International Society of the Learning Sciences

Cognitive Processing in Online Communities of Inquiry

Samuel J. Bullard, Keisha Varma,
bulla049@umn.edu, keisha@umn.edu
University of Minnesota

Abstract: The Community of Inquiry framework has been a major feature guiding research on the cognitive processes facilitated by student discourse in online learning environments. The primary analytic approaches used to assess cognitive presence have been via manual content analysis or student questionnaires. However, advancements in technology-enhanced methods offer new opportunities for expanding the analytical scope of research in online learning environments. This study sampled student discussion posts from an online undergraduate college course to explore relationships between the surface-level linguistic features observed in student posts with their level of cognitive presence. To do so, we evaluated a multi-level model to identify which phases of cognitive presence best fit students' use of cognitive processing language in their online discussion posts. Of the four cognitive presence phases, *resolution* and *exploration* were the most effective predictors of cognitive processing language in their online discussion posts.

Background

One of the most widely used frameworks for studying students' cognitive engagement in online contexts is through the Community of Inquiry framework developed by Garrison et al. (1999). A Community of Inquiry is composed of three constitutive "presences", including the cognitive, social, and teaching aspects afforded by a particular learning environment. The present study focuses its attention on students' progression of cognitive presence in online discussions. Cognitive presence is described as the progression of cognitive processes necessary for construction of knowledge and higher-order thinking skills (Garrison et al., 2001) and has been found to be a major influence in students' perceived and actual learning outcomes in an online course (Akyol & Garrison, 2011; Galikyan & Admiraal, 2019). Cognitive presence is primarily operationalized by the Practical Inquiry (PI) model, which describes a cyclical process composed of four phases: (1) a triggering event, in which a problem or dilemma arises, (2) exploration of the problem, including brainstorming and information exchange, (3) integration or synthesis of ideas generated in the prior phases, and finally (4) resolution of the initial problem or inquiry consisting of testing hypotheses or application to real-world situations (Garrison et al., 1999). However, the assessment of students progression through the PI model has in prior research been limited to the measurement capabilities of individual research teams. The measurement of cognitive presence typically consists of using PI phase-level indicators to manually code discussion forum transcripts, which is often a labor-intensive and time-consuming process (Kovanović et al., 2016). Although this analytic approach may be feasible for researchers in the learning sciences, such demands raise concerns regarding their applicability to real-world educational contexts. In the present study, we utilize the Practical Inquiry model to identify the progression of students' cognitive presence in asynchronous online discussions according to the Community of Inquiry framework. In addition, we analyze these same posts using a validated Language Inquiry software (LIWC; Pennebaker et al., 2015) to quantitatively measure the frequencies of words related to cognitive processes used in these discussion posts. In combination, these two analytical approaches will be used to answer the following research question: *To what extent can the linguistic features of students' online discussion posts be explained by the phases of cognitive presence, as defined by the Practical Inquiry model?*

Method

Research context, participants, and data sources

The research context for this study was an online introductory course for undergraduate college students studying educational psychology at a large midwestern university in the United States. A total of 20 undergraduate students enrolled in this 14-week long course during the Fall of 2020 for academic credit. While in the past the course was typically held in a hybrid format, public health restrictions required the course to be held completely online. In this modified version of the course, the instructor utilized asynchronous discussion threads for students to collaboratively reflect on the assigned weekly readings. At the top of each discussion thread, the instructor provided several suggested prompts for students to focus their posts on. These prompts encouraged students to raise questions or issues, draw connections, propose alternative solutions or explanations, and seek clarification when necessary. Across 12 discussion threads, students contributed a total of 235 posts (M = 11.75, SD = 1.97) throughout the semester. We collected these discussion transcripts by manually copying and pasting student responses from the university's online learning platform to word processing software.

Procedure

Unit of analysis

We underwent a procedure to segment each student post into thematic units, or 'units of meaning' (Rourke et al., 2001), described as a sentence or series of sentences which convey a single idea or theme (Kleinheksel et al., 2020). After carefully reading the entire passage of a students' discussion post, a team of two coders independently attempted to identify the primary arguments or ideas brought up by the author of the post. These coders would then segment these passages at specific points, often indicated by linguistic markers such as "However,", "On the other hand," or "Additionally,". Following an acceptable level of segmentation reliability (IRR = .83), the two coders segmented the remaining data independently. After segmenting 235 total discussion posts, the final sample consisted of 726 thematic units (M = 36.3, SD = 7.23) which were used for all subsequent analyses.

Practical Inquiry coding

The same team of coders from the segmentation procedure were then trained on conducting the manual content analysis adopted from prior coding schemes based on the Practical Inquiry model (Garrison et al., 1999; Park, 2009). Specifically, this study initially utilized the set of Practical Inquiry indicators used in Park's (2009) replication study. This set was chosen based on the similarity to the original set of indicators established by Garrison et al. (1999) as well as its specificity for implementation purposes. A set of 55 thematic units from our sample was utilized to train coders on this coding scheme. After two rounds of training, an excellent degree of coder reliability was met (IRR = .87). After this finding, the two coders independently coded the remaining data.

Linguistic analysis

The Language Inquiry and Word Count (LIWC; Pennebaker et al., 2015) software counts words from text inputs and classifies them according to validated dictionaries related to a variety of psychological, social, and emotional constructs. It then calculates a numerical measure by dividing the total number of words related to these individual constructs by the total number of words in the text input. The word category of interest to this study was that of cognitive processing, an aggregate of several subcategories including insight, causation, discrepancies, differentiation, and tentativeness/certainty. We processed 726 individual thematic units through the LIWC software and received numerical values for cognitive processing words which would then be used as a dependent variable to model according to the four phases of cognitive presence in the Practical Inquiry model. See Table 1 to see a side-by-side comparison of the two analyses described in this section.

Table 1
Exemplar Practical Inquiry Coding for Cognitive Presence and Analysis of Cognitive Processing Language

Example Text	Practical Inquiry Code	Cogproc
The discussion and **conclusion** sections of this paper were super interesting in my **opinion**. I **thought** it was interesting **how** he looked at the actual structure of **lessons** and **how** that impacts **learning** and **not** just focusing on the content. He describes **how all** students have prior **knowledge** that they bring to the classroom, and **how** these prior foundations have an impact on **how** students **learn** the material that is being presented.	Triggering Event	19.44
I **thought** this article was **really** interesting **because** throughout school there have been times when I've been assigned to teach the class about a **certain** topic, and I **always** hated doing it. **However**, I **definitely** did **learn** more and **remembered** the material better after having taught it.	Exploration	19.15
Furthermore, I have **learned** in previous classes before as well as **some** of the **other** readings in this course, that children are "little scientists". The prevalence of mental models supports this **idea** by showing the child **making changes** to their original **theory** in order to **make sense** of the new **questions** they are being asked.	Integration	18.18
As a future teacher this reminds me of **how** important it is to put effort into **making sure** your students are **all** on the same page before **learning** new **information**, **otherwise** it can be misinterpreted and harder to **correct** later on. **Information** can be difficult to **understand** when the **perspective** of the teacher does **not** align with the **perspective** of a student regarding a topic.	Resolution	20.00

Note. Bold formatted text in first column represents words categorized as cognitive processing by the LIWC software. Values contained under the "Cogproc" column represent real LIWC output, first calculated by the quotient of cognitive processing words divided by the total word count in the thematic unit, then multiplied by 100.

Results

First, we examined preliminary results and computed descriptive statistics of our predictor variables (see Table 2). Regarding the cognitive presence predictors, the most frequent phase coded for in the Practical Inquiry model was the integration phase, with a median of 11.5 across the student sample. The next most observed phase from the model was *triggering events*, followed by *exploration*, and finally *resolution*. Note that during the coding process, there were a handful of thematic units for which we were unable to determine its alignment with any of the phases of cognitive presence. These codes ("Other"; n = 25) typically consisted of students greeting their peers, or statements which were wholly unrelated to the learning material. Seeing their irrelevance to the study's research question, which inquires specifically about the phases of cognitive presence, we did not include them as variables in our multi-level model. Second, we inspected general trends in cognitive processing language scores as measured by LIWC software. Across all 726 thematic units, cognitive processing language values ranged from 0.00 to 43.48, with a grand mean of 19.14, indicating that on average, about 19% of words used in each unit were reflective of students' cognitive processes. Next, we calculated descriptive statistics for cognitive processing language according to the unit's corresponding cognitive presence code.

Table 2
Correlation Coefficients and Descriptive Statistics of Model Predictors and Outcome Variables.

Code	Coefficient					Predictor (Cognitive Presence)		Outcome (Cognitive Processing)
	1	*2*	*3*	*4*	*5*	*N*	*Median (CI)*	*M (SD)*
Triggering Event	-					166	9.5 (4.8, 11)	18.42 (6.89)
Exploration	.31	-				173	8.5 (7.8, 11)	19.97 (6.15)
Integration	.22	-.13	-			257	11.5 (9, 14.3)	18.75 (5.08)
Resolution	-.29	-.10	-.34	-		105	4.5 (2.8, 7.5)	20.20 (5.29)
Other	.11	.36	.01	-.16	-	25	1.0 (0.0, 2.0)	17.63 (7.81)

Multi-level modeling

Next, we conducted a multilevel regression analysis to investigate the relationships between individual cognitive presence phases and cognitive processing language used in online discourse. These analyses began with an examination of the fully specified mixed effects model, which was comprised of two primary components: (1) the fixed effects of cognitive presence phases, and (2) the random effects structure. The fixed effects included in the full model consisted of a binary variable for whether the unit was coded as either a Triggering Event, Exploration, Integration, or Resolution phase of the Practical Inquiry Model. Because this was intended as an exploratory, rather than confirmatory study, we included all four of these phases as fixed effects to avoid making any faulty presumptions about which phase of cognitive presence would be more influential to cognitive processing language. The second component of the fully specified model was the random effect's structure. Due to the inherently nested nature of the data, we made an a priori decision to include random effects for both student-level variation (i.e., students wrote multiple discussion posts throughout the course) and nested post-level variation (i.e., each students discussion posts were segmented into multiple thematic units for analyses). In this model, we found the strongest effects of the *resolution* and *exploration* phase of cognitive presence on cognitive processing language (see Table 3).

Table 3
Coefficients and Standard Errors of Model-Level Predictors

Fixed Effects	β	*SE*
Intercept	17.599	1.195
Resolution	2.578	1.307
Integration	1.035	1.225
Exploration	2.412	1.248
Triggering Event	0.657	1.253

Discussion

The results of this study explored the research question: *To what extent can the linguistic features of students' online discussion posts be explained by the phases of cognitive presence?* In this sample, students' use of language reflecting their cognitive processes was dependent on the phase of cognitive presence coded in their discussion posts. Specifically, this study found that the *exploration* and *resolution* phases of cognitive presence were most impactful in the explanation of variation in cognitive processing language. The highest levels of cognitive processing language in this study were observed when student messages were coded at the *resolution* phase. In the Community of Inquiry framework, *resolution* is an indicator of higher-order thinking and is characterized by consensus-building, hypothesis testing, and the application of knowledge generated from prior discourse (Garrison et al., 2001). Prior work has demonstrated an association between the frequencies of this phase in discourse and student achievement (Galikyan & Admiraal, 2019), suggesting that students' reaching this level of cognitive presence in online discourse is a consequential aspect for their learning. The exploration phase of cognitive presence occurs after a *triggering event*, or when problem or dilemma is recognized by members within a Community of Inquiry. During *exploration*, students are shifting from a state of internal critical reflection to one of brainstorming or exchanging information with peers in online discourse (Garrison et al., 2001). Although the present study did not investigate the subcategories of the LIWC cognitive processing variable, our findings align with prior work which has demonstrated associations between the *exploration* phase and two *cognitive processing* subcategories, specifically *causal* (e.g., "because", "effect") and *exclusive* (e.g., "but", "without") words (Joksimović et al., 2014). Higher frequencies of words in these subcategories may potentially explain the generally higher level of cognitive processing language observed during the exploration phase of the practical inquiry model. Continued inquiry into these questions would both benefit the field of computer-supported collaborative learning and expand our theoretical understanding of the Community of Inquiry framework.

References

Akyol, Z., & Garrison, D. R. (2011). Understanding cognitive presence in an online and blended community of inquiry: Assessing outcomes and processes for deep approaches to learning: Cognitive presence in an online and blended community of inquiry. *British Journal of Educational Technology*, *42*(2), 233–250. https://doi.org/10.1111/j.1467-8535.2009.01029.x

Galikyan, I., & Admiraal, W. (2019). Students' engagement in asynchronous online discussion: The relationship between cognitive presence, learner prominence, and academic performance. *The Internet and Higher Education*, *43*, 100692. https://doi.org/10.1016/j.iheduc.2019.100692

Garrison, D. R., Anderson, T., & Archer, W. (1999). Critical Inquiry in a Text-Based Environment: Computer Conferencing in Higher Education. *The Internet and Higher Education*, *2*(2–3), 87–105. https://doi.org/10.1016/S1096-7516(00)00016-6

Garrison, D. R., Anderson, T., & Archer, W. (2001). Critical thinking, cognitive presence, and computer conferencing in distance education. *American Journal of Distance Education*, *15*(1), 7–23. https://doi.org/10.1080/08923640109527071

Joksimović, S., Gašević, D., Kovanović, V., Adesope, O., & Hatala, M. (2014). Psychological characteristics in cognitive presence of communities of inquiry: A linguistic analysis of online discussions. *The Internet and Higher Education*, *22*, 1–10. https://doi.org/10.1016/j.iheduc.2014.03.001

Kleinheksel, A. J., Rockich-Winston, N., Tawfik, H., & Wyatt, T. R. (2020). Demystifying Content Analysis. *American Journal of Pharmaceutical Education*, *84*(1), 7113. https://doi.org/10.5688/ajpe7113

Kovanović, V., Joksimović, S., Waters, Z., Gašević, D., Kitto, K., Hatala, M., & Siemens, G. (2016). Towards automated content analysis of discussion transcripts: A cognitive presence case. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge - LAK '16*, 15–24. https://doi.org/10.1145/2883851.2883950

Park, C. L. (2009). Replicating the Use of a Cognitive Presence Measurement Tool. *Journal of Interactive Online Learning*, *8*(2), 16.

Rourke, L., Anderson, T., Garrison, D. R., & Archer, W. (2001). Methodological Issues in the Content Analysis of Computer Conference Transcripts. *International Journal of Artificial Intelligence in Education*, *11*. https://auspace.athabascau.ca/handle/2149/715

