

Modeling and Analyzing Collaborative Dialogue States in Mathematics Problem-Solving through ONA and Hidden Markov Models

Shan Zhang, University of Florida, zhangshan@ufl.edu
 Andrés Felipe Zambrano, University of Pennsylvania, azamb13@upenn.edu
 Hongming Li, University of Florida, hli3@ufl.edu
 Seiyon Lee, University of Florida, leeseiyon@ufl.edu
 Ji-Eun Lee, Singapore University of Technology and Design, jieun_lee@sutd.edu.sg
 Anthony F. Botelho, University of Florida, abotelho@coe.ufl.edu

Abstract: Prior research has shown that collaborative dialogue reflects students' cognitive, social, and regulatory processes, which support deeper learning. However, less is known about how these moves unfold sequentially during collaborative problem solving (CPS), especially in math contexts. This study leverages Markov Chains, Hidden Markov Models (HMM), and Ordered Network Analysis (ONA) to capture and analyze recurring patterns and transitions in dialogue that are representative of collaboration dynamics and observe how these relate to students' perceptions of their interactions, collaboration quality, and performance. We conduct our analyses with data from 22 college students (11 dyads) who worked collaboratively to solve algebra problems in a digital learning environment. Results show that dialogue within problems cycled through agreement, clarification, and self-explanation, indicating stable interaction patterns across five HMM latent states. Findings from ONA suggest that while explanatory talk anchors collaboration, task-grounding and co-constructive exchanges may be more directly associated with effective CPS.

Introduction

A student's capacity to express their mathematical reasoning through oral and written communication is considered essential to demonstrating mastery (Partnership for 21st Century Skills, 2007). One effective way to foster such communication is by engaging students in *Collaborative Problem Solving* (CPS), an instructional approach in which groups of learners work together to solve novel problems by sharing their understanding and coordinating their efforts (Graesser et al., 2020; Andrews-Todd & Forsyth, 2023). As CPS involves students engaging in sustained, multimodal interaction (e.g., combinations of verbal and non-verbal exchanges to achieve a shared goal), this complex phenomenon poses challenges for researchers seeking to capture and analyze patterns of collaboration and discourse quality. Prior studies have often examined dialogue in isolation (Tausczik et al., 2014), focusing on individual communicative moves such as building, monitoring, or repairing shared understanding (e.g., Hansen, 2022), rather than exploring how these acts unfold sequentially to shape the dynamics of collaborative reasoning. Understanding how these dialogue acts co-occur and form sequences or latent states is essential, as such patterns reveal the dynamic processes through which students shift between productive and unproductive forms of collaboration (Earle-Randell et al., 2023). With advancements in Computer-supported Collaborative Learning (CSCL) tools and data analytics, researchers have begun applying advanced modeling techniques, such as Hidden Markov Models (HMM) (Earle-Randell et al., 2023) or Ordered Network Analysis (ONA), to capture these temporal and structural dynamics more precisely (Tan et al., 2023). Yet, despite these methodological developments, relatively few studies have applied such approaches within mathematics education, where the cognitive and representational complexity of problem solving demands fine-grained analysis. As a result, considerable uncertainty remains about how best to unpack the intricate interplay of dialogue, reasoning, and regulation that characterizes CPS in this context.

To address these gaps, the present study examines the dynamic structure of collaborative dialogue during mathematical problem solving, focusing on how latent interaction states relate to collaboration quality, prior experience, perception of interaction, and learning outcomes. Specifically, it investigates the following research questions: RQ1: (a). How do observable dialogue acts transition from one to another during collaboration? (b). What latent problem-solving states emerge from students' dialogue during collaborative tasks, and how do students transition among these hidden states? RQ2: To what extent are HMM state proportions associated with the outcomes (i.e., overall learning satisfaction, perceptions of collaboration, problem-solving correctness, collaboration quality) as well as prior math learning collaborative experience and prior collaborative skills? RQ3:

How do patterns of collaborative dialogue acts differ between groups with high versus low quality of collaborative problem-solving?

Methods

This study involved 22 undergraduate students (11 dyads) from a university located in the northeastern U.S. Of these students, 55% ($n = 12$) were male and 45% ($n = 10$) female. In terms of race/ethnicity, 64% ($n = 14$) identified as White, 32% as Asian ($n = 5$), and the remaining students ($n = 3$) as other categories. In this study, we collected data through Graspable Math (GM, <https://activities.graspablemath.com/>), a dynamic mathematics notation platform grounded in perceptual learning and embodied cognition theories. Prior to the task, students completed a pre-survey assessing collaborative experience and skills, followed by a brief practice phase. Dyads then worked together for 30 minutes to solve five algebra problems, after which they completed a post-survey measuring collaboration perceptions and satisfaction. Group-level problem-solving correctness was also computed. Multimodal data, including dialogue transcripts and interaction logs, were collected and temporally aligned to capture collaborative problem-solving processes. In total, 11 groups completed 49 problems, and all analyses were conducted at the group level.

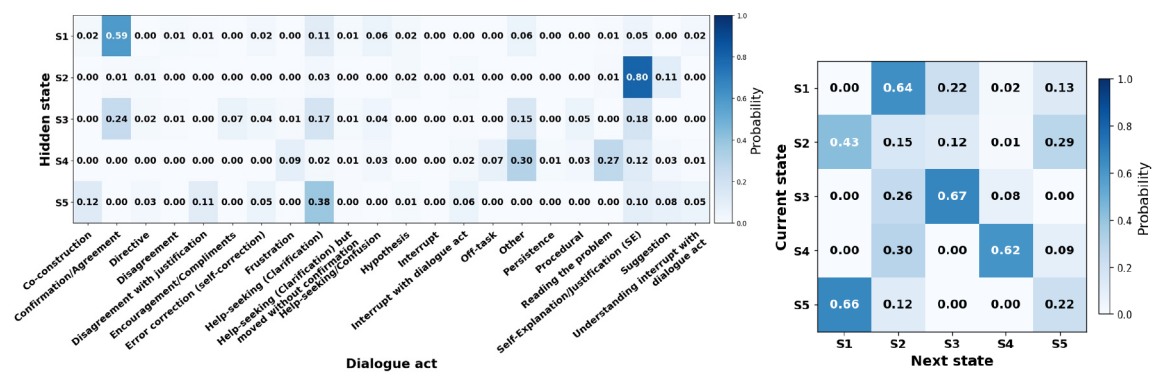
To capture interaction processes, dialogue transcripts were coded using an established scheme (Zhang et al., 2024) encompassing cognitive, metacognitive, social, and affective dimensions ($\kappa > .75$). In parallel, collaboration quality was evaluated using a six-dimension rubric on a three-point scale, achieving high inter-rater reliability (QWK = .84). Each group's overall score was calculated by averaging its total problem scores. Finally, we used the median of these group-level overall scores to create two categories: high-quality collaboration (scores greater than or equal to the median; $N = 6$) and low-quality collaboration (scores below the median; $N = 5$). Detailed descriptions of the learning environment, coding scheme, and rubric are available on Open Science Framework: https://osf.io/583g2/overview?view_only=3120f65ebca541378e5c9d6ce00db05e. To analyze collaboration dynamics, we first applied first-order Markov chains to model transitions between dialogue acts (Thomas et al., 1983), followed by a Hidden Markov Model (Boyer et al., 2011) with five latent states to identify underlying interaction patterns from time-ordered sequences, with model selection based on Bayesian Information Criterion with five-fold cross-validation. We then examined emission probabilities, transition patterns, and run-lengths, and derived group-level features such as state proportions (time spent in each state). Next, to examine associations with outcomes, we computed Spearman correlations between hidden-state proportions and measures of collaboration perceptions, satisfaction, correctness, collaboration quality, and prior experience and skills. Finally, we conducted Ordered Network Analysis using the ONA Web Tool (version 1.7.0; Marquart et al., 2018) at the group level to compare structural dialogue patterns between high- and low-collaboration groups (median split), retaining key dialogue acts based on connection weights ($CW > .02$) and testing group differences using Mann-Whitney U on unit means after means rotation.

Results

In addressing RQ1, the Markov chain analysis revealed systematic and asymmetric transition patterns among dialogue acts. Confirmation/Agreement functioned as a central hub, frequently following Directives ($P = 0.42$), Help-seeking ($P = 0.31$), and Self-Explanation/Justification (SE) ($P = 0.24$). In turn, SE was often followed by Confirmation/Agreement ($P = 0.49$) or Help-seeking ($P = 0.33$), suggesting that confirmation and clarification prompted further reasoning. Help-seeking also commonly emerged after Disagreement with justification ($P = 0.20$), Hypothesis ($P = 0.28$), and Suggestions ($P = 0.15$), supporting the resolution of uncertainty. In contrast, disagreement-related moves were typically short-lived and often transitioned into Agreement ($P = 0.14$ - 0.37), indicating a tendency to quickly re-establish consensus. Less frequent acts (e.g., Frustration, Persistence, Off-task) appeared sporadically and showed limited continuation. Overall, collaboration was characterized by recurring cycles of agreement, help-seeking, and self-explanation.

Figure 1

HMM Emission Means (Left; $N = 5$) and Group-level HMM Transition Probabilities (Right)



HMM analysis identified five latent states underlying CPS (Figure 1). State 1 (run length = 1) reflected brief agreement-dominant exchanges, while State 2 (run length = 1) captured frequent self-explanation and suggestion (SE; $P = 0.80$). State 3 (run length = 4) represented more sustained, mixed interaction combining agreement, clarification, and partial explanation. State 4 (run length = 3) reflected task grounding (e.g., reading, filler talk, occasional frustration), and State 5 (run length = 1) captured help-seeking, co-construction, and justified disagreement. Transition patterns showed strong cycling between States 1 and 2, shifts from State 2 to State 5, and high self-transitions in States 3 and 4, indicating more stable interaction phases.

To address RQ2, as shown in Table 1, the correlation analysis revealed several connections between hidden-state proportions and collaborative-related pre- or post-outcomes. State 4 (Reading the problem/Other) showed a moderate positive correlation with overall collaboration perceptions and a weaker, non-significant trend with satisfaction. State 5 (*Help-seeking/Co-construction/Disagreement with Justification*) was positively associated with group-level collaboration quality, suggesting that students spending more time in this state perceived their groups as more collaborative. State 2 (*SE-dominant*) was strongly associated with prior experience in math collaboration, indicating that students with richer backgrounds were more likely to engage in *Self-Explanation/Justification*. By contrast, State 3 (*Mixed*) consistently showed negative, non-significant correlations across outcomes. Overall, only States 4, 5, and 2 showed statistically significant associations with collaboration outcomes or prior experience.

Table 1
Spearman Correlations between Hidden-state Proportions (S1–S5) and Collaborative Measures

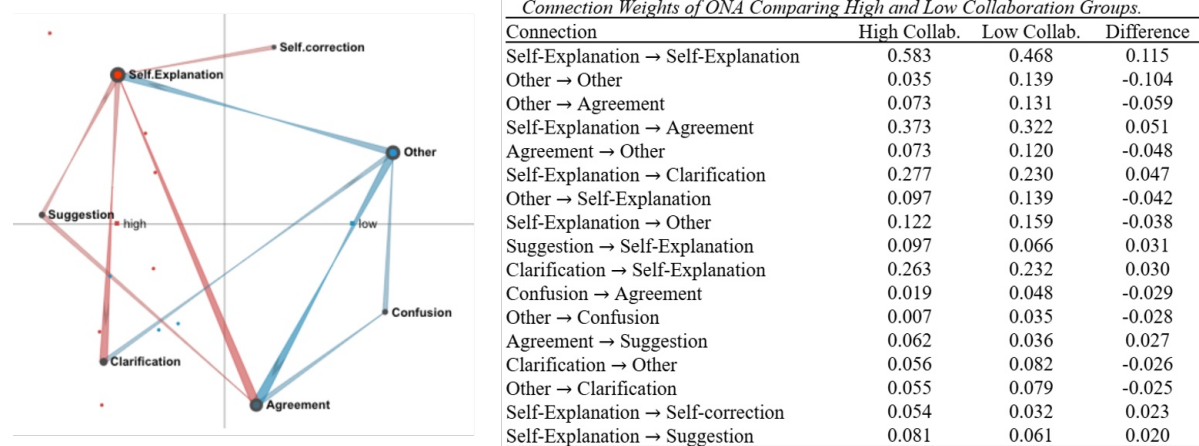
	S1	S2	S3	S4	S5
Collaborative learning					
Satisfaction	0.19	0.18	-0.21	0.4	0.14
Overall perceptions of collaboration	0.29	0.15	-0.38	0.43*	0.12
Overall problem-solving correctness	0.12	0.05	0.27	-0.03	-0.04
Group-level collaboration quality	0.18	0.27	-0.24	0.09	0.46*
Prior math learning collaboration experience	-0.04	0.53**	-0.36	-0.14	0.11
Prior collaborative skills	0.22	0.11	-0.17	-0.12	-0.2

Note. * $p < .05$, ** $p < .01$, *** $p < .001$.

In addressing RQ3, Figure 2 presents the ONA difference model comparing high-collaboration (red, $N = 6$) and low-collaboration (blue, $N = 5$) groups. A Mann–Whitney U test confirmed significant differences in overall dialogue patterns ($U = 27.00$, $p = .03$). High-collaboration groups engaged more actively in self-explanation (SE; $CW = 0.583$ vs. 0.468) and showed stronger bidirectional connections between explanation and clarification ($CW = 0.277$ and 0.263 vs. 0.232 and 0.230). These groups were also more likely to transition from explanation to agreement ($CW = 0.373$ vs. 0.322), error correction ($CW = 0.054$ vs. 0.032), and suggestion ($CW = 0.081$ vs. 0.061), forming sustained cycles of co-construction. In contrast, low-collaboration groups produced more “Other” utterances ($CW = 0.139$ vs. 0.039), reflecting filler talk, and more often surrounded explanations with such contributions. Their agreements were more frequently preceded by non-substantive utterances ($CW = 0.131$ vs. 0.073), suggesting shallower consensus. They also showed higher tendencies toward confusion ($CW =$

0.035 vs. 0.007) and clarification without deeper elaboration (CW = 0.079 vs. 0.055), further distinguishing their interaction patterns from high-collaboration groups.

Figure 2
ONA Difference Model of High (Red) vs. Low (Blue) Quality Groups



Discussion and implications

This study combined HMM, correlational analyses, and ONA to examine how dialogue patterns shape students' experiences and outcomes in CPS. Across methods, results highlight that productive collaboration is driven by multiple, interacting dialogue processes rather than a single dominant behavior.

The HMM revealed how structural patterns of dialogue acts distinguish collaboration quality by uncovering latent states and their transitions. For example, State 2 (SE-dominant) emerged as the dominant hub, occupying the largest share of time and mediating transitions across states. This state aligns with Mercer's category of *exploratory talk* (2000), which involves making reasoning explicit and ensuring mutual understanding. Prior work showed that exploratory talk dominates collaborative dialogue and sustains the longest dwell times (Earle-Randell et al., 2023). While exploratory talk more broadly includes justification, in this study, we found that questioning, elaboration, self-explanation, and clarification also function as its core markers. Similarly, the ONA results revealed self-explanation to be central to high-quality collaboration, aligning with prior work exploring mediators of cognitive challenge (Liang & Fung, 2020; Patterson, 2018).

The remaining states played complementary roles. State 1 captured agreement-dominant exchanges characteristic of *cumulative talk* (e.g., uncritical agreement, repetition, and passive acceptance), which may preserve social harmony but contribute little to deeper understanding (Hennessy et al., 2016; Mercer, 1995, 2000; Mercer & Wegerif, 1999). State 3 reflected mixed dialogue, State 4 focused on task grounding through reading and reframing, and State 5 combined help-seeking with co-construction and disagreement with justification. Correlational analyses further revealed that State 4 was positively associated with students' perceptions of collaboration, suggesting that grounding activities strengthen shared understanding, whereas State 5 correlated with collaboration quality, indicating that help-seeking and co-construction, especially when paired with justified disagreement, act to indicate sustained productive interaction. State 2 was linked to students' prior collaborative experience in mathematics, implying that explanatory and clarifying dialogue reflects transferable strategies from earlier contexts.

Unlike HMM, which captures temporal transitions among latent dialogue states, the ONA comparison highlighted structural differences between high- and low-quality groups. High-quality groups exhibited recurring cycles of self-explanation, clarification, agreement, and suggestion, consistent with Mercer's framework of exploratory talk, in which students justify reasoning, co-construct knowledge, and critically engage with peers' ideas. In contrast, low-quality groups relied more heavily on filler Other utterances and expressions of confusion, often surrounding explanations with non-substantive talk, aligning with Earle-Randell et al.'s (2023) finding that confusion can lead to disputational talk and conflict. These findings suggest that while *explanatory talk* (State 2) anchors collaboration, it is not sufficient alone. Task-oriented grounding (State 4) enhances shared perceptions of collaboration, and problem-focused exchanges involving help-seeking, co-construction, and justified disagreement (State 5) directly support collaboration quality. Overall, our results indicate that high-quality collaboration is best characterized not by a single dominant dialogue move but by the interplay of explanatory, grounding, and co-constructive processes.

References

- Andrews-Todd, J., & Forsyth, C. M. (2023). Assessment of collaborative problem solving skills. In *International Encyclopedia of Education (Fourth Edition)* (pp. 494–503). Elsevier. <https://doi.org/10.1016/B978-0-12-818630-5.10062-4>
- Boyer, K. E., Phillips, R., Ingram, A., Ha, E. Y., Wallis, M., Vouk, M., & Lester, J. (2011). Investigating the relationship between dialogue structure and tutoring effectiveness: a hidden Markov modeling approach. *International Journal of Artificial Intelligence in Education*, 21(1-2), 65-81.
- Earle-Randell, T. V., Wiggins, J. B., Ruiz, J. M., Celepkolu, M., Boyer, K. E., Lynch, C. F., Israel, M., & Wiebe, E. (2023). Confusion, Conflict, Consensus: Modeling Dialogue Processes During Collaborative Learning with Hidden Markov Models. In N. Wang, G. Rebolledo-Mendez, N. Matsuda, O. C. Santos, & V. Dimitrova (Eds.), *Artificial Intelligence in Education* (Vol. 13916, pp. 615–626). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-36272-9_50
- Graesser, A. C., Greiff, S., Stadler, M., & Shubeck, K. T. (2020). Collaboration in the 21st century: The theory, assessment, and teaching of collaborative problem solving. *Computers in Human Behavior*, 104, 106134. <https://doi.org/10.1016/j.chb.2019.09.010>
- Hansen, E. K. S. (2022). Students' agency, creative reasoning, and collaboration in mathematical problem solving. *Mathematics Education Research Journal*, 34(4), 813–834. <https://doi.org/10.1007/s13394-021-00365-y>
- 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. <https://doi.org/10.1016/j.lcsi.2015.12.001>
- Liang, W., & Fung, D. (2020). Development and evaluation of a WebQuest-based teaching programme: Students' use of exploratory talk to exercise critical thinking. *International Journal of Educational Research*, 104, 101652. <https://doi.org/10.1016/j.ijer.2020.101652>
- Marquart, C. L., Hinojosa, C., Swiecki, Z., Eagan, B., & Shaffer, D. W. (2021). Epistemic network analysis (Version 1.7.0) [Software]. Available from <http://app.epistemicnetwork.org>
- Mercer, N. (1995). *The Guided Construction of Knowledge: Talk Amongst Teachers and Learners*. Multilingual Matters.
- Mercer, N. (2000). *Words and Minds*. Routledge. <https://doi.org/10.4324/9780203464984>
- Mercer, N., & Wegerif, R. (1999). Is 'exploratory talk' productive talk? In *Learning with Computers*. Routledge.
- Patterson, E. W. (2018). Exploratory talk in the early years: Analysing exploratory talk in collaborative group activities involving younger learners. *Education* 3-13, 46(3), 264–276. <https://doi.org/10.1080/03004279.2016.1243141>
- Partnership for 21st Century Skills (2007). *Partnership for 21st century skills - Core content integration*. https://www.marietta.edu/sites/default/files/documents/21st_century_skills_standards_book_2
- Tan, Y., Ruis, A. R., Marquart, C., Cai, Z., Knowles, M. A., & Shaffer, D. W. (2023). Ordered Network Analysis. In C. Damşa & A. Barany (Eds.), *Advances in Quantitative Ethnography* (Vol. 1785, pp. 101–116). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-31726-2_8
- Tausczik, Y. R., Kittur, A., & Kraut, R. E. (2014). Collaborative problem solving: A study of MathOverflow. *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 355–367. <https://doi.org/10.1145/2531602.2531690>
- Thomas, A. P., Roger, D., & Bull, P. (1983). A sequential analysis of informal dyadic conversation using Markov chains. *British Journal of Social Psychology*, 22(3), 177–188. <https://doi.org/10.1111/j.2044-8309.1983.tb00582.x>
- Zhang, S., Earle-Randell, T. V., Shen, Q., Botelho, A. F., Israel, M., Boyer, K. E., ... & WIEBE, E. (2024, November). Predicting and Analyzing Students' Higher-Order Questions in Collaborative Problem-Solving. In *International Conference on Computers in Education*. APSCE.

Acknowledgements

We would like to thank the National Science Foundation (#2331379), Institute of Education Sciences (#R305B230007), Gates Foundation (#078981), Support from the Learning Engineering Tools Competition, and other anonymous philanthropy.