International Society of the Learning Sciences

My Partner was a Good Partner: Investigating the Relationship between Dialogue Acts and Satisfaction among Middle School Computer Science Learners

Gloria Ashiya Katuka, University of Florida, gkatuka@ufl.edu
Richard T. Bex, University of Florida, rbex@ufl.edu
Mehmet Celepkolu, University of Florida, mckolu@ufl.edu
Kristy Elizabeth Boyer, University of Florida, keboyer@ufl.edu
Eric Wiebe, North Carolina State University, wiebe@ncsu.edu
Bradford Mott, North Carolina State University, bwmott@ncsu.edu
James Lester, North Carolina State University, lester@ncsu.edu

Abstract: Collaborative dialogue provides a rich information source for understanding the effectiveness of student interactions. While many studies emphasize the importance of productive dialogue behaviors, the impact of those behaviors on learners' perceptions of their partners is not yet understood. This paper examines a dialogue corpus of 18 pairs of middle school students as they engage in block-based coding activities. We tagged the corpus with a collaborative dialogue act taxonomy and identified sequences of one to two dialogue acts (*n*-grams) that are significantly associated with partner satisfaction during collaborative learning. Six *n*-grams were found to be significant predictors: *n*-grams that were positively associated with satisfaction included some questions and clarifications. In contrast, *n*-grams that were negatively associated with satisfaction included off-task utterances, pairs of consecutive questions, and unexpectedly, positive feedback. These findings contribute to our understanding of how learners prefer to interact with their partners and how that interaction impacts collaborative experiences.

Introduction

Collaborative dialogue constitutes one of the main channels for students to exchange information and co-construct knowledge (Wegerif, 2011; Mercer et al., 2019; Major et al., 2018) and has attracted considerable interest among computer-supported collaborative learning (CSCL) researchers (e.g., Madaio et al., 2017; Stahl, 2015; Rosé et al., 2008). Dialogue provides numerous cues and opportunities for understanding the effectiveness of collaboration, and thus there is a growing body of research concerning the types of dialogue behaviors that lead to better learning (Chi & Wylie, 2014). On the other hand, analyzing collaborative dialogue is a challenging process due to the dynamics and complexity of group interactions. There is still a need for developing instruments and methodologies to understand how certain dialogue moves occur and how they impact students' learning (Howe, 2017; Hennessy et al., 2016). In recent years, CSCL research has investigated collaborative dialogue for understanding students' socio-metacognitive dialogue patterns (Borge et al., 2019), dialogue transactivity and epistemic quality (Schmitt & Weinberger, 2017), reasoning processes (Snyder et al., 2019), and how students express and address uncertainty (Rodríguez et al., 2017).

In this paper, we extend this body of research by investigating the collaborative dialogue patterns that lead to higher partner satisfaction among middle school students in the context of pair programming. In pair programming, students take on structured roles: the driver's role is to control the mouse and keyboard and focus their effort on building and editing code, while the navigator's role is to observe the work being done by the driver to identify potential errors, provide suggestions, and ask clarification questions (Williams & Kessler, 2003). Pair programming holds great promise for supporting students' learning and engagement in K-12 settings (Campe et al., 2020; Denner et al., 2014), yet, several studies have reported that the demanding nature of collaborative learning can lead to challenges for younger learners, who lack effective collaboration skills (Deitrick et al., 2016; Lewis & Shah, 2015). If these challenges are not addressed, students may develop negative dispositions toward collaboration in the future (Schultz et al., 2010). Thus, it is crucial to deeply examine student dialogues during collaborative learning activities to reveal what kind of dialogue patterns are present and how those dialogue patterns are related to learners' perceptions of their partners. Identifying dialogue patterns that are predictive of learners' satisfaction with their partners can help researchers and educators to understand and facilitate more positive collaborative learning experiences. This study focuses on two research questions: (1) *What dialogue acts emerge during collaborative dialogue within pairs of middle school students during coding activities*, and (2) *How are the dialogue acts associated with outcomes related to partner satisfaction?*

To investigate these research questions, we first developed a taxonomy consisting of 15 dialogue acts, which provides a high-level representation of the underlying meaning of student dialogues, based on a corpus of collaborative dialogue from 36 middle school students who completed a coding activity in pairs. Next, we examined sequences of dialogue acts of length one to two (*n*-grams) and generated a linear regression model which used the frequency of the *n*-grams to predict partner satisfaction, with the *n*-grams as predictors of a derived *satisfaction* outcome (the average of partner-related post-survey items). The results showed significant associations between six dialogue *n*-grams and the learners' satisfaction with their partners. Learners reported higher partner satisfaction when their partners were more engaged, such as by asking questions, seeking clarifications, and actively talking about the task. The results also showed that when their partner frequently responded with positive feedback or when both collaborators engaged in off-task dialogues, learners reported lower partner satisfaction. These findings provide us with a better understanding of how learners would prefer their partners to interact with them and how they prefer to interact with their partners when participating in pair programming activities.

Background: Dialogue Analysis

Within dialogue analysis, representing conversations at the utterance level, such as through *dialogue acts*, has long been studied. Dialogue acts are a higher-level representation of the intention of the user (Austin, 1975), and dialogue act tagging involves labeling an utterance with a predefined dialogue act that provides information about that utterance. Each dialogue turn is considered as one utterance; thus, an utterance can serve as a smaller unit of communication that describes a single event (Polanyi et al., 2004). An utterance can be an incomplete or grammatically broken sentence but still have a role in conversation depending on the context (Bakhtin, 2010). A dialogue act expresses the nature of a communicative behavior between a sender and addressee that has an effect on the context of understanding the behavior (Bunt, 2005). Previous research has investigated the ways in which dialogue acts are associated with learning outcomes (Dubovi & Lee, 2019; Olsen & Finkelstein, 2017) and motivation (Meier et al., 2007). In this paper, the goal of dialogue act tagging is to classify the utterances to show collaborative patterns that are associated with partner satisfaction.

Methods

Participants and context

This work is part of a larger project aimed at developing computer science knowledge and deepening understanding of science concepts through computationally rich science activities for middle school students (Celepkolu et al., 2020). To achieve this goal, the research team collaborated with a middle school science teacher to implement a series of computer coding lessons as part of their regular classroom activities. The students learned about the fundamentals of coding, such as loops, conditionals, and variables, and applied their coding knowledge to create science models and simulations, such as homeostasis and evolution, using the Snap*!* block-based programming environment. The researchers explained the driver and navigator roles in pair programming, the expectations for each role, and reminded students to switch roles regularly (12-15 minutes). Data was collected as part of an IRB-approved study that included written parental consent and student assent. The researchers implemented the activities during a science class in two semesters (Spring and Fall 2019), which was taught by the same teacher and followed the same structure. Out of 204 students, 145 students provided assent and parental consent, and we randomly selected 19 pairs (38 students) to audio/video record their interactions during the coding activities (24 students in Spring 2019 and 14 students in Fall 2019). Out of these students, there were 23 girls (60.5%) and 15 boys (39.5%). The distribution of race/ethnicities was 14 White (36.8%), 2 Hispanic (5.3%), 7 Asian (18.5%), 10 Multiracial (26.2%), and 5 Other (13.2%). The mean age was 12.1, with ages ranging from 11 to 13, and 53% of students reported having had some prior coding experience at the beginning of the semester.

Procedure

In every class, researchers assisted the teacher by presenting an introduction to the science topics and providing students with a copy of the written instructions. Next, students worked on activities for 35-40 minutes with a randomly assigned partner. During these activities, the teacher and researchers were available to help students with their questions. After pairs participated in the collaborative work sessions, students were asked to individually complete a post survey. We developed the post survey items because, to the best of our knowledge, there is no existing survey that captures partner satisfaction within the pair programming context. From the post survey, this paper utilizes the following six questions for analysis: *(1) "My partner answered my questions well," (2) "My partner listened to my suggestions," (3) "My partner often cut my speech" (which was reversed scored* *for accurate computation), (4) "My partner was comfortable asking me questions," (5) "My partner asking questions helped me think about things differently," and (6) "Overall, my partner was a good partner."* Responses followed a 5-point Likert scale, with 1 representing "strongly disagree" and 5 representing "strongly agree." Figure 1 shows the distribution of the survey responses from both studies. Most students agreed or strongly agreed that their collaboration with their partner was successful.

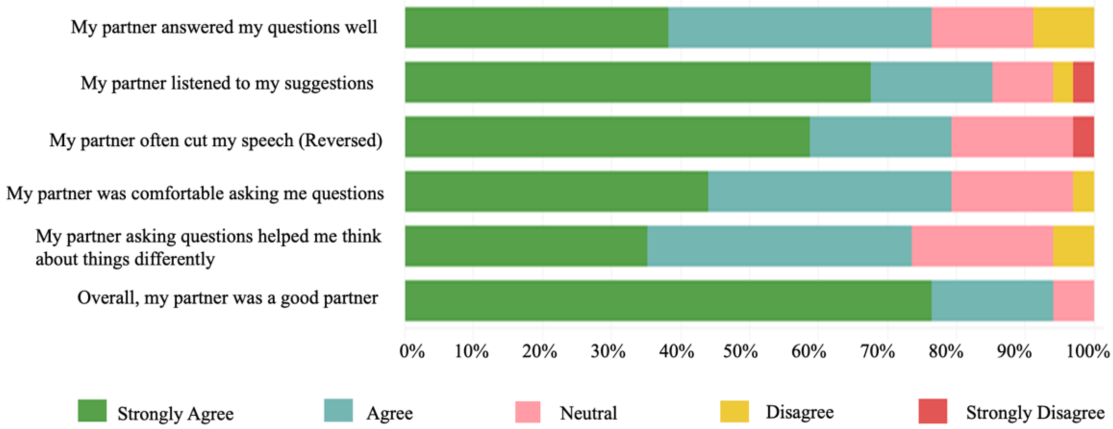

Figure 1. Post survey responses related to partner satisfaction

Dialogue Corpus and Annotation

We manually transcribed the 19 video recordings of students collaborating, which resulted in 8,940 dialogue utterances. Next, we tagged each utterance based on the function of the information in the dialogue. Prior to tagging our dataset, we filtered and removed utterances directed toward anyone other than the learner's partner (teachers, researchers, and other students). Next, we removed one session that contained large amounts of chatter and indistinguishable utterances. Lastly, we also removed all utterances that were untranscribable due to audio quality. Our final student-student dialogue corpus included 18 sessions (36 students) and 4,859 utterances with a mean of 242 utterances per session (SD = 118, Min = 93, Max = 526) and a mean of 121 utterances per student (SD = 114, Min = 42, Max = 264).

To develop a taxonomy for our corpus, we reviewed the existing taxonomies within closely related fields and age groups and considered relevant taxonomies (Core & Allen, 1997; Rodríguez et al., 2017; Tsan et al., 2018). Previous research has established dialogue act tags for pair programming among college students (Rodríguez et al., 2017) as well as among elementary school students (Tsan et al., 2018), and our work took the union of these two taxonomies as its starting point, producing 19 initial tags. Several iterations of dialogue tag application and refinement revealed, as expected, that some of the tags from the taxonomies were not present in the current middle school corpus, and that some new or modified tags were needed. A process of iterative refinement of the tagging scheme in several rounds of collaborative and then independent tagging produced a final dialogue act taxonomy of 15 tags (Table 1). Eight of these tags were adopted from Rodríguez et al. (2017): *Statement*, *Acknowledgement*, *Uncertain*, *Meta comment*, *Positive Feedback*, *Non-Positive Feedback* and *Off-task*. From the Tsan et al. (2018) scheme, three tags were adopted: *Make Suggestions*, *Acceptance*, and *Rejection*. The newly developed tags are *Next Step*, *Seek Clarification*, *Question*, and *Seeking Attention*. Both annotators independently tagged 23% of the dataset and achieved an inter-rater agreement score Cohen's kappa of .83 (Landis & Koch, 1977) indicating "almost perfect" agreement. The two annotators then each tagged half of the remaining utterances so that the entire corpus was tagged.

Data Analysis

Our next goal was to discover the ways in which student dialogue acts were related to the outcomes reported on six post survey items. To determine whether to treat these six post survey items as a single item or multiple items, we conducted a principal component analysis (PCA). The results of PCA suggested proceeding with only one derived outcome variable, which we refer to as *satisfaction*. The single component explains 52% of the variation across all six survey items with eigenvalue 3.15. The distribution of the *satisfaction* outcome shows 76% of the learners agreed or strongly agreed that they were satisfied with the overall interaction with their partner.

Our overarching goal was to identify the ways in which dialogue acts (or sequences of them) were associated with partner satisfaction. From our tagged dialogue corpus, we proceeded to extract sequences of dialogue acts, known as *n*-grams, which will be treated as predictors within a regression model. To extract the sequence of *n*-grams, we applied standard practices from previous dialogue analyses (Forbes-Riley & Litman, 2005). In our work, we generated *n*-grams of dialogue act tags for each learner's dialogue using a sliding window of *n*=1 (unigrams) and *n*=2 (bigrams). We used a sliding window approach and assigned each dialogue act tag a student or partner subscript (*e.g.*, $Statement_{stu}$, $Question_{par}$) to indicate whether the utterance originated from the student or their partner. Every learner was tagged as a student to ensure that we extracted each *n*-gram from each learner's perspective. Each row in the resulting dataset corresponds to a student whose own dialogue moves contain the subscript "*stu*" and whose respective partner's dialogue moves contain the subscript "*par.*" Each student played the role of a "driver" as well as a "navigator" during the pair programming task, and these roles are not indicated within the bigrams. We extracted 563 *n*-grams, 30 distinct unigrams and 533 distinct bigrams. Unigram frequencies are shown in Table 1. The most frequent bigrams were ($Statement_{stu}$, $Statement_{par}$), ($Question_{stu}$, $Statement_{par}$), and ($Statement_{stu}$, $Question_{par}$), which occurred 677, 291, and 235 times, respectively.

Table 1: Taxonomy of Dialogue Acts

Dialogue Act	**Frequency**	**Description**	**Example(s)**
Statement	1622	Makes a statement of information, an explanation, or a response to an inquiry	*"This looks like it's not moving at all."* *"Oh, we forgot to put repeat forever."*
Off-task	733	Interacts with someone other than their partner or off-topic conversations with their partner	*"I have a, we have an orchestra test today."* *"You like my new look?"*
Question	604	Asks partner for help or information seeking some feedback from the partner with regards to the task.	*"Do I put this in here?"* *"Do we just have to put it together now?"*
Directive	405	Provides an explicit instruction to their partner	*"Push the restart button."* *"Click on the amplitude variable."*
Acknowledgment	216	Accepts or acknowledges the previous statement or utterance	*"Okay."*
Meta Comment	193	Makes a meta response to something relating to the task	*"Um, uh…"* *"Oh my gosh."*
Uncertainty	133	States an opinion or indication of uncertainty or confusion	*"Maybe. I don't know."* *"I'm a little confused."*
Seek Clarification	112	Asks for further clarification on something mentioned earlier or referred to in the text	*"What?"* *"Which one?"* *"What do you mean?"*
Positive Feedback	88	Provides positive feedback related to a task action completed by themselves or their partner	*"There! We finally did it."* *"Oh, ours is good."* *"Yeah that's good, it's good."*
Make Suggestion	81	Makes a suggestion or contributes an idea without explicitly asking the partner to do something	*"Maybe make a new forever loop just for that."* *"Let's go back to the directions because it will tell us what code to use."*
Non-positive feedback	54	Provides negative feedback on the task or something incorrectly done by themselves or their partner	*"Wait, try the, oh that's not gonna work."* *"We don't need that."*
Next Step	52	Makes a suggestion for what they believe should be the next step to be completed in the near future	*"And then I think you're supposed to put it in."* *"And then change variables.."* *"And then we can do the operators."*
Acceptance	44	Accepts or acknowledges their partner's idea, suggestion, or directive. (Follows a MS, NS, or D)	*"Yes."* *"Right."*
Seeking Attention	12	Seeks partner's attention while working on task	*"Hello?"* *"Bro."*
Rejection	10	Rejects a direct instruction or idea or suggestion	*"No."*

Results

To determine the *n*-grams that were significant predictors of the *satisfaction* outcome, we conducted a regression analysis using the JMP statistical software. To mitigate the problem of a large number of *n*-gram predictors for a smaller sample size (563>>36) we included only *n*-grams that occurred in at least half of the sessions. The remaining 78 predictors included 30 unigrams and 48 bigrams. We provided these 78 *n*-grams as predictors and the derived *satisfaction* variable as the outcome variable to a generalized regression model. We selected the best subset estimation method, which uses an exhaustive algorithm that fits and assesses all possible models and chooses the best subset to predict the outcome variable. We used the AIC (Akaike information criterion) statistic as the goodness-of-fit measure. Table 2 shows the regression results, including the six *n*-grams that satisfied the test for statistical significance ($p < .05$). The regression model passes the test for multicollinearity with all resulting variance inflation factor (VIF) values less than 2 (VIF values greater than 5 often indicate multicollinearity). The adjusted R^2 of .74 shows that the model explains 74% of the variance in partner satisfaction.

Table 2: Generalized Regression Model (Best subset method) of *n*-grams as predictors of partner satisfaction

Dialogue Act *n*-gram	**Estimate**	**Standardize Estimate**	**Std Error**	**VIF**
Intercept	3.853	1.477	0.126	0
Question_{stu}, $\text{Seek Clarification}_{par}$	1.945	0.601	0.091	1.99
Directive_{stu}, Question_{par}	1.092	0.209	0.025	0.878
Statement_{par}	2.119	0.681	0.002	1.122
$\text{Positive Feedback}_{par}$	-1.406	-0.938	0.019	1.817
Question_{par}, Question_{stu}	-1.736	-0.941	0.018	1.162
Off-task_{stu} Off-task_{par}	-1.413	-0.941	0.002	1.064

Note: The model only contains significant *n*-gram predictors with $p<.001$.

As the parameter estimates in Table 2 show, three *n*-grams are positively related to partner satisfaction: (1) a question by the student followed by their partner seeking clarification, (2) a directive by the student followed by a question from their partner, and (3) a statement from their partner. In contrast, the model also revealed that three *n*-grams are negatively associated with partner satisfaction: (1) the student initiates a conversation not related to the task and their partner responds and continues with the unrelated conversation, (2) positive feedback from their partner, and (3) a question from their partner followed by a question from the student.

Discussion and Implications

The overarching goal of this study was to develop a better understanding of how dialogue acts are associated with partner satisfaction for middle school students during collaborative coding. The model identified six statistically significant *n*-grams, including three bigrams and one unigram that may indicate an interactive partnership: ($\textit{Question}_{stu}$, $\textit{Seek Clarificaton}_{par}$), ($\textit{Directive}_{stu}$, $\textit{Question}_{par}$), ($\textit{Question}_{par}$, $\textit{Question}_{stu}$), and $\textit{Statement}_{par}$. Conversely, ($\textit{Off-task}_{stu}$, $\textit{Off-task}_{par}$) and—perhaps counterintuitively—$\textit{Positive Feedback}_{par}$, may suggest a tendency for reduced engagement or distractions in this context. This section discusses these findings in turn.

Three of the significant *n*-grams within the model include asking questions. The literature has clearly established the role of questions in collaborative learning as a means of establishing and sustaining mutual understanding (Spada et al., 2005). Asking questions also elicits a constructive engagement between collaborators by presenting an avenue to generate new ideas (Chi & Wylie, 2014). When learners ask their partner questions, they create a channel for dialogue interaction by taking the first step to access information and resolve confusion (Chin & Osborne, 2008). The analysis results indicate that question-related dialogues are significant indicators of partner satisfaction. Higher occurrence of bigrams where learners ask their partner a question followed by their partner seeking clarification are associated with that learner reporting higher partner satisfaction. This finding is likely related to the importance of understanding a question before attempting to answer it. For example, one student said, *"Why is the wavelength N-A-N?"*($\textit{Question}_{stu}$) and their partner replied, by *"Nan what?"*($\textit{Seek Clarification}_{par}$). Here, the partner is making an effort to better understand the student, and the student subsequently reported higher satisfaction with that partner. Similarly, the results show a positive correlation between partner satisfaction and higher occurrences of instances when a partner asks a question after receiving a directive/instruction from the student. For example, a student said, *"Okay. Now, create a variable."*($\textit{Directive}_{stu}$) and their partner asked, *"Named what?"*($\textit{Question}_{par}$). Here the question *"Named what?"*($\textit{Question}_{par}$) refers to seeking new information. This is different from seeking clarification, which refers to a question or information already stated previously.

Not all occurrences of *Question* were positively associated with the satisfaction outcome. The more often collaborators asked back-to-back questions without a response to the first question, the less likely the student reported a high satisfaction rating. For example, a student asked, *"Why did it set the generation to zero?"(Question*$_{stu}$*)* and their partner asked, *"So, wait is this what we're supposed to do?"(Question*$_{par}$*)*. Unanswered questions and unresolved uncertainty have been linked to less positive outcomes in other work on collaborative coding for dialogue as well (Rodríguez et al., 2017).

Another dialogue act whose frequency was highly predictive of satisfaction is a statement from the partner. Statements are one of the most prominent conversational dialogue moves in the corpus. The findings in this study are consistent with previous results where statements were shown to be associated with effective collaboration (Rodríguez et al., 2017). Statements can indicate more active engagement by the partner, which improves learning outcomes by facilitating advancement from *constructive* to *interactive* behavioral modes (Chi & Wylie, 2014). In the current corpus, the most frequent occurrences of statements were in response to a question, directive, or acknowledgment by the student. For example, a student asked, *"Right or does it not get longer?" (Question*$_{stu}$*)* and their partner responded with *"It doesn't get longer."(Statement*$_{par}$*)*. The second most common occurrence of a statement from the partner is as a response to a directive from the student. For example, one student said, *"Wait, increase the clone counter by one,"(Directive*$_{stu}$*)* and their partner responded, *"I think they do the same thing."(Statement*$_{par}$*)*. In the third most common occurrence of a statement, a statement is followed by an acknowledgment. For example, an *"Okay" (Acknowledgement*$_{stu}$*)* from the student was followed by *"Um, when a clone is spawned it should increase the clone generation counter by one, uh clone generation counter…" (Statement*$_{par}$*)* by their partner.

In addition to the previously mentioned sequence of back-to-back *Questions*, two others emerged as negative predictors of partner satisfaction: *Positive Feedback*$_{par}$, and (*Off-task*$_{stu}$*, Off-task*$_{par}$*)*. Feedback within peer collaboration has been shown to positively enhance interpersonal behaviors and social performance (Phielix et al., 2010), but a potential explanation for the negative association of a partners' positive feedback with a student's perception of that partner might be the possibility of the partner compensating for lower participation, which eventually becomes apparent (Prinsen et al., 2007). In the corpus, the most common occurrence of positive feedback by the partner followed a statement. For example, one student said, *"Now, we're going to do this. There we go." (Statement*$_{stur}$*)* and their partner responded with *"Yay."(Positive Feedback*$_{par}$*)*. It is also possible that a positive feedback response might function as the partner doubtfully accepting the student's assertions. This could be due to the partner not fully understanding their role, the task, or their ability to effectively contribute to the collaboration.

As for the off-task bigram's role in predicting partner satisfaction, recent CSCL research investigated the impact of off-task exchanges during collaborative problem solving such as lower participation and distraction from the task (Cheng et al., 2020). In the corpus, we see threads of *Off-task* utterances that can pose a distraction to the collaborators. This may result in the collaborators not completing their tasks and lower satisfaction in their interaction. For example, the utterance, *"It's so surprising because my parents don't believe in bath and body works."(Off-task*$_{stu}$*)* by the student followed by *"Really? They don't believe in bathworks."* (*Off-task*$_{par}$*)* by their partner sets the tone for more off-task dialogue. This exchange shows a mutual distraction between collaborators that can deviate the conversation from the task at hand. These pairs of off-task utterances are associated with lower partner satisfaction in the current context.

Implications

The findings discussed here hold several potential implications for research and practice. The findings have shown that dialogue moves indicating an interactive give-and-take, including questions, clarification questions, and elaboration, are positively associated with partner satisfaction while other phenomena such as sequences of off-task dialogue acts are negatively associated. Some seemingly positive moves, such as positive feedback from the partner, were negatively associated with a learner's satisfaction with that partner, and these phenomena warrant deeper investigation for several reasons. For example, they tell us that as we move toward using natural language processing to automatically analyze and support real-time collaboration, we must take great caution in interpreting utterances at face value: positive sentiment, whether in on-task or off-task utterances, may express a wide variety of underlying states and different levels of engagement. Additionally, while a tremendous body of literature shows the importance of certain dialogue moves including question asking, the results here suggest that the ways in which these questions are incorporated into collaborative dialogue could have a significant impact on outcomes.

Conclusion

The overarching goal of this study was to explore the relationship between dialogue patterns and partner satisfaction during pair programming activities. The findings suggest that collaborative dialogue acts that reflect interactive partnerships and active participation between learners are associated with higher satisfaction ratings, whereas dialogue acts that reflect lower participation and distraction during collaborative activities are associated with lower satisfaction ratings. This research contributes to a better understanding of the ways in which learners' and their partners' interaction during CSCL activities impact the collaborative learning process. Several limitations of this work are important to note. First, our resulting model uses partner satisfaction as its primary outcome, rather than measures of learning or process-oriented metrics of collaboration. This intentional choice is due to the importance of learners' affective and motivational states during collaboration, for which satisfaction with a partner is an important component. Second, a limitation of this work is that the relationship between dialogue acts and partner satisfaction is correlational and not causal. Finally, an additional limitation is that the studies were only conducted with middle school students from the southeastern United States and important cultural differences in other contexts may influence the generation of dialogue moves and findings.

There are several promising directions for future work. First, while this work investigated the relationship between dialogue patterns and partner satisfaction, it is also important to examine whether these patterns are also associated with learning outcomes or process-level collaborative metrics. Secondly, the relationship between partner satisfaction and effective learning outcomes should be further examined. Moreover, there is a need for examining the dialogue patterns for different pair compositions by characteristics such as gender, experience level, and personality. Additionally, deeper qualitative analysis can shed further light on how these dialogue patterns influence partner satisfaction. Finally, these findings can inform the design of adaptive support for computer-supported collaborative learning technologies, which use rich data from student dialogues.

References

Austin, J. L. (1975). *How to Do Things with Words* (2nd ed.). Clarendon Press.

Bakhtin, M.M. (2010). *Speech Genres and Other Late Essays*. University of Texas Pres.

Borge, M., Aldemir, T., & Xia, Y. (2019). Unpacking Socio-Metacognitive Sense-Making Patterns to Support Collaborative Discourse. *In Proceedings of Computer Supported Collaborative Learning (CSCL),* 320-327.

Bunt, H. (2005). A framework for dialogue act specification. *In Proceedings of SIGSEM WG on Representation of Multimodal Semantic Information,* Tilburg, Netherlands.

Campe, S., Denner, J., Green, E., & Torres, D. (2020). Pair programming in middle school: variations in interactions and behaviors. *Computer Science Education, 30*(1), 22–46.

Celepkolu, M., Fussell, D. A., Galdo, A. C., Boyer, K. E., Wiebe, E. N., Mott, B. W., & Lester, J. C. (2020). Exploring Middle School Students' Reflections on the Infusion of CS into Science Classrooms. *In Proceedings of the 51st ACM Technical Symposium on Computer Science Education* (pp. 671-677).

Cheng, X., Fu, S., de Vreede, G.-J., & Li, Y. (2021). Using Collaboration Engineering to Mitigate Low Participation, Distraction, and Learning Inefficiency to Support Collaborative Learning in Industry. *Group Decision and Negotiation, 30,* 171-190.

Chi, M. T. H., & Wylie, R. (2014). The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes. *Educational Psychologist, 49*(4), 219–243.

Chin, C., & Osborne, J. (2008). Students' questions: a potential resource for teaching and learning science. *Studies in Science Education, 44*(1), 1-39.

Core, M. G., & Allen, J. (1997). Coding dialogs with the DAMSL annotation scheme. *In AAAI Fall Symposium on Communicative Action in Humans and Machines* (Vol. 56, pp. 28-35).

Deitrick, E., Shapiro, R. B., & Gravel, B. (2016). How do we assess equity in programming pairs? *In Proceedings of the 12th International Conference of the Learning Sciences (ICLS)* (pp. 370–377).

Denner, J., Werner, L., Campe, S., & Ortiz, E. (2014). Pair Programming: Under What Conditions Is It Advantageous for Middle School Students? *Journal of Research on Technology in Education, 46*(3), 277–296.

Dubovi, I., & Lee, V. (2019). Comparing the Effectiveness of Supports for Collaborative Dialogic Sense-Making with Agent-Based Models. *In Proceedings of Computer Supported Collaborative Learning (CSCL)* (pp. 88-95).

Forbes-Riley, K., & Litman, D. J. (2005). Using bigrams to identify relationships between student certainness states and tutor responses in a spoken dialogue corpus. *Proceedings of the 6th SIGDIAL Workshop on Discourse and Dialogue* (pp. 87-96).

Hennessy, S., Rojas-Drummond, S., Higham, R., Márquez, A. M., Maine, F., Ríos, R. M., García-Carrión, R., Torreblanca, O., & Barrera, M. J. (2016). Developing a coding scheme for analysing classroom dialogue across educational contexts. *Learning, Culture and Social Interaction, 9*, 16–44.

Howe, C. (2017). Advances in research on classroom dialogue: Commentary on the articles. *Learning and Instruction, 48,* 61-65.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174.

Lewis, C. M., & Shah, N. (2015). How Equity and Inequity Can Emerge in Pair Programming. *Proceedings of the Eleventh Annual International Conference on International Computing Education Research,* 41–50.

Madaio, M., Cassell, J., & Ogan, A. (2017). The impact of peer tutors' use of indirect feedback and instructions. *Proceedings of the 12th International Conference on Computer Supported Collaborative Learning (CSCL),* 383–390.

Major, L., Warwick, P., Rasmussen, I., Ludvigsen, S., & Cook, V. (2018). Classroom dialogue and digital technologies: A scoping review. *Education and Information Technologies, 23*(5), 1995–2028.

Meier, A., Spada, H., & Rummel, N. (2007). A rating scheme for assessing the quality of computer-supported collaboration processes. *International Journal of Computer-Supported Collaborative Learning, 2*(1), 63–86.

Mercer, N., Hennessy, S., & Warwick, P. (2019). Dialogue, thinking together and digital technology in the classroom: Some educational implications of a continuing line of inquiry. *International Journal of Educational Research, 97*, 187–199.

Olsen, J. K., & Finkelstein, S. (2017). Through the (thin-slice) looking glass: An initial look at rapport and co-construction within peer collaboration. *Proceedings of the 12th International Conference on Computer Supported Collaborative Learning (CSCL)* (pp. 511–518).

Phielix, C., Prins, F. J., & Kirschner, P. A. (2010). Awareness of group performance in a CSCL-environment: Effects of peer feedback and reflection. *Computers in Human Behavior, 26*(2) 151–161.

Polanyi, L., Culy, C., van den Berg, M., Thione, G. L., & Ahn, D. (2004). Sentential structure and discourse parsing. *In Proceedings of the 2004 ACL Workshop on Discourse Annotation - DiscAnnotation '04*.

Prinsen, F., Volman, M. L. L., & Terwel, J. (2007). The influence of learner characteristics on degree and type of participation in a CSCL environment. *British Journal of Educational Technology, 38*(6), 1037–1055.

Rodríguez, F. J., Price, K. M., & Boyer, K. E. (2017). Exploring the Pair Programming Process: Characteristics of Effective Collaboration. *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education* (pp. 507–512).

Rosé, C., Wang, Y.-C., Cui, Y., Arguello, J., Stegmann, K., Weinberger, A., & Fischer, F. (2008). Analyzing collaborative learning processes automatically: Exploiting the advances of computational linguistics in computer-supported collaborative learning. *International Journal of Computer-Supported Collaborative Learning, 3*(3), 237–271.

Schmitt, L. J., & Weinberger, A. (2017). Collaborative learning on multi-touch interfaces: scaffolding elementary school students. *Proceedings of the 12th International Conference on Computer Supported Collaborative Learning (CSCL)* (pp. 9–16).

Schultz, J. L., Wilson, J. R., & Hess, K. C. (2010). Team-Based Classroom Pedagogy Reframed: The Student Perspective. *American Journal of Business Education, 3*(7), 17–24.

Snyder, C., Biswas, G., Emara, M., Grover, S., & Conlin, L. (2019). Analyzing students' synergistic learning processes in physics and CT by collaborative discourse analysis. *Proceedings of the 13th International Conference on Computer Supported Collaborative Learning (CSCL)* (pp. 360–367).

Spada, H., Meier, A., Rummel, N., & Hauser, S. (2005). A new method to assess the quality of collaborative process in CSCL. *Proceedings of Computer Supported Collaborative Learning (CSCL)* (pp. 622–631).

Stahl, G. (2015). A decade of CSCL. *International Journal of Computer-Supported Collaborative Learning, 10*(4), 337–344.

Tsan, J., Lynch, C. F., & Boyer, K. E. (2018). "Alright, what do we need?": A study of young coders' collaborative dialogue. *International Journal of Child-Computer Interaction, 17*, 61–71.

Wegerif, R. (2011). Towards a dialogic theory of how children learn to think. *Thinking Skills and Creativity, 6*(3), 179–190.

Williams, L., & Kessler, R. R. (2003). *Pair Programming Illuminated*. Addison-Wesley Professional.

Acknowledgments

This material is based upon work supported by the National Science Foundation through grant DRL-1640141. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.