“I Thought We Said”: Perceived Peer Support, Discourse Cohesion, and Regulation in Engineering Design

Ha Nguyen, University of California, Irvine, thicn@uci.edu
Kyu Yon Lim, Ewha Woman’s University, klim@ewha.ac.kr
Lily Wu, University of California, lwu@uci.edu
Christian Fischer, Eberhard Karls University of Tübingen, christian.fischer@uni-tuebingen.de
Mark Warschauer, University of California, Irvine, markw@uci.edu

Abstract: There is limited research on how perceived peer network influences student collaboration in project-based instruction. Research based on interviews or content analyses may overlook the semantic structure of discourse. In this study, we combine content analysis with computational linguistics to explore the collaboration patterns of 22 first-year students during face-to-face group design (n = 7,514 conversational turns) in a project-based engineering course. We find that students who reported smaller peer network generally produced discourse that provided new information, but less cohesion, compared to students with perceived median and large peer networks. Overall, students with small networks also engaged in group planning, evaluation, and shared regulation less frequently compared to the other two groups. Findings have implications for adjusting group arrangement in design activities. The study also illustrates the potential of incorporating computational approaches to detecting discourse.

Introduction

The strength of students’ existing social ties influences how they co-construct knowledge and seek resources in collaborative learning (Dawson, 2008). Students who possess more close-knit ties tend to engage more and produce more cohesive communication, with some evidence of better academic outcomes (Putnik et al., 2016). Understanding how students with varied perceptions of social ties regulate motivation and behaviors in design groups is important in engineering education, where some student populations (e.g., female, underrepresented minorities) have reported experiencing a lack of peer support (Hurtado, Newman, Tran, & Chang, 2010).

However, there has been limited research on regulation processes, particularly the shared regulation of goals and progress towards task completion (Malmberg, Järvelä, & Järvenoja, 2017). Content analyses of discourse provide fine-grained details about learning, but may not capture the sequential nature and semantic structure within situated interactions (Strijbos, Martens, Prins, & Jochems, 2006). This study extends the research base by combining content analyses with computational linguistics to examine the interconnection of regulation patterns in discussion of engineering design. **RQ1.** To what extent does discourse cohesion vary among students whose expected peer support differs? **RQ2.** To what extent do regulation processes vary among student groups?

Regulation in collaborative engineering design

Project-based curricula foster student *self-regulation* of cognition, motivation, and behavior towards task completion (Schmidt, Vermeulen, & Van Der Molen, 2006). In collaborative learning, from a social cognitive view, students also practice *co-regulated learning* to influence other team member’s goals and beliefs (i.e., “you” perspective) and *shared regulation* to plan and monitor progress towards collective goals (i.e., “we” perspective; Järvelä & Hadwin, 2013; Malmberg, Järvelä, & Järvenoja, 2017). These patterns are critical in engineering design, as there exists a positive relation between the amount of time that learners spend on planning and evaluating design alternatives and the resultant quality of their designs (Atman, Cardella, Turns, & Adams, 2005). However, examination of collaborative discourse in science and engineering suggests that opportunities for knowledge co-construction are not equally distributed within groups (Cohen & Lotan, 2014). Notably, few studies have examined how regulation types (e.g., self, shared regulation) and processes unfold in tandem (Malmberg et al., 2017). Existing frameworks of regulation processes cycle among task understanding, goal setting, implementation, and evaluation (Malmberg et al., 2017; Sinha, Rogat, Adams-Wiggins, & Hmelo-Silver, 2015). Less effective planning, followed by limited monitoring, hinders group problem-solving (Sinha et al., 2015). Thus, understanding how groups alternate between regulatory phases may help instructors adjust tasks when needed.

Techniques to analyze discourse data

There has been growing effort to apply computational approaches to reduce the processing time of large discourse datasets. Manual coding is time-consuming, resource-intensive, and may not attend to the sequence of discourse (Strijbo et al., 2006). Meanwhile, automated processes such as social network analyses only capture surface-level interactions (Dowell, Nixon, & Graesser, 2019). Thus, scholars have combined analytics strategies to examine learner roles and regulation in collaborative discourse. For example, Dowell et al. (2018) combine computational linguistics and sequential analyses to detect roles in regulation, social coordination, and meaning-making in discussion. Shaffer and colleagues (2017) develop nCoder, using regular expressions to automatically classify qualitative data. The classifications are incorporated into epistemic network analysis (ENA) to display the structure of learning domains over time (Shaffer, 2017). These examples illustrate the potential to apply computational approaches to examine interaction data, although questions of validity and reliability between human-annotated and automated output remain (Dowell et al., 2019; Lee, Gui, Manquen, & Hamilton, 2019).

Methods

This study took place in a two-term first-year engineering course in the Southwestern U.S. The course consisted of 1-hour lectures and 2-hour lab sections over ten weeks for students to design, build, and test an autonomous device. The first term introduced students to fundamental engineering principles. In the second term, students formed teams of five to six to work on design projects. Prior to team formation, we surveyed students about their peer network in the class (n = 211). We asked students to identify the peers who they would turn to for resources and help with tasks, and the weekly frequency that they leveraged peer resources. Peer network ties were calculated as (*number of peers)*(frequency of interaction*). On average, students reported having 4.6 ties.

Analyses draw from discussion transcripts of four teams (22 students) during the last three labs in the second term (24 hours of audio). The selected teams presented a range of perceived network: *Below median* (n = 7, M = 2.33, SD = 1.97), *median* (n = 10, M = 5.00, SD = 0.00), and *above median* (n = 5, M = 7.00, SD = 1.00). The teams had a mix of students from each group, with about half in the *median*. The sample represents overall course demographics (22.73% female, 72.72% underrepresented minorities). The three peer network groups had similar compositions of female to male ratio and percentage of underrepresented minority students.

Table 1. Codes for regulation types (*italicized*) and processes; Malmberg et al. (2017); Jarvela & Hadwin (2016)

Code	Definition	Example
Task Understanding	Activate prior knowledge. Think about tasks' purpose and value. Discuss instruction.	**That's the original one** and then we are gonna remodify it.
Strategic Planning	State what needs to be done (e.g., discuss available resources). Divide work.	**Would you guys want me** at open lab? We just **need to attach the arm** right?
Motivational Beliefs	Share feelings of motivation regarding tasks Discuss the group's capabilities, challenges.	S1: If you want I can finish that one off. S2: **It's good. It's actually pretty fun.**
Control & Collaboration	Discuss tasks. Write/build together. Encourage group members. Ask for help.	S1: Are we saying it needs to be 2 inches below? S2: I have to add 0.08.
Progress Monitoring	Praise/evaluate an idea, a solution, or the group's progress regarding time and goals.	S1: **We're going to be done today right**? S2: I think that landing gear is **done.**
Reflection	Evaluate if the group has reached its goals. Discuss challenges in the performance.	It's pretty hard at first doing this, but **we have gone far** this week **with the codes**.
Off-task	Talk not relevant to ongoing discussion	I'm so hungry. Whose pencil is this?
Self-Regulation	Individuals ("I" perspective) about task perception, knowledge, goals, motivations.	Oh yeah, **I did it last quarter**; it's not bad. **I'll do it if you guys need.**
Shared regulation	"Our" perception of tasks; suggestions; actively constructing knowledge together. Instances where students are responding to questions and not suggesting strategies are not coded.	S3: No, **we have to do 4 ESC's** S2: **You could do like 2 or 3** S2: Oh yeah, **I thought you were talking about powering all of the other electric.**

RQ1. Discourse data—each individual's contribution to group discussion, per weekly session—were preprocessed (i.e., convert to lower cases, remove stop words, remove number, stemming), and analyzed using TAACO (The Tool for the Automatic Analysis of Text Cohesion; Crossley, Kyle, & McNamara, 2016). The unit of analysis is each turn of talk by a student (n unit = 7,514). An explanatory factor analysis was conducted on the TAACO output. The scree plot and percentage of variance explained suggested a four-factor model as the most parsimonious. Based on prior literature, we grouped the factors into *density*, *semantic*, *connectives*, and *givenness* (see Crossley et al., 2016). We retained items with factor loading > .40 and calculated factor scores for each subscale. All subscales had acceptable internal reliability (Cronbach's α for *density* = .88, *semantic* = .98, *connectives* = .89, and *givenness* = .70). We performed Kruskal-Wallis tests to examine whether there was a difference in factor scores among the groups (*below median, median, above median* peer support). Following

results that indicated statistical significance at p = .05, pairwise Mann-Whitney U tests were conducted, with Bonferroni corrections for multiple comparisons, to examine which group differed from the others.

RQ2. A set of codes for regulation types and processes was created based on prior work and iterative coding cycles using 15% of the dataset (Table 1; Marmberg et al., 2017; Jarvela & Hadwin, 2013). Because the study focused on knowledge co-construction, we coded for shared regulation ("we" perspective) instead of co-regulation ("you" perspective). Each unit received one code for regulation type and one for process. The first author and a research assistant coded 10% of the dataset, with acceptable inter-rater agreement (Cohen's $\kappa = .81$). The first author coded the rest of the data. The procedure to determine group differences is similar to RQ1. We applied ENA (Shaffer, 2017) to explore code co-occurrences. The ENA models were based on a binary matrix that reflected the presence of each pair of regulation types and process (e.g., plan-self-regulate) within 4-unit moving windows. ENA normalized the networks and visualized co-occurrences along a two-dimensional space.

Findings

Table 2. Regulation types and process (proportions) and cohesiveness (factor scores), average across sessions

		Below median (1)		Median (2)		Above median (3)		Mann-Whitney U		
		M	SD	M	SD	M	SD	1:2	1:3	2:3
Total talks per session		89.53	68.03	130.90	82.30	146.14	97.46	.17	.07	.64
Cohesion	Density	.49	1.24	-.19	.86	-.18	.77	**.02***	**.04***	.71
	Semantic	-.02	.55	-.06	1.35	.16	.15	.68	.68	.91
	Connectives	-.25	.85	.10	.96	.08	.71	.29	.76	.98
	Givenness	-.51	1.01	.19	1.04	.20	.65	**.03***	**.05***	.97
Regulation Process	Understanding	.04	.19	.07	.25	.05	.22	**.02***	.26	**.01***
	Motivation	.02	.14	.22	.41	.16	.37	***	***	***
	Strategic plan	.10	.29	.12	.32	.18	.38	.16	***	***
	Collaboration	.74	.44	.49	.50	.50	.50	***	**	.27
	Monitoring	.06	.24	.07	.26	.06	.25	1	1	1
	Reflection	.00	.06	.11	.31	.06	.24	***	**	***
Regulation Types	Self-regulate	.08	.28	.07	.26	.13	.34	.75	.27	***
	Shared	.35	.48	.24	.43	.55	.46	**	**	**.01***

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

RQ1. Text analytics reveals that generally, students with *below median* peer network has individual discourse of higher density (i.e., more unique words; M = .49, SD = 1.24) and lower givenness (e.g., fewer third-person pronouns, less information restatement; M = -.51, SD = 1.01; Table 2). These results suggest disjointed utterances where students were providing new information but not engaging in coherent knowledge construction.

RQ2. Planning and shared regulation differ by peer network expectations. Pairwise Mann-Whitney U tests indicated that students who perceived *below median* peer support tended to express task understanding, motivation, planning, reflection, and co-regulation less frequently, compared to students in the other two groups (Table 2). We applied ENA to examine the cooccurrences of regulation processes and types. The average networks of each student group (Figure 1) show students who reported *median* and *above median* peer network (blue and purple) engaged in multiple processes in tandem with self- and shared regulation, rather than just building (red).

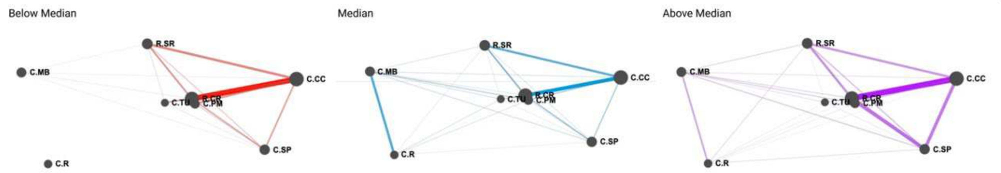

Figure 1. Epistemic network of regulation type and process co-occurrences

Discussion

Research on group knowledge construction in engineering has established a positive link between perceived social roles and participation (Bianchini, 1997). We did not find any significant difference in the total turns of talk among groups. However, differences in semantic structure highlight the need to examine patterns that are more productive in collaborative design. Students' more sophisticated design process emerged from deeper discourse at certain steps like planning, motivation, and evaluation. Prior research indicates that more socially collaborative learners, who contribute in ways that are coherent with and productive to group discourse, have significant learning gains, compared to less socially engaged learners (Dowell et al., 2018). Findings about different regulation patterns are noteworthy in design engineering. Designers who systematically set goals, frame problems, and choose among alternatives generate high-quality solutions, compared to those focused on prototyping (Atman et al., 2005).

Conclusion

This exploratory study attempts to understand what types of discourse and regulation learners embrace. Findings show that students who reported peripheral social ties seemed to engage in less socially cohesive discourse and less shared regulation of task understanding, motivation, reflection, and planning, relative to the other groups. Interpersonally cohesive discourse and systematic design evaluation have been associated with learning gains (Atman et al., 2005, Dowell et al., 2018). A practical implication is to attend to learners' perceived support systems early on and embed in instruction the types of scaffolds for regulation and discourse conducive to collaboration. From a methodological standpoint, we illustrate the potential of using computational approaches to detect co-occurrences and semantic structure beyond those that can be detected through traditional content analysis. Follow-up studies can explore the relations between cohesion and regulation features to develop training sets for automated classifiers. Existing tools that requires manual refinement of term lists are time-consuming and may inflate Type II error rates (Lee et al., 2019). Future research can also examine the individual and group-level regulation and learning outcomes in groups with different compositions. Although the three peer network groups in this study did not differ in student demographics, future work should explore in depth which students reported peripheral social network and uncover potential barriers to productive learning in collaborative design spaces.

References

Atman, C. J., Cardella, M. E., Turns, J., & Adams, R. (2005). Comparing freshman and senior engineering design processes: An in-depth follow-up study. *Design Studies*, *26*(4), 325–357.

Cohen, E. G., & Lotan, R. A. (2014). *Designing Groupwork: Strategies for the Heterogeneous Classroom Third Edition*. Teachers College Press.

Crossley, S. A., Kyle, K., & McNamara, D. S. (2016). The tool for the automatic analysis of text cohesion (TAACO): Automatic assessment of local, global, and text cohesion. *Behavior Research Methods*, *48*(4), 1227–1237. https://doi.org/10.3758/s13428-015-0651-7

Dowell, N. M. M., Nixon, T. M., & Graesser, A. C. (2019). Group communication analysis: A computational linguistics approach for detecting sociocognitive roles in multiparty interactions. *Behavior Research Methods*, *51*(3), 1007–1041. https://doi.org/10.3758/s13428-018-1102-z

Hurtado, S., Newman, C. B., Tran, M. C., & Chang, M. J. (2010). Improving the rate of success for underrepresented racial minorities in STEM fields: Insights from a national project. *New Directions for Institutional Research*, *2010*(148), 5–15. https://doi.org/10.1002/ir.357

Järvelä, S., & Hadwin, A. F. (2013). New Frontiers: Regulating Learning in CSCL. *Educational Psychologist*, *48*(1), 25–39. https://doi.org/10.1080/00461520.2012.748006

Lee, S. B., Gui, X., Manquen, M., & Hamilton, E. R. (2019). *Use of Training, Validation, and Test Sets for Developing Automated Classifiers in Quantitative Ethnography*.

Malmberg, J., Järvelä, S., & Järvenoja, H. (2017). Capturing temporal and sequential patterns of self-, co-, and socially shared regulation in the context of collaborative learning. *Contemporary Educational Psychology*, *49*, 160–174. https://doi.org/10.1016/j.cedpsych.2017.01.009

Putnik, G., Costa, E., Alves, C., Castro, H., Varela, L., & Shah, V. (2016). Analysing the correlation between social network analysis measures and performance of students in social network-based engineering education. *International Journal of Technology and Design Education*, *26*(3), 413–437.

Shaffer, D.W. (2017). Quantitative Ethnography. Madison, WI: Cathcart Press.

Schmidt, H. G., Vermeulen, L., & Van Der Molen, H. T. (2006). Longterm effects of problem-based learning: A comparison of competencies acquired by graduates of a problem-based and a conventional medical school. *Medical Education*, *40*(6), 562–567. https://doi.org/10.1111/j.1365-2929.2006.02483.x

Sinha, S., Rogat, T. K., Adams-Wiggins, K. R., & Hmelo-Silver, C. E. (2015). Collaborative group engagement in a computer-supported inquiry learning environment. *International Journal of Computer-Supported Collaborative Learning*, *10*(3), 273–307. https://doi.org/10.1007/s11412-015-9218-y

Strijbos, J. W., Martens, R. L., Prins, F. J., & Jochems, W. M. G. (2006). Content analysis: What are they talking about? *Computers and Education*, *46*(1), 29–48. https://doi.org/10.1016/j.compedu.2005.04.002

Acknowledgements

We thank the anonymous reviewers, the students who participated in this study, the lab facilitator for helping us set up the recorders. This work is supported by the National Science Foundation, Award 1535300.