Shih, JL. et al. (Eds.) (2023). Proceedings of the 31st International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

Epistemic Network Analysis to assess collaborative engagement in Knowledge Building discourse

Aloysius Ong[a*], Chew Lee Teo[a], Alwyn Vwen Yen Lee[a] & Guangji Yuan[a]
[a]*National Institute of Education, Nanyang Technological University, Singapore*
*aloysius.ong@nie.edu.sg

Abstract: Knowledge Building (KB) is an established learning sciences theory that seeks to promote innovative ideas and idea improvement among students via collaborative engagement in productive discourse. KB discourse supports students to make constructive discourse moves such as questioning, explaining with evidence, adding new information and so on, to advance the collective inquiry. However, current understanding on KB discourse remains limited to students' online participation. Although small group discussion is a common practice, there is little understanding on the role of verbal discussions to support KB discourse. This paper attempts to address this line of inquiry by assessing student engagement in KB discourse supported by both online and verbal discussions. Data is retrieved from a group of six students in a Grade 6 Social Studies class. The group participated in a 2.5hr lesson designed with opportunities for discussions on the Knowledge Forum (online) and in small groups (verbal). Group talk was transcribed, and Knowledge Forum notes were coded for its semantic level of contribution, with the codes being analysed for weighted connections using Epistemic Network Analysis (ENA). The ENA analysis revealed clear differences in both group and individual engagement between the online and verbal discourse. Notably, students' contributions on Knowledge Forum showed an apparent pattern of stronger connections among codes of higher semantic levels, suggesting that students were more cognitively engaged in the online discussion than their group verbal talk. Implications for research and practice are discussed.

Keywords: Knowledge Building, Discourse Analysis, Epistemic Network Analysis

1. Introduction

Knowledge Building (KB) is an established learning sciences theory that seeks to promote innovative ideas and idea improvement among students via collaborative engagement in productive discourse. Knowledge Building brings students together to work as a community to tackle real-world problems and creates a discursive environment that supports students to develop a deeper collective understanding of the problem at hand (Scardamalia & Bereiter, 2006). When a community of students is engaged in Knowledge Building discourse to advance collective understanding, they contribute important moves such as questioning, explaining with evidence, adding new information and so on, in order to advance the collective inquiry. The understanding of what and how students contribute to improve ideas and their engagement in KB discourse have been of interest among researchers. In fact, to understand student engagement in KB discourse, various frameworks have been proposed to characterise student discourse moves such as theory building moves (Chuy et al., 2010) and idea complexity (Zhang et al., 2007). Recent works suggest quantitative modelling techniques such as Epistemic Network Analysis (ENA) to assess connections between student moves (Ma et al., 2019). Insights from such analyses provide a means to understand the quality of student ideas from discussions as well as how they develop collective inquiry (build on ideas). However, current understanding on KB discourse remains limited to students' online participation. Although small group discussion is a common practice, there is little understanding on the role of verbal discussions to support KB discourse. This paper attempts

to address this line of inquiry by assessing student engagement in KB discourse supported by both online and verbal discussions. Specifically, we aim to understand the similarities and differences in the students' discourse moves and connections between their ideas as they participate in online and verbal discussions in KB lessons.

2. Literature

2.1 Knowledge Building discourse

Knowledge building discourse is characterised by the generation of novel and challenging student ideas, questions, and different perspectives that strive to improve the collective understanding of the problem at hand. It focuses on the generative and creative nature of a classroom discussion, moving away from a more typical and traditional classroom discussion that focuses on getting the single correct answer (Ong et al., 2021; Teo et al., 2022). The goal of Knowledge Building classrooms is beyond information sharing and divide and conquer to put a product together (Scardamalia & Bereiter, 2006). The essence of knowledge building classroom lies in the Knowledge Building discourse. In KB discourse, students are encouraged to constantly question and refine their ideas collectively, find their motivation to connect with the big problem and help each other progress in the inquiry and knowledge. KB discourse is usually supported by the Knowledge Forum which is an online platform to facilitate students in posting and building on ideas (Scardamalia & Bereiter, 2006). Figure 1 shows a view of the Knowledge forum where students interact and contribute to the discussion together. Each node (square box) represents a note (contains ideas) from a student, and the arrows show the build-on from one node to another. However, it is also common for students to work in small groups and discuss verbally to improve ideas. Figure 1 shows a group of four students in a face-to-face setting with computer devices to allow students to engage in both online and verbal discussions.

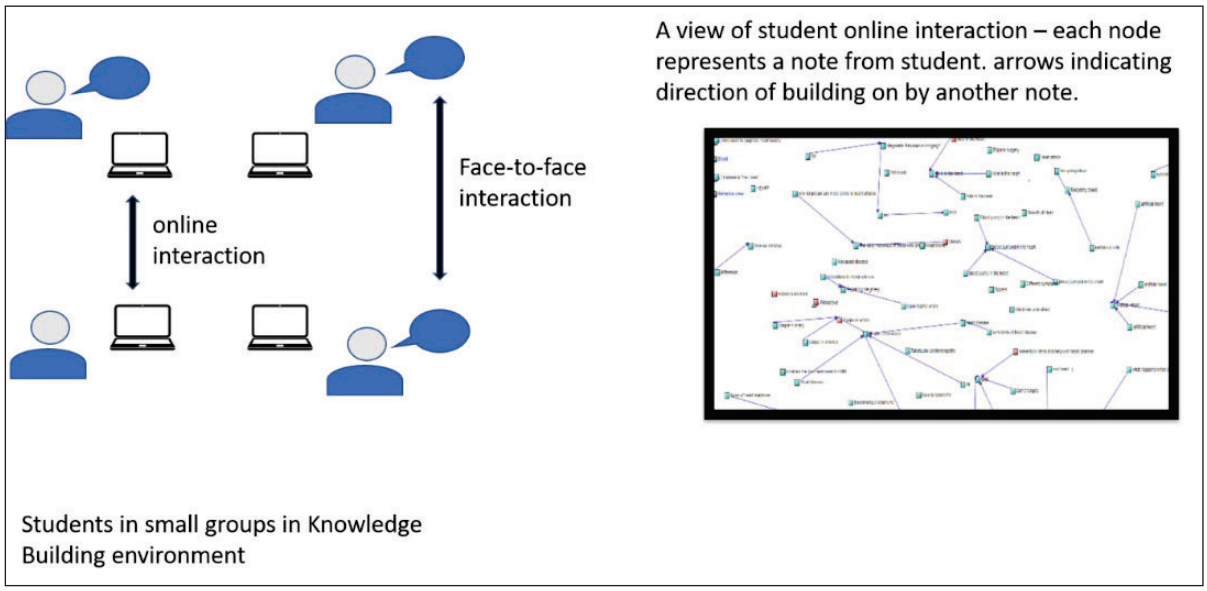

Figure 1. Face-to-Face and Online discussion modes in Knowledge Building

2.2 Understanding student engagement in KB discourse

Several ways of understanding student contributions in KB discourse have been proposed in the field. One notable area of focus is on the characterisation of students' ideas or discourse moves (e.g., Chuy et al., 2010; Zhang et al., 2007; 2022). Notably, Chuy and colleagues proposed the theory building moves to code student discourse moves. Specifically, they developed an empirically grounded list of six major ways of contributing to a productive dialogue for KB discourse. These categories comprise: asking thought-provoking questions,

theorizing, experimenting, working with evidence, creating syntheses and analogies, and supporting discussion. From another study, Zhang and colleagues (2007) looked at elementary students' (Grade 4) contributions in KB discourse from the quality of their individual ideas. They devised the idea complexity scheme to categorise students' ideas. Questions may be categorised into explanation-seeking versus fact-seeking questions and single-area versus cross-area questions. According to Zhang et al. (2007), students' contributions can also be assessed based on the level of scientific sophistication (pre-scientific; hybrid; basically scientific; and scientific) and the level of epistemic complexity (unelaborated facts; elaborated facts; unelaborated explanations; elaborated explanations).

Another area of focus is on the epistemic understanding via patterns of how students build on each other's ideas to progress knowledge (e.g., Lin & Chan, 2018; Ong et al., 2021). For instance, Lin and Chan introduced the thread analysis to code the quality of student online discourse. The authors found and categorised different patterns of how students build on each other's ideas into four epistemic levels of increasing collective knowledge advancement. The basic level, Fragmented Discussion, showed no connection or a weak connection between student ideas which do not contribute to knowledge advancement. The next level, Knowledge Sharing, showed connections between student ideas mainly for exchanging information and may not be essential to solving the problem. Knowledge Construction represented a higher level where there is a building of connections to generate questions and explanations to construct theories. Finally, Knowledge Building represents the highest epistemic level where there are coherent connections between ideas to bring about new understanding and progressive development of students' theories. Lin and Chan's work highlighted the importance and need to consider idea connection when assessing student engagement in KB discourse, however their coding scheme remains restricted by few epistemic levels.

Recent works suggest the potential of modelling techniques to understand connections between student discourse moves in Knowledge Building. For instance, Ma and colleagues (2019) explored the Epistemic Network Analysis (ENA) to assess Knowledge Building discourse, specifically looking into "ways of contributing" to the discourse such as questions and theories. The authors argued that there is potential for ENA to model contribution dynamics during cycles of knowledge advancement. Ma et al.'s study demonstrated the potential of ENA as a feasible technique to identify similarities and differences in discourse patterns between groups. For example, one group may differ from another in terms of certain connections such as theorizing and questioning moves. As explained by Ma and colleagues, such network models can be used to understand complementary engagement in group discussions and indications of collaborative engagement in both online and verbal discourse.

While there is an extensive development of techniques to assess student contributions in KB discourse, current understanding on KB discourse remains limited to students' online participation. Existing studies tend to focus on student participation in online discussions supported by the Knowledge Forum. There is little understanding on the role of verbal discussions to support KB discourse, even though it is a common practice among teachers to plan for small group discussions to allow students to talk about their ideas verbally (see Ong et al., 2021). Elsewhere beyond KB contexts, there are studies showing that verbal discussions do not necessarily compliments nor supports online discourse. For example, a recent study by Bagheri and Zenouzagh (2021) highlighted that students engaged differently in face-to-face and online conversations. The researchers investigated the discourse from a group of 30 student participants and they found differences in verbal, paralinguistic and functional level between the students' face-to-face and online conversations (computer-mediated communication). Their study revealed that student engagement in online discourse was more productive compared to the physical discourse as (i) the students showed active control of the talk and engaged in richer exchanges of ideas such as argumentation and (ii) students have more tendency to appeal to the teacher for help and displayed silences and pauses in the physical discourse (Bagheri & Zenouzagh, 2021). As student engagement in both online and verbal discussion may impact on the quality of their collective inquiry in KB, it is thus important for ongoing research to delve deeper on this understanding.

In this paper, we hope to build on this line of inquiry by assessing students' ways of contributing in both online and verbal discussion to better understand their engagement during

Knowledge Building discourse. We report on the interactional patterns in verbal and online discussions from a group of students engaged in KB collective inquiry over a 2.5hr lesson. Specifically, we aim to understand the similarities and differences in the students' discourse moves and connections between their ideas as they participate in online and verbal discussions in KB lessons by using Epistemic Network Analysis. Our research question therefore is "What connections of student contribution can we see during verbal and online discussion using ENA?"

3. Research design and Analysis

3.1 Methods

To capture students' verbal interaction during face-to-face discourse, we used a 360-degree camera placed at the center of the group to video-record the group interaction as shown in Figure 4. We flattened the video data to enable us to see the interaction of each participant in the group. To ensure high quality of verbal data is captured, we used another audio recorder with lapel microphones for every individual student to capture their verbal data (Figure 2). This setup was part a multimodal recording setup that also collected students' self-reports on their emotions and physiological data from Empatica E4 wearables. However, we focused on the description of the setup on the video and audio recording.

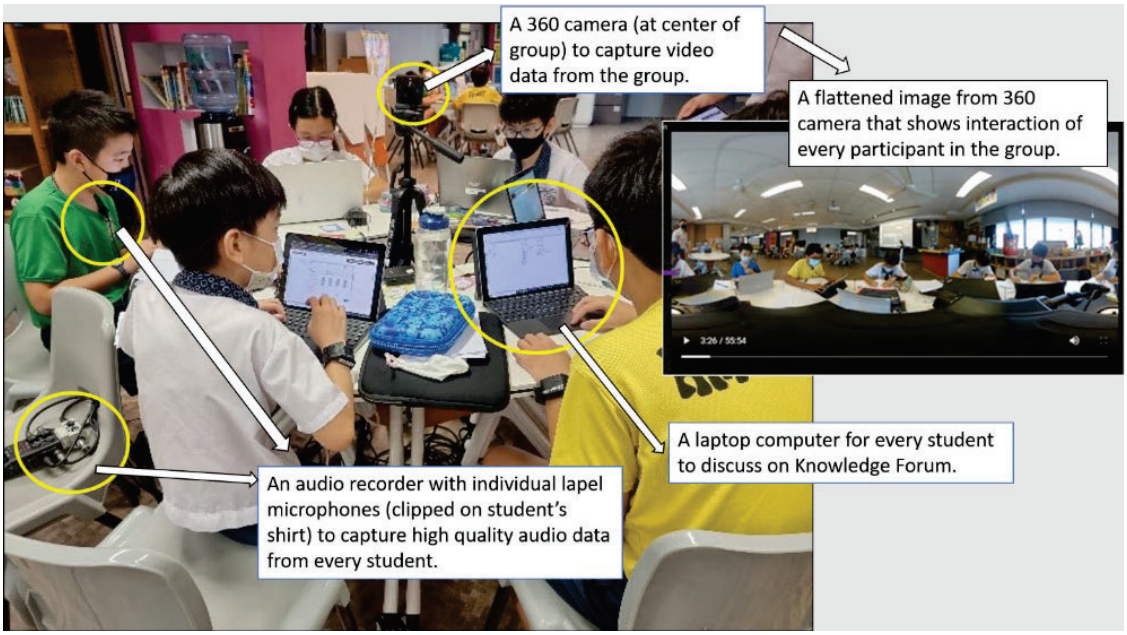

Figure 2. Recording setup for f2f and online discussion in classroom

3.2 Lesson design and implementation

In this case study, we collaborated with a teacher (experienced in knowledge building) to implement a lesson designed with Knowledge Building activities to support collective inquiry and idea improvement. Table 1 shows the teacher's planned lesson activities that engaged students to conduct idea generation and idea building about a real-world problem (based on Social Studies). Students will be tasked to play the roles of locals situated in three Southeast Asian areas (Sarawak Rainforest, Mekong River, and Anak Krakatoa) and they will continue to discuss the pros and cons of relocation and to provide their final stand with justifications.

Before the lesson, the students were pre-assigned to one of the three geographical regions. Each group was tasked to discuss and find solutions to existing problems (habitat loss, active volcano, and deforestation) that threaten the respective residents' livelihoods and environments. Table 1 shows the flow of the activities that were planned to provide

opportunities for student discussion both online and face-to-face. A total of 17 Primary Five (Grade Five) students participated in this lesson that lasted for 2.5 hours.

Table 1. *Activities from the teacher's lesson plan*

Lesson phase	Teacher's instructions and class activities	Students' KF activities
Activating Activity	● Students played the roles of locals in three Southeast Asian areas (Sarawak Rainforest, Mekong River, and Anak Krakatoa). ● Students introduced to the KB task, which is a Town Hall discussion anchored on the question: "Should our community stay or relocate from our area?"	No KF activities recorded
Idea Generation	● Students discussed the anchor question in their respective groups. ● Students posted notes on KF stating their stand and with supporting reasons.	Discuss and post different perspective (stand) with reasons on KF
Idea Improvement – Build On	● Students discussed and built on their peers' ideas on the Knowledge Forum using build-on scaffolds. ● The teacher used learning analytics such as word clouds and an in-built Scaffold Tracker to conduct a meta-discussion and highlight value-adding contributions.	Read and build on groupmates' viewpoints, such as clarifying, inquiring, or proposing new elaborations.
Idea Improvement – Rise Above	● Teacher introduced a Rise-Above scenario: Students took on newer roles as community representatives to potentially implement a solution to the respective area's issues.	Negotiate as a group, weighing pros and cons of different solutions to develop a solution.
Idea Assessment	● Students consolidated their problems and solutions that were then shared with the class for comments and suggestions.	Post synthesis of group ideas and peer feedback for improvement.
Reflection	● Students responded to three reflection questions on KF.	Post reflection on KF

3.3 Analysis

In this study, we used the ENA Web Tool (Shaffer et al., 2009; Shaffer et al., 2016) to analyse both Knowledge Forum data and verbal discussion data. ENA models the weighted structure of connections in discursive data (Shaffer & Ruis, 2017). The method was designed to make sense of connections between ideas and actions using a repertoire of network-based methods that include visualizations and statistical modelling. For example, if a student group is discussing a particular topic, they share important elements such as the areas of their interest, their questions, information that they know, and so on. They may have a few conversations during the lesson, and a key part of understanding their discourse is modelling how they think about the connections between the different elements of their inquiry. ENA quantifies the co-occurrence of, for example, codes in discourse or elements of interaction in a chat and produces a weighted network visualization that shows connections for each unit of analysis (e.g., individual speakers or sub-groups of speakers) (Shaffer & Ruis, 2017). This paper reports the preliminary analysis with one student group (six students).

Based on 42 Knowledge Forum notes from the group, we coded for student contribution based on five levels of semantic patterns (adapted Zhang et al., 2022). These included casual talk (CT), fact-seeking question (Q1), explanation-seeking question (Q2), explanation with simple points or simple statement without elaboration (E1), and explanation with elaborations or rise-aboves (E2). Table 2 provides the description and sample coding. We use each student idea as the unit of analysis. This is typically in the form of a question or an explanation (statements). Where a student note or utterance contains a question and explanation, we coded for both respectively. We coded student contribution from their verbal talk as well. In total, we captured approximately 1.5hr of talk time from the student group and this data was transcribed (2416 utterances) and coded with the same coding scheme and unit of analysis. In total, we coded 51 codes from the KF notes and 2548 codes from the verbal data.

We next applied Epistemic Network Analysis (ENA) to identify and quantify connections between these codes for the group. We created the epistemic networks by firstly using the discourse type (KF and verbal) as *units* for ENA to construct networks. The two networks were aggregated using a binary summation to reflect the presence or absence of the co-occurrence of each pair of codes within a stanza window of 6 lines in the *conversations*. The *conversations* are collections of lines that ENA uses to model connections between concepts, and this was based on the student utterances or KF notes in our analysis. Finally, codes are considered concepts and its patterns of association are what we want to model. In our case, we applied all the five semantic levels of the student contribution.

Table 2. *Description and examples of semantic patterns of KF notes*

Codes of semantic patterns	Description/Examples
Casual Talk (CT)	Information or inputs generally irrelevant to the collective inquiry, e.g., “I need the toilet.”
Fact-seeking question (Q1)	Questions that elicit factual information, using when, where, or who, e.g., “Where do you want to go?”
Explanation-seeking question (Q2)	Questions which inquire into relations between facts and elicit elaborations, such as why, how, and what, e.g., “Why should we be affected by other people living in the modern world?”
Explanation with simple points/Simple statement without elaboration (E1)	Simple statement without elaboration, e.g., “You are contradicting yourself. These are the exact reasons why we should move.”
Explanation with elaborations/rise-aboves (E2)	Statement with elaborations to provide reasons, relationships, or synthesis of ideas, e.g., “We don't really have a government. Without a government, there is no law, no order. And with a government, we suffer under a dictatorship. People who are weaker will abuse power.”

4. Findings

Answering the research question “What connections of student contribution can we see during face-to-face and online discussion using ENA?”, Figure 3 shows the epistemic network between the group’s verbal talk and KF discussion throughout the entire lesson. The positions of the network graph nodes are fixed, and the nodes correspond to the codes applied in the analysis. ENA determined these positions using an optimization routine that minimises the difference between the plotted points and their corresponding network centroids (Shaffer et al., 2016). The lines reflect the relative frequency of co-occurrence, or connection, between

two codes. The weight of each code is represented by the size of the node in the network graph. The results show that the group demonstrated a relatively unique profile between the two discourse types. Student contributions from their KF discussion shows centrality on elaborated questions (node with highest weight) and strongest connections between this code with explanations (simple and elaborated). Student contributions from their verbal discussion however shows centrality on simple explanation and strongest connections with casual talk (thickest line) and questions (simple and elaborated) than elaborated explanations. This preliminary analysis from this group suggest that students seem to engage in a deeper discussion of their ideas when they interact on KF compared to their verbal interaction (despite that they were given time to discuss verbally).

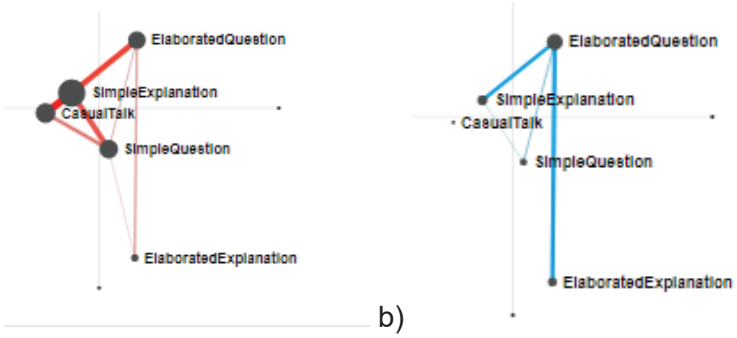

Figure 3. Epistemic networks showing connection of codes (semantic levels) for (a) verbal and (b) online discussion from the student group.

This finding was corroborated when we looked at the distribution of codes for the verbal and online discussion for the group. As shown in Figure 4, there was also a considerable difference in the distribution of codes between the two discourse modes. Student contributions in their KF discussion showed highest occurrence of simple explanation followed by elaborated explanations and then elaborated questions and lower occurrences of casual talk and simple questions. Student contributions in verbal discussion also showed highest occurrences of simple explanations but followed by casual talk and lower occurrences of the codes of higher semantic levels (elaborated explanations and elaborated questions).

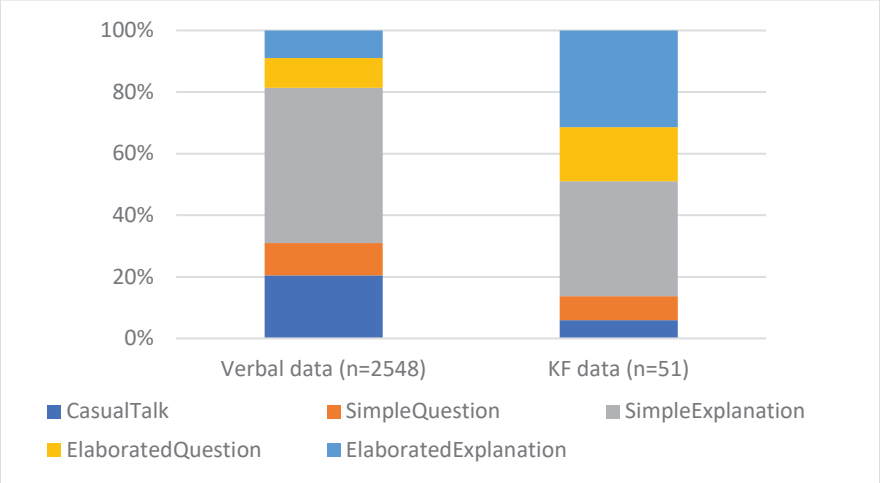

Figure 4. Comparison of codes (semantic levels) between verbal and online discussion from the student group.

However, we found interesting observations when we further explore the epistemic network of individual students. As shown in Figure 5, 3 students in the group showed connections between the codes for their online discussion for the group but their network profile were not totally consistent. Specifically, two students, S1 and S5, showed strongest

connections between elaborated question and elaborated explanation (thickest line), while student S4 showed strongest connections between elaborated question and simple explanation and question. This result suggests that S1 and S5 were consistently using higher semantic levels like elaborated questions and explanations to deepen the collective inquiry. Note that the ENA did not picked up connections for the other students (as these students contributed very few notes).

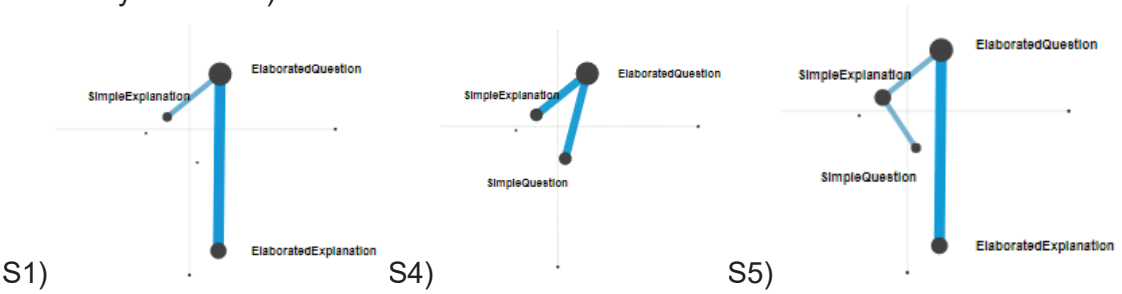

Figure 5. Epistemic networks showing connection of codes (semantic levels) for online discussion from individual students in the group.

Similarly, their individual network profile was also different in the verbal discussion. Figure 6 shows connections between the codes for their verbal discussion. This time, students S1 and S4 showed strongest connections between elaborated question and simple explanation, whereas students S1 and S5 displayed strongest connections between elaborated question and elaborated explanation (thicker line). The other three students S2, S3, and S6 showed weaker connections between these codes but strongest connection between simple explanation and casual talk. Thus, students S1, S4 and S5 seemed more engaged in a deeper discussion of their ideas during their verbal discussion, which is a consistent find with their contributions on KF discussion.

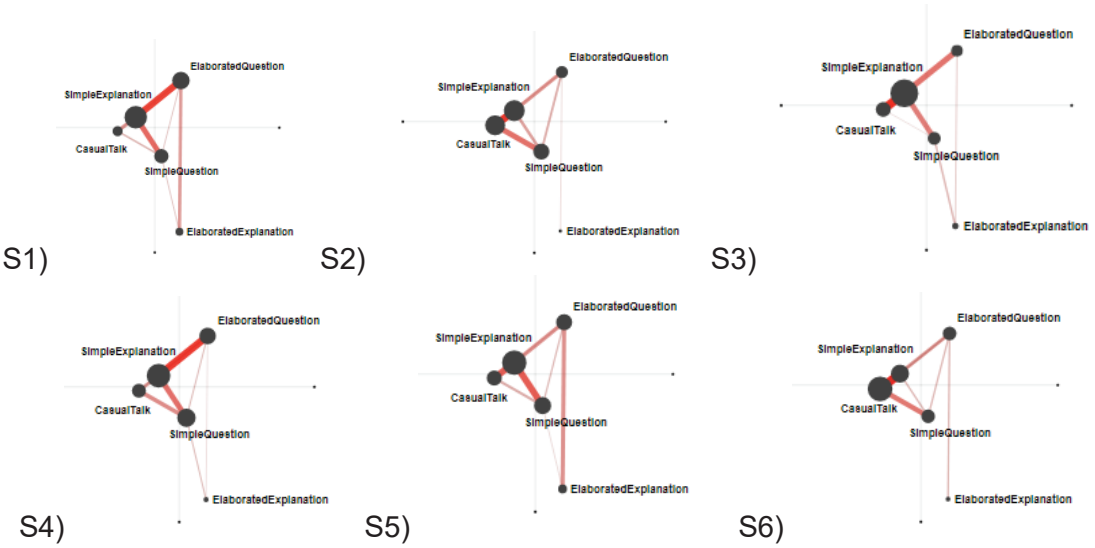

Figure 6. Epistemic networks showing connection of codes (semantic levels) for verbal discussion from individual students in the group.

5. Discussion and conclusion

Our preliminary finding from the ENA analysis is consistent with those by Ma et al. (2019) as we found that students deepened their collective inquiry as they engaged in Knowledge Building discourse. As shown in our analysis, this group of students reflected a strong connection between higher semantic levels of contributions elaborated questions and elaborated explanations when given time to discuss on KF. However, we also found different patterns of their contributions during verbal discussion with the ENA analysis. Specifically, in their group verbal talk, there was a strong presence of casual talk and connections between

simple explanations and questions which suggest that the collective inquiry was more superficial during verbal discussions. This claim that students engage more productively in online communication is consistent with the report from Bagheri and Zenouzagh (2021). However, we think that such a result may also be due to the affordances of the online KF platform such as the support of KB scaffolds and the chance to read (even multiple times) and consider multiple peer ideas. Unlike KF, students tend to observe turn-taking during verbal discussion, which may restrict them in responding to certain ideas and giving longer responses such as elaborated explanations. Therefore, future work can consider looking into the impact of different support features in both online and verbal discussion.

An interesting finding was the interactional patterns from individual students. The ENA analyses suggested that students who were engaged in a deeper discussion of their ideas on KF also contributed ideas of higher semantic levels during their verbal interaction. This finding has an implication for practitioners particularly on the lesson design and facilitation. For example, teachers can first plan for an online discussion to allow students to generate questions or explanations of higher semantic levels and then allowing them to continue building their ideas via small group discussions. As it is easy to extract student notes from the KF, teachers can also run the data using the ENA tool and use the analysis to discuss with students on how their ideas are connected and ways to improve the quality of their talk. Such teacher facilitation has been shown to be beneficial for students to develop new inquiry (Zhu & Lin, 2023). However, we recognise that our findings were based on a very small dataset and our claims warrant further investigation with more data.

Moving forward, we have collected new datasets, and additional analyses are underway to compare findings from other groups and to validate the claims. In conclusion, this study highlights that the understanding of student engagement in KB discourse is not straightforward, as student interactions can occur within and between groups as well as via online or verbal discussion. As we found that the type of and connections between student discourse moves were less sophisticated in their verbal discussions, a key implication from this study is that students need more support during verbal talk to engage in deeper collective inquiry and advancing the community knowledge. Our study also shows ENA as a promising approach to assess Knowledge Building discourse in both online and verbal mode. ENA affords modelling of both group and individual networks to enable the understanding of differences in student engagement during KB discourse. Future studies can explore on the automation of such feedback for teachers and students. For example, the real-time capturing and coding of verbal data which is integrated with the ENA tool to generate feedback on the discourse quality from the verbal discussion.

Acknowledgements

This study was funded by the Ministry of Education (MOE), Singapore, under the Education Research Funding Programme (OER 19/19 TCL) and administered by National Institute of Education (NIE), Nanyang Technological University, Singapore. The views expressed in this study are the authors' and do not necessarily represent the views of the host institution. The authors thank the teacher and student participants involved in this study. The authors also express their appreciation to Associate Professor Justin Dauwels and his team for expert advice and support on the multimodal recording setup.

References

Bagheri, M., & Mohamadi Zenouzagh, Z. (2021). Comparative study of the effect of face-to-face and computer mediated conversation modalities on student engagement: speaking skill in focus. *Asian-Pacific Journal of Second and Foreign Language Education*, *6*(1), 1-23

Chuy, M., Resendes, M., & Scardamalia, M. (2010, August). Ways of contributing to knowledge building dialogue in science. Paper presented at the Knowledge Building Summer Institute, Toronto, Canada.

Lin, F., & Chan, C. K. (2018). Promoting elementary students' epistemology of science through computer-supported knowledge-building discourse and epistemic reflection. *International Journal of Science Education*, *40*(6), 668-687.

Ong, A., Lee Teo, C., Tan, S., & Song Kim, M. (2021). A knowledge building approach to primary science collaborative inquiry supported by learning analytics. *Education 3-13*, *49*(3), 371-386.

Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. In K. Sawyer (Ed.), *Cambridge Handbook of the Learning Sciences* (pp. 97-118). New York: Cambridge University Press.

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.

Shaffer, D. W., Hatfield, D., Svarovsky, G. N., Nash, P., Nulty, A., Bagley, E., ... & Mislevy, R. (2009). Epistemic network analysis: A prototype for 21st-century assessment of learning. *International Journal of Learning and Media*, *1*(2), 33-53.

Shaffer, D., & Ruis, A. (2017). Epistemic network analysis: A worked example of theory-based learning analytics. *Handbook of learning analytics*.

Teo, C. L., Ong, A., & Lee, V. Y. A. (2022, June). Exploring students' epistemic emotions in knowledge building using multimodal data. In Weinberger, A., Chen, W., Hernandez-Leo, D., and Chen, B. (Eds.), *Proceedings of the 15th Computer-Supported Collaborative Learning (CSCL) 2022* (pp. 266-273). Hiroshima, Japan: International Society of the Learning Sciences. https://2022.isls.org/proceedings/

Zhang, J., Scardamalia, M., Lamon, M., Messina, R., & Reeve, R. (2007). Socio-cognitive dynamics of knowledge building in the work of 9-and 10-year-olds. *Educational Technology Research and Development*, *55*, 117-145.

Zhang, J., Tian, Y., Yuan, G., & Tao, D. (2022). Epistemic agency for costructuring expansive knowledge-building practices. *Science Education*.

Zhu, G., & Lin, F. (2023). Teachers scaffold student discourse and emotions in knowledge building classrooms. *Interactive Learning Environments*, 1-18.

Ma, L., Matsuzawa, Y., & Scardamalia, M. (2019). Using epistemic network analysis to ExploreWays of contributing to knowledge building discourse. In *Proceedings of the 13th international conference on computer-supported collaborative learning* (pp. 881–882). Lyon, France: International Society of the Learning Sciences.