increased researchers' interest in exploring text classification techniques in online discussions. For example, [7] combined n-grams and various machine learning (ML) algorithms to automatically classify Collaborative Problem Solving (CPS) events based on text-chat messages. To the best of our knowledge, research on using text classification techniques to classify SSR phases remains scarce, particularly using speech data collected in the context of authentic CPS practice courses.

2. Related Works

To deeply explore essential activities, accumulating research in engineering education has stressed the value of learners' conversation during CPS. For example, through epistemic network analysis, [8] found that different engineering design behavioural patterns under two instructional approaches.

These studies demonstrated the great value of a fine-grained analysis of learners' collaborative learning process when using CPS in engineering courses. But most of these studies focused on conceptual-design course in engineering, not practice courses. On the other hand, compared to the online peer discussion, the synchronous nature and proximity of fellow students in face-to-face settings stimulate learners to operate their thinking deeply and produce longer and reciprocal conversations [9]. This means that simply extrapolating the findings of SSR from one collaboration mode to another is not wise [10], the online setting and face-to-face setting might unravel different SSR profiles [2].
As a classical problem in natural language processing (NLP), text classification targets at assigning labels or tags to textual units [11]. Most researchers using text classification in CSCL involved large datasets from online discussion. For example, in English language, using 19,105 sentences from online inquiry-based discussion, [12] employed different models to classify cognitive presence, social presence, and teaching presence, like Logistic Regression (LR), Naïve Bayes (NB), Random Forest (RF), and Bidirectional Encoder Representations from Transformers (BERT). [7] used NB and k-nearest-neighbours (KNN) classifiers in two online chat datasets (one included 5,045 events and one included 15,950) to classify social dimension and cognitive dimension in virtual simulation CPS environment. In Chinese language, using 17,118 online discussion transcripts, [13] compared the performance of different models (BERT, SVM, NB, LR) to classify cognitive, metacognitive, behavioural, emotional engagement. For speech conversation data, [14] used BERT to classify cognitive and social CPS skills from 8,860 utterances (English as language) in the context of videoconferencing to collaboratively solve physics and math problems. It is understudied that how these models perform in a small speech transcription dataset from authentic face-to-face engineer practice course. Neither about in the SSR phase identification.
Therefore, the contribution of the present study is twofold. First, based on the prior code scheme about SSR phase, a new one adjusted based on authentic practice course was developed. Second, text classification techniques have been used to automatically identify SSR phases using speech data from authentic classroom. Although group awareness was not included in this study, the current findings provide valuable guidelines on how to identify utterances of group metacognition during collaborative learning, which might help to conceptually refine collaborators shared focus and assigned role adopting intragroup regulation strategies in virtual laboratory environments.
In summary, several research gaps still exist in the research about sequential SSR behaviours in authentic face-to-face engineering practice settings. Based on these, some research questions were proposed:
RQ1: Compared to other CPS environments, what other activities can be observed in authentic face-to-face practice courses?
RQ2: To what extent can text classification techniques identify different SSR phases?

3. Methods

3.1 Participants and Learning Context

36 undergraduates participated from a Chinese public university formed 18 dyads groups. This lesson was the practical course (one week) after their theoretical course (four weeks). In the CPS environment, the learning activities were IP sending and receiving using the Internet Control Message Protocol (ICMP) package. All groups had simple CSCL scripts to support them monitor their progress. Group members were required to use their own computers to take turns being the IP sender and receiver.

3.2 Data Collection and Analysis

The session lasted nearly 3 h 50 minutes and the operation session lasted approximately two hours with an instructor present throughout. Speech dialogues were recorded during the operation session.
The speech recordings data were transcribed manually, and then content analysis and qualitatively coded transcription were conducted. Two researchers independently coded the data, with one coding 20% of the data and the other coding all of them. The agreement of the coding results achieved 80%.

SSR code scheme
Considering the research questions, the features of learning materials and peer interaction, the coding instrument from [15] was employed as the initial version of code schemes. To adjust the existing code

scheme to match better the discussion content of the collaborative operating tasks in formal classes, the study followed the first four steps of the process for the thematic analysis [16]. After multiple rounds of listening to the speech recordings and discussion with the course instructor and coauthors, based on [15] framework, we developed the main categories of the coding scheme of the SSR activities in CPS engineer practice course with several subcodes respectively. Table 1 demonstrate the coding schemes that include the Orientation, Planning, Support strategies, Monitoring, as well as Evaluation and reflection phases and further activities of SSR (the content in italics is new content added in this study).

Table 1. The SSR coding schemes in face-to-face practical course.

SSR phases	Event	Activities
Orientation	Task Analysis	Exploring task demands
		Processing task demands/learning objectives
	Content Orientation	Generating hypotheses
		Activating prior knowledge
		Becoming aware of task perceptions
Planning	Planning in advance	Formulating problem solving plan (planning in advance)
		Selecting problem solving plan (planning in advance)
	Interim Planning	*Formulating problem solving plan repeated (interim planning)*
		Formulating problem solving plan new (interim planning)
		Peers' formulating problem solving plan new (interim planning)
		Teacher's formulating problem solving plan new
		Selecting problem solving plan
		Questioning the problem solving plan
Support strategies	*Peer Interaction*	*Asking for peers' support*
		Replying to peers' help asking
		Asking for teacher's support
	Online Searching	*Searching Online solutions*
Monitoring	Comprehension Monitoring	Noting lack of comprehension
		Checking comprehension by repeating
		Checking comprehension by elaborating
	Monitoring of Progress	*Checking of task execution process*
		Checking of task execution results
		Checking of progress (initiated by instructor)
		Checking of peers' progress
		Spontaneous checking of progress
		Reflecting on progress (initiated by script)
		Spontaneous reflecting on progress
		Writing progress
	Monitoring of Collaboration	Commenting on collaboration (Monitoring of collaboration)
		Reflecting on collaboration (Monitoring of collaboration)
Evaluation and reflection	Evaluation Learning Outcomes	Checking learning outcomes
		Elaborating on learning outcomes
	Evaluating Learning Process	Commenting on learning process
		Reflecting on learning process (initiated by script)
		Spontaneous reflecting on learning process
	Evaluating Collaboration	Commenting on collaboration (Evaluating collaboration)
		Reflecting on collaboration (Evaluating collaboration)

Text classification

5 SSR phases were used as labelling in our dataset, namely Orientation, Planning, Support strategies, Monitoring, Evaluation and reflection. The dataset was split into training and test datasets at 8:2 ratio.

Several common text classifiers were built, including Logistic Regression (LR), Naïve Bayes (NB), Random Forest (RF), k-nearest-neighbours (KNN), Support Vector Machine (SVM), XGBoost, BERT. Feature extraction used two steps, including "Jieba" Chinese text segmentation and Term Frequency-Inverse Document Frequency (TF-IDF). Then the outputs were put into the text classifier. This study performed all text classification processing with Python and its available package except BERT model used TensorFlow. For evaluation metrics, accuracy, precision, recall and F1 score were included. A simplified workflow can be seen on Figure1. Detailed code solutions can be found at the OSF link: https://osf.io/87jzy/?view_only=4744cd2a9d2948c29ead081e7034abeb

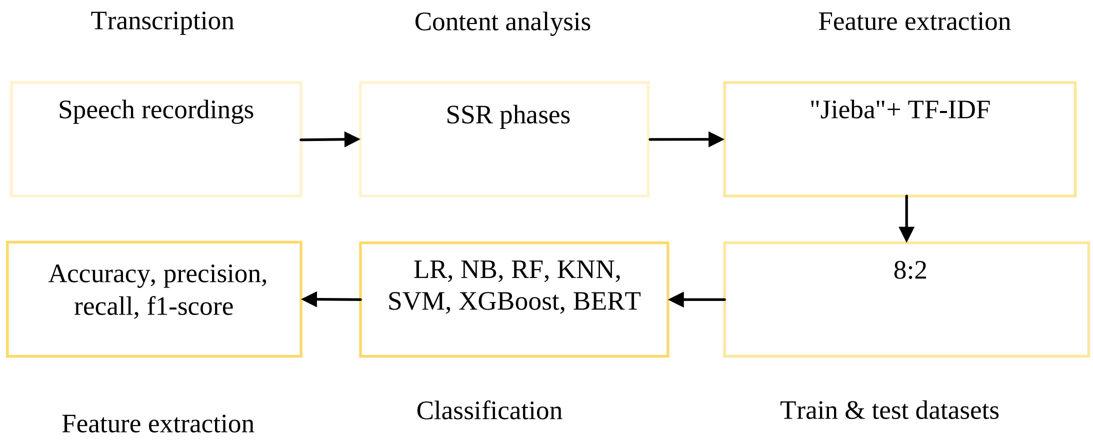

Figure 1. Speech recordings analysis procedure.

4. Results and Discussion

4.1 RQ1: Compared to Other CPS Environments, What Other Activities Can Be Observed in Authentic Face-To-Face Practice Courses?

Compared to the SSR code scheme from [15], we added new items under each dimensions according to the detailed activities extracted from our speech recordings in the entire authentic engineering CPS course (see Table 1, the content in italics is new content added in this study). The Planning dimension encompassed problem solving solutions either at the commencing phase or fine-tuning solutions based on operation results. In this dimension, the original term “Formulating problem solving plan” under “Interim Planning” were divided into another five new planning activities. To record the solution seeking activities when groups confront challenges beyond their abilities at that moment, we added a totally new dimension, “Support strategies”. The Monitoring dimension aimed at inconsistencies identifying and solutions modifying, including Comprehension Monitoring, Monitoring of Progress, and Monitoring of Collaboration. The original term “Checking of Progress” were divided into another five new monitoring activities and “Reflecting on Progress” were divided into another three new. The Evaluation and reflection dimension referred to students’ assessment of the completion of task. The original term “Reflecting on Learning Process” in this dimension were divided into another two new reflection activities. The coding schemes adjusted for this study can be furthermore used to assess or compare SSR behaviours in other domains or collaboration settings as well as provide educators cues when evaluating interventions in CPS.
Out of 36 students, the valid speech recordings were selected from 14 groups with 18 hours. All SSR phases were 4258, including 280 (6.6%) for Orientation, 686 (16.1%) for Planning, 200 (4.7%) for Support strategies, 2687 (63.1%) for Monitoring, and 405 (9.5%) for Evaluation and reflection. The high ratio of monitoring is different from prior studies [17]. The reason might lie in the task difference. The task we used in this study is a highly collaborative operation task, not conceptual-design task (e.g., [17]). The operation results of one student directly impact the next operation action of the other group member while the group performance is evaluated based on the group produced knowledge artifacts, which intensively raise the responsibility of all group members to monitoring the task completion.

4.2 RQ2: To What Extent Can Text Classification Techniques Identify Different SSR Phases?

Among the 7 classifiers (Table 2), SVM model overall produced better performance (Confusion matrix heatmap can be seen in Figure 2). The model trained with TF-IDF features outputs the accuracy and recall of 0.67, precision of 0.66. Our results are different from some previous findings. For example, using Chinese as language, [13] achieved accuracy of 0.85 at label of metacognitive engagement and accuracy of 0.76 at label of cognitive engagement using BERT. One important point is that their dataset is large, including 17,118 online discussion transcripts while our dataset only contained 4,258 labelled textual units. This might be explained through the mechanism of SVM. SVM do better at higher-order data using kernel functions [18].

Table 2. Summaries of model performance.

Classifier	Accuracy	Precision	Recall	F1-score
LR	0.66	0.65	0.66	0.59
NB	0.64	0.5	0.64	0.5
RF	0.66	0.62	0.66	0.59
KNN	0.64	0.61	0.64	0.57
SVM	**0.67**	**0.66**	**0.67**	0.59
XGBoost	0.66	0.62	0.66	**0.62**
BERT	0.64	0.48	0.64	0.5

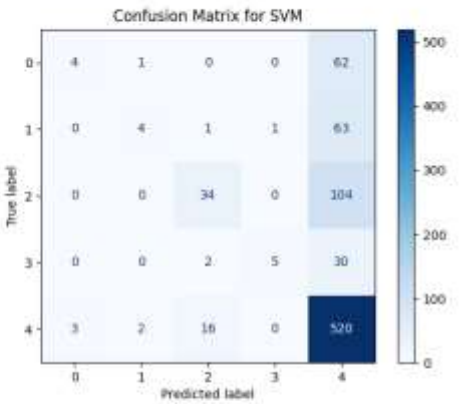

Figure 2. Confusion matrix heatmap of the best performance of classification algorithms (SVM)

This study provides an example to use text classification to identify SSR phases using speech transcription from authentic face-to-face engineer practice course. The model achieved a comparatively good performance (accuracy of 0.67) considering our dataset is small and there were five categories. This shows that text classification techniques also work good at speech transcription data not only in text chat data [7, 12, 13]. This would give some clues for learning analytics research in authentic face-to-face practice class.

5. Conclusion and Future Research

Based on the existing SSR phases, another 16 activities were found in the authentic face-to-face practical course in this study. In terms of code schemes used for specific tasks at the fine-grained levels in CPS, researchers need adjust original one based on their specific context.
Though the small group sample made it possible to conduct a deep process-oriented analysis of students SSR behaviour in authentic engineering CPS course, this study is limited by several shortcomings. First, the sample size of this study was small which limits the possibility for a deeper statistical analysis. Second, the code schemes used in this study provides novel insights into the identification of SSR behaviours in authentic face-to-face engineering practice course, which might not be suitable for concept-design courses. The code schemes have not been tested in other languages or cultural learning context, which reduces the generalization of the findings.

For future research, the data models and research methods can be diverse. The data in this study was only speech recordings collected to help analyse sequential patterns of SSR. Multimodal datasets are recommended to be collected to do a deep analysis of the collaboration process in CPS to provide a granular and comprehensive understanding of sequential SSR, such as through speech rates, eye tracking, and body movement [19] In addition, with large language models like ChatGPT becoming prevalent, it is also one research direction about how to combine the CSCL script with large language models in CPS. Peer feedback analysis in CPS using natural language processing display an approach to automated content analysis for extracting specific categories [20], including SSR behaviours.

REFERENCES

[1] Sharma K., Nguyen A., Hong Y., "Self - regulation and shared regulation in collaborative learning in adaptive digital learning environments: A systematic review of empirical studies", British Journal of Educational Technology, City, Publishing House, 2024, 1398-1436.
[2] De Backer L., Van Keer H., Valcke M., "The functions of shared metacognitive regulation and their differential relation with collaborative learners' understanding of the learning content", Learning and Instruction, City, Publishing House, 2022, 77, 101527.
[3] Chien Y.-H., Liu C.-Y., Chan S.-C., Chang Y.-S., "Engineering design learning for high school and college first-year students in a STEM battlebot design project", International Journal of STEM Education, City, Publishing House, 2023, 10(1).
[4] Lyu Q., Chen W., Su J., Heng K. H., "Collaborate like expert designers: An exploratory study of the role of individual preparation activity on students' collaborative learning", The Internet and Higher Education, City, Publishing House, 2023, 59, 100920.
[5] Lämsä J., Hämäläinen R., Koskinen P., Viiri J., Lampi E., "What do we do when we analyse the temporal aspects of computer-supported collaborative learning? A systematic literature review", Educational Research Review, City, Publishing House, 2021, 33, 100387.
[6] Vaughan N., Wah J. L., "The Community of Inquiry framework: Future practical directions shared metacognition", International Journal of E-Learning & Distance Education/Revue internationale du e-learning et la formation à distance, City, Publishing House, 2020, 35(1).
[7] Flor M., Andrews - Todd J., "Towards automatic annotation of collaborative problem - solving skills in technology - enhanced environments", Journal of Computer Assisted Learning, City, Publishing House, 2022, 38(5), 1434-1447.
[8] Lyu Q., Chen W., Su J., Heng K. H., "Collaborate like expert designers: An exploratory study of the role of individual preparation activity on students' collaborative learning", The Internet and Higher Education, City, Publishing House, 2023, 59, 100920.
[9] Fehrman S., Watson S. L., "A systematic review of asynchronous online discussions in online higher education", American Journal of Distance Education, City, Publishing House, 2021, 35(3), 200-213.
[10] Iiskala T., Volet S., Lehtinen E., Vauras M., "Socially shared metacognitive regulation in asynchronous CSCL in science: Functions, evolution and participation", Frontline Learning Research, City, Publishing House, 2015, 3(1), 78-111.
[11] Minaee S., Kalchbrenner N., Cambria E., Nikzad N., Chenaghlu M., Gao J., "Deep learning--based text classification: a comprehensive review", ACM Computing Surveys (CSUR), City, Publishing House, 2021, 54(3), 1-40.
[12] Ba S., Hu X., Stein D., Liu Q., "Assessing cognitive presence in online inquiry - based discussion through text classification and epistemic network analysis", British Journal of Educational Technology, City, Publishing House, 2023, 54(1), 247-266.
[13] Zheng L., Long M., Niu J., Zhong L., "An automated group learning engagement analysis and feedback approach to promoting collaborative knowledge building, group performance, and socially shared regulation in CSCL", International Journal of Computer-Supported Collaborative Learning, City, Publishing House, 2023, 18(1), 101-133.
[14] Pugh S., Subburaj S. K., Rao A. R., Stewart A. E., Andrews-Todd J., D'Mello S. K., "Say what? Automatic modeling of collaborative problem solving skills from student speech in the wild", Proceedings of the 14th International Conference on Educational Data Mining, International Educational Data Mining Society, City, Publishing House, 2021, 55-67.

[15] De Backer L., Van Keer H., Moerkerke B., Valcke M., "Examining evolutions in the adoption of metacognitive regulation in reciprocal peer tutoring groups", Metacognition and Learning, City, Publishing House, 2016, 11, 187-213.

[16] Braun V., Clarke V., "Using thematic analysis in psychology", Qualitative Research in Psychology, City, Publishing House, 2006, 3(2), 77-101.

[17] Zabolotna K., Malmberg J., Järvenoja H., "Examining the interplay of knowledge construction and group-level regulation in a computer-supported collaborative learning physics task", Computers in Human Behavior, City, Publishing House, 2023, 138, 107494.

[18] Hassan S. U., Ahamed J., Ahmad K., "Analytics of machine learning-based algorithms for text classification", Sustainable Operations and Computers, City, Publishing House, 2022, 3, 238-248.

[19] Spikol D., Ruffaldi E., Dabisias G., Cukurova M., "Supervised machine learning in multimodal learning analytics for estimating success in project - based learning", Journal of Computer Assisted Learning, City, Publishing House, 2018, 34(4), 366-377.

[20] Castro M.S.d.O., et al., "Understanding Peer Feedback Contributions Using Natural Language Processing", In: Viberg O., Jivet I., Muñoz-Merino P., Perifanou M., Papathoma T. (eds) Responsive and Sustainable Educational Futures, EC-TEL 2023, Lecture Notes in Computer Science, Springer, Cham, 2023, 14200.