International **R**eview of **R**esearch in **O**pen and **D**istributed **L**earning
Volume 27, Number 3

August – 2026

Structure and Content of Interactions: Social Presence in Asynchronous Online Whole-Class vs. Small-Group Discussions

K. Bret Staudt Willet[1], Chufeng Bai[1], Jaesung Hur[1], and Mete Akcaoglu[2]
[1]Florida State University; [2]Georgia Southern University

Abstract

This study investigates how the format of asynchronous online class discussions (whole class vs. small group) affects the structure and content of student interactions. Data were collected from 8 weeks of discussion forums in a graduate-level, asynchronous, online course at a comprehensive university in the southeastern United States. Social network analysis (SNA) metrics, such as response rate, reciprocity, transitivity, diameter, density, in-degree, and closeness centrality, were used to evaluate interaction structure. Linguistic Inquiry and Word Count (LIWC) software was used to analyze the linguistic features of the interaction content. Findings revealed differences between the formats: whole-class discussions exhibited greater diameter and variability in response rate, emphasizing formal and analytical communication. In contrast, small-group discussions displayed higher reciprocity and transitivity, as well as more support and socioemotional engagement. Findings underscore the importance of designing and facilitating collaborative online learning experiences by strategically employing discussion formats to meet specific learning objectives. Implications for practice include incorporating both whole-class and small-group discussions to leverage the strengths of each format, such as creativity and critical thinking in larger groups and mutuality and support in small groups. Future research should examine these effects across different educational levels and subject areas, considering instructional contexts and personal learner characteristics.

Keywords: online learning, class discussion, group size, social network analysis, natural language processing

Introduction

Online courses are increasingly viewed as an important part of modern higher education. With the advancement of technology and changing times, more students are participating in online courses. According to the National Center for Education Statistics, more than half of students in higher education in the United States (approximately 10 million students, or 54.2%) participated in online courses in fall 2022 (2023), twice as many as a decade earlier, with five million students in fall 2013 (2015).

Online learning presents both advantages and challenges. A primary advantage is its flexibility and accessibility, allowing students from various regions to access quality education and teaching materials without the constraints imposed by physical distance and time (Ho et al., 2023). In addition, online learning supports the integration of diverse teaching tools, including multimedia, simulations, and interactive modules, thereby enriching the educational experience. However, online education also presents persistent challenges, including technical support, assessment integrity, and the necessity for additional teacher training (Palau et al., 2021). Furthermore, learners' experiences of social presence are often diminished in online learning. Online learning often features less robust communication and interaction than face-to-face instruction, which may result in decreased motivation, participation, and satisfaction (Sung & Mayer, 2012). Less engagement can lead to some learners feeling less connected and becoming less actively involved in their online educational journey (Ho et al., 2023).

Discussion forums are one of the most common course activities used to increase interaction and social presence in asynchronous online courses. Beyond their social benefits, discussion forums also help students understand course objectives, provide real-world applications, and can be used to promote student engagement (Wickershaw & Dooley, 2006). Although instructor feedback can help increase student engagement, peer feedback is equally important (Salter & Conneely, 2015).

Group discussions can enhance students' social skills and their awareness of societal issues (Akcaoglu & Lee, 2016; Steponavičius et al., 2023), but the size of these groups influences student engagement. For instance, larger discussion groups have been associated with lower student participation, poorer individual performance, and weaker team cohesion; conversely, smaller groups tend to yield better outcomes (Zhan et al., 2022). Recent research in this area has tended to focus on the impact of group size differences on learners' knowledge and skills (Elaish et al., 2023), critical thinking (Yang et al., 2022), and engagement levels (Luo et al., 2023; Zhan et al., 2022).

In our study, we extended research on online discussion group size to include the nature of the discussions themselves. Specifically, we compared the discussion group size between the format of whole-class discussions versus those in small groups. The purpose of this study was to investigate how the format of asynchronous online class discussions affects the structure and content of student interactions.

Framework

We have framed this study of online class discussions in terms of Kreijns et al.'s (2022) conceptualization of *social interaction*. They began with Short et al.'s (1976) foundational definition of *social presence*, which aimed to explain how different media types shape interpersonal interactions—specifically, the degree to

which an individual is perceived as "real" and salient in mediated communication. Garrison et al. (2000) applied social presence to online learning contexts, emphasizing learners' ability to project social and emotional aspects of themselves to help other students to perceive them as real people. Rourke et al. (1999) made the concept more concrete, operationalizing social presence through the assessment of interactive content in asynchronous online discussions, categorized as *affective* (e.g., expression of emotions), *interactive* (e.g., continuing a thread), and *cohesive* (e.g., referring to the group using inclusive pronouns) indicators. Finally, Kreijns et al. (2022) argued that with advances in communication technologies since the time of Short et al.'s (1976) original definition, social presence is now better understood in terms of *being present*—that is, in the here and now—rather than in terms of being real.

Recent research presents social interaction as a multidimensional construct that includes factors in addition to social presence. Even three decades ago, Gunawardena (1995) noted that social presence depends not only on the medium itself but also on how participants communicate within it. Kreijns et al. (2022) further clarified this point by reconceptualizing social interaction as social presence in dynamic interplay with social space and sociability. They defined *social space* as the quality of interpersonal and emotional connections between group members and *sociability* as the affordance of the medium or technology to allow for the expression of social presence. In our study, we investigated social interaction in online discussion groups by considering social space through indicators of *interaction structure* (e.g., frequency and mutuality of communication between students) and sociability through indicators of *interaction content* (e.g., the affective tone in the expression of information).

Literature Review

Social Presence in Online Education

Social presence, as a key aspect of social interaction (Kreijns et al., 2022), plays a critical role in relationship-building for student participation and online learning inclusion (Lowenthal et al., 2020). Social presence is influenced not only by technology (e.g., communication media) but also by social context and personal factors (Gunawardena, 1995; Kim et al., 2011). Social presence brings specific benefits to group learning, such as comparing notes and correcting misunderstandings, which helps build trust and learner communities (Lim, 2023; Richardson et al., 2017). Considering the seven main motivations for students to participate in online courses, Van Wart et al. (2020) found that although social presence was a relatively low-ranking factor among the seven motivations, it was closely related to online course satisfaction, course participation, and interaction between students. Similarly, Richardson et al.'s (2017) meta-analysis showed that students who experience higher levels of social presence in online environments experience higher levels of learning satisfaction, engagement, academic self-efficacy, self-regulated learning, and learning performance. Student-student interactions and teacher-student interactions increase social presence, which then reduces learners' social isolation and loneliness and makes students more engaged in learning activities (Phirangee & Malec, 2017). However, despite the importance of social presence, online learning has been criticized for lacking social presence and having lower sociality compared to face-to-face learning environments (Adnan & Anwar, 2020).

To increase social presence in online learning environments, discussion forums can be established as a space for interaction between students—a viable alternative to traditional face-to-face focus groups (Lowenthal & Dunlap, 2020). However, past research has shown that group size in computer-mediated environments influences students' learning activity participation and learning performance (Luo et al., 2023; Yang et al., 2020). Group size can impact group cohesion, student learning, and social perception; smaller group sizes have been connected to greater perceived social presence, more sense of community, and better learning outcomes (Akcaoglu & Lee, 2016; Michaelsen et al., 2023). Small groups allow members to take more responsibility and be more invested in outcomes (Qiu et al., 2014); students in this format view and comment on their peers' posts more than when in whole-class discussions (Yang et al., 2020). Students in small groups experience less information overload in addition to more social presence; when assigned to small groups, following discussion threads is easier compared to when in whole-class discussions (Qiu & McDougall, 2015; Yang et al., 2020). The reduced reading load, in turn, creates more time for reflection and thoughtful responses.

Measuring Interaction Structure Through Social Network Analysis

Although there are various methods to measure social interaction in online courses, social network analysis (SNA) has been shown to be an effective quantitative approach. Specifically, SNA examines the relationships between social entities and their influence within a network (Dado & Bodemer, 2017; Norz et al., 2023) by considering a set of objects (i.e., nodes) and their connections (i.e., edges). A node is an individual engaged in networks, and an edge represents an interaction between individuals (Dado & Bodemer, 2017; Ye & Pennisi, 2022). Through SNA, researchers can identify an influential node that has more connections, examine the strength of collaboration in the network, and illustrate the structure of connections in the network (Ye & Pennisi, 2022). Because SNA is calculated from observations of a learning environment, it reflects the structure of real interactions rather than perceived ones (Oh et al., 2025).

Past SNA research has used a variety of metrics to examine social networks, including measures of individuals' roles within the network and measures of the overall network structure (e.g., reciprocity, transitivity, diameter, density). In terms of individuals' roles, there are two kinds of centrality: degree and closeness. *Degree centrality* quantifies the connections between nodes; a higher degree-centrality score means that a node has more direct links to other nodes. Furthermore, *in-degree* refers to the number of incoming connections pointing to a node (e.g., how many responses a student receives to a discussion post), and *out-degree* refers to the amount of outgoing connections starting from a node, such as the total number of times a student replies to any of their classmates (Ye & Pennisi, 2022). In addition, *closeness centrality* calculates the proximity of a node to others, which can indicate the key influencers in a network (Ye & Pennisi, 2022).

As a measure of the overall network structure, *reciprocity* measures the relationship between a pair of nodes in terms of the probability of mutual connections; higher reciprocity indicates more back-and-forth conversations, whereas lower reciprocity suggests more one-directional conversations in the network (Staudt Willet & Carpenter, 2020). *Transitivity* is a measure of clustering and group cohesion in the network; higher transitivity indicates that adjacent nodes are more likely to be connected, and lower transitivity means that discussions are more evenly distributed across the whole network (Staudt Willet &

Carpenter, 2020). *Diameter* indicates the size of the network, as understood by the number of nodes required to traverse from one end of the network to the other. Diameter provides insight into the proximity of two network participants by representing the maximum distance between two nodes (Tabassum et al., 2018). Finally, *density* calculates the number of edges present between nodes out of all connections possible in the network (Norz et al., 2023). Density is an indicator of how actively individuals communicate or collaborate with each other; higher density indicates a highly connected network, and lower density implies a loosely connected network.

Researchers have conceptualized online learning communities as social networks and used SNA to investigate the complexity of relationships and connections among network members—that is, students' participation and interactions (Dado & Bodemer, 2017; Lim, 2023). To date, numerous studies have established SNA as an effective tool for mapping social interactions within text-based asynchronous online discussions (Ye & Pennisi, 2022); researchers have also used SNA metrics to analyze social presence, specifically, in online learning environments (e.g., Norz et al., 2023). For instance, Lim (2023) examined SNA metrics such as out-degree, in-degree, betweenness centrality, closeness centrality, and eigenvector centrality to understand the relationship between social presence and students' cognitive or affective learning outcomes. Similarly, Satar and Akcan (2018) found a statistically significant correlation between SNA metrics and social presence indicators.

Measuring Interaction Content Through Linguistic Features

In online discussions, students communicate through written or verbal language. Researchers have claimed that the content of students' interactions is more important than the frequency of posts and comments (Gasell et al., 2022). Previous research has studied students' interactions through qualitative content analysis or thematic analysis (e.g., Rourke et al., 1999). However, Pennebaker et al. (2007) proposed a different approach: to computationally analyze the content of text data by examining psychological characteristics evident through linguistic features. To accomplish this, they developed Linguistic Inquiry and Word Count (LIWC), a software program that categorizes text according to more than 80 linguistic features, such as emotional tone, thinking styles, and linguistic functions (https://www.liwc.app/). By examining linguistic features, researchers can gain information on one's attention, emotional status, social relationships, social hierarchy, social coordination and group processes, honesty and deception, close relationships, thinking styles, or individual differences (Tausczik & Pennebaker, 2010).

Although computationally analyzing linguistic features loses the nuanced insights of qualitative coding by hand, it gains the advantage of scalability, automating the assessment of linguistic features across potentially large amounts of text data. This approach can still provide insights into group interactions and dynamics (Scholand et al., 2010), and researchers have used LIWC to study social and cognitive presence in online environments. For instance, Joksimović et al. (2014) used LIWC to analyze 1,747 messages in the online discussion forum of a graduate-level course, finding that linguistic features can serve as indicators of different phases of cognitive presence: triggering, exploration, integration, and resolution. Comparing the linguistic features in user comments on two online platforms (i.e., YouTube and edX) at a larger scale of 6,938 comments from 2,889 contributors, Gao et al. (2024) found that social presence appeared mostly

similar on both platforms. Differences in sentiment were associated with community characteristics and expectations related to the technological affordances of each platform.

Purpose and Research Questions

Although previous research has investigated social interactions and social presence in online learning environments, it remains unclear whether group size and format, such as assigning students to small-group discussions, are associated with the nature of those discussions. The literature on group size in online discussions has largely focused on learning outcomes, participation frequency, or cognitive performance; few studies have examined how discussion formats (e.g., whole class vs. small group) influence the interaction structure and content in asynchronous environments.

To fill this research gap, the purpose of this study was to investigate how the format of asynchronous online class discussions (whole class vs. small group) affects the structure and content of student interactions. This study went beyond self-reported explanations of social interaction and social presence in asynchronous online discussions; we followed a computational text analysis approach to investigate real, observable behavior. To accomplish this purpose, we posed two research questions:

- RQ1. How does the structure of social interactions differ between small-group and whole-class discussions?
- RQ2. How does the content of social interactions differ between small-group and whole-class discussions?

Method

This research followed a quantitative case study design where the case was an online, graduate-level, asynchronous course. Students participated in whole-class discussions in weeks 1–4 and in randomly generated small groups for Weeks 5–8; each small group consisted of four or five class members. Discussion prompts asked students to summarize their learning from the week's materials and reflect on how it related to their practice. They were also required to comment on at least one class or group member's original post.

Data Collection

The data for the current study involved a retrospective analysis of existing student discussion posts that were originally produced as part of regular assignments in the graduate-level course Assessment and Data Analysis in Teaching at a comprehensive university in the southeastern United States. Online class discussions occurred asynchronously in the learning management system for the course, BrightSpace. No data were accessed or analyzed until Institutional Review Board (IRB) approval was obtained following the conclusion of the course. Upon IRB approval, all discussions were fully deidentified prior to analysis. The instructor (a co-author on this paper) exported the discussion forums as text data containing the content of posts and responses, the discussion format, and information regarding senders and receivers. Notably, although the data came from the fall 2014 version of the course, the core structure of this asynchronous online course remains unchanged in that it still covers the same content, same assignments, and discussion structure. In total, the dataset included 460 posts and responses from 19 students.

Data Analysis

We used R (Version 4.5.2) for data cleaning and to conduct all statistical analyses. To answer our first research question regarding the structure of social interactions, we used sender-and-receiver information (i.e., posts and responses) to investigate the social network structures of these discussions as indicated by SNA metrics. The unit of analysis was the number of posts and responses, as well as their connections to each other. We calculated descriptive statistics (e.g., mean, standard deviation, range, median) of the weekly number of posts, response rate (the percentage of posts that received at least one response), reciprocity (the likelihood of mutual connections), transitivity (how clustered interactions were), diameter (how disconnected the furthest two participants were), density (how many actual interactions there were relative to all possible interactions), in-degree (how connected participants were through receiving responses), and closeness centrality (proximity of participants to others, indicating influence). We checked the skewness and kurtosis of each measure; nearly all had a kurtosis score indicating a non-normal distribution. Therefore, we descriptively compared the medians when reporting the findings.

To answer our second research question regarding the content of social interactions, we created a data visualization to compare the likelihood that the most common terms appeared in whole-class versus small-group discussions. We calculated log-odds ratios for each term to show in which discussion format the term was more likely to appear. We also used the 2022 version of LIWC (Pennebaker et al., 2022) to analyze the linguistic features of discussion posts and responses. With LIWC, we calculated the median of numerous linguistic features (Table 1; Boyd et al., 2022) across 460 messages in the online discussion forum. We compared the presence of LIWC features in whole-class versus small-group discussions using a series of Mann-Whitney U tests, specifically Wilcoxon rank sum tests as the special case for two samples (Hollander et al., 2014). We used these non-parametric tests because LIWC features often violated the assumption of normality required for parametric *t*-tests.

Table 1

LIWC Categories and Descriptions

Category	Description
Word count	Total number of words
Words per sentence	Mean number of words per sentence
Big words	Percentage of words seven letters or longer
Dictionary words	Percentage of words captured by LIWC
Analytical thinking	Words indicating logical, formal thinking
Clout	Words indicating leadership, status
Authentic	Words indicating honesty, genuineness
Emotional tone	Degree of positive or negative tone
Linguistic	Words including pronouns (e.g., I, they) and function words (e.g., the, to, and)
Drives	Words indicating power, affiliation, achievement
Cognition	Words indicating thinking processes: insight, causation, differentiation

Category	Description
Affect	Words indicating emotional processes (positive and negative): anxiety, anger, sadness
Social processes	Words indicating social processes: communication, politeness, conflict
Culture	Words indicating facets of human life: politics, ethnicity, technology
Lifestyle	Words indicating leisure, home, work, school, money, religion
Physical	Words indicating health, wellness, illness, mental health, sex, food, death
Perception	Words indicating attention, motion, sensory
Conversational	Words indicating netspeak, assent, fillers (e.g., yeah, oh, okay)

Note. LIWC = Linguistic Inquiry and Word Count software; *netspeak* means informal, online language.

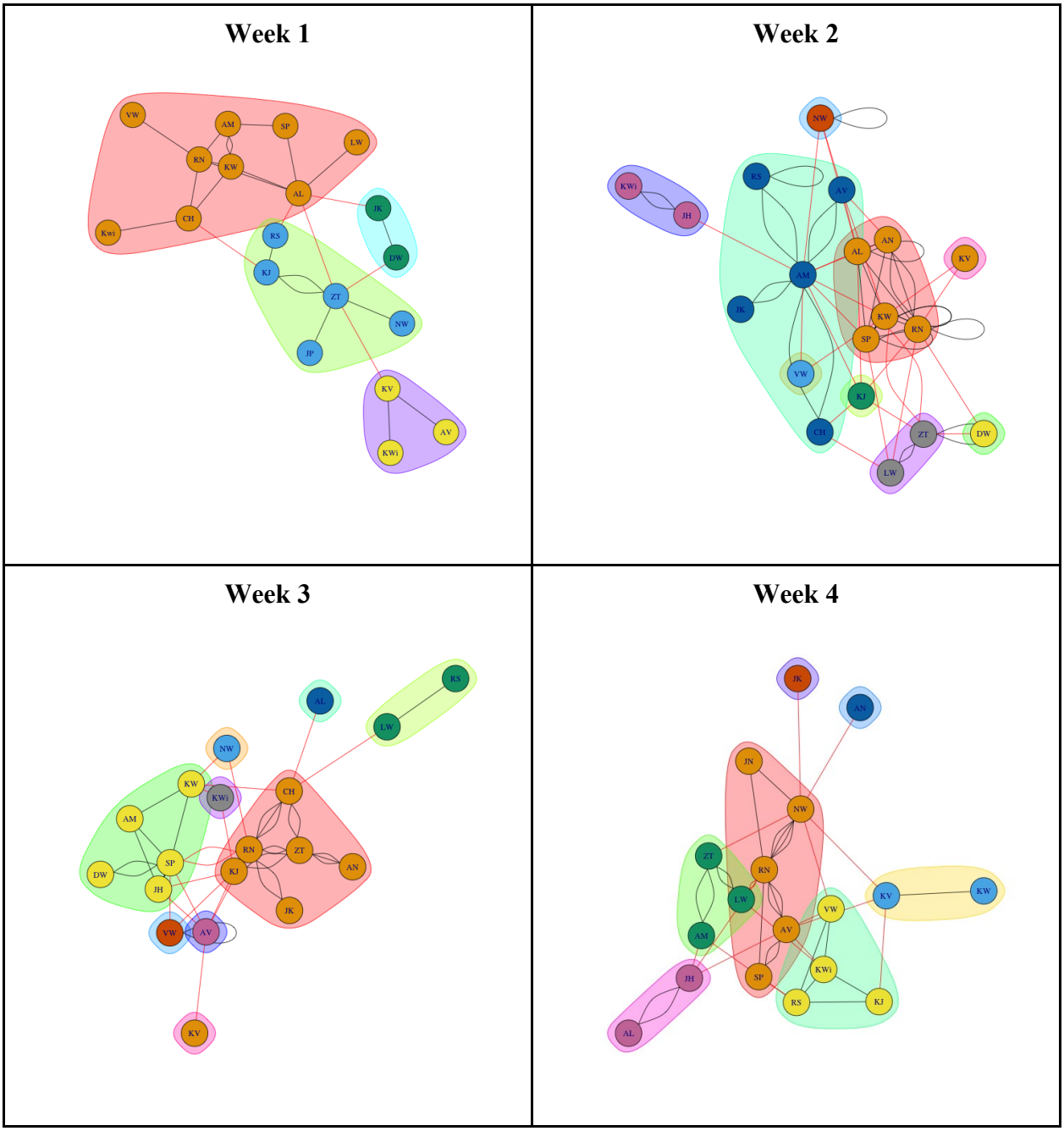
Results

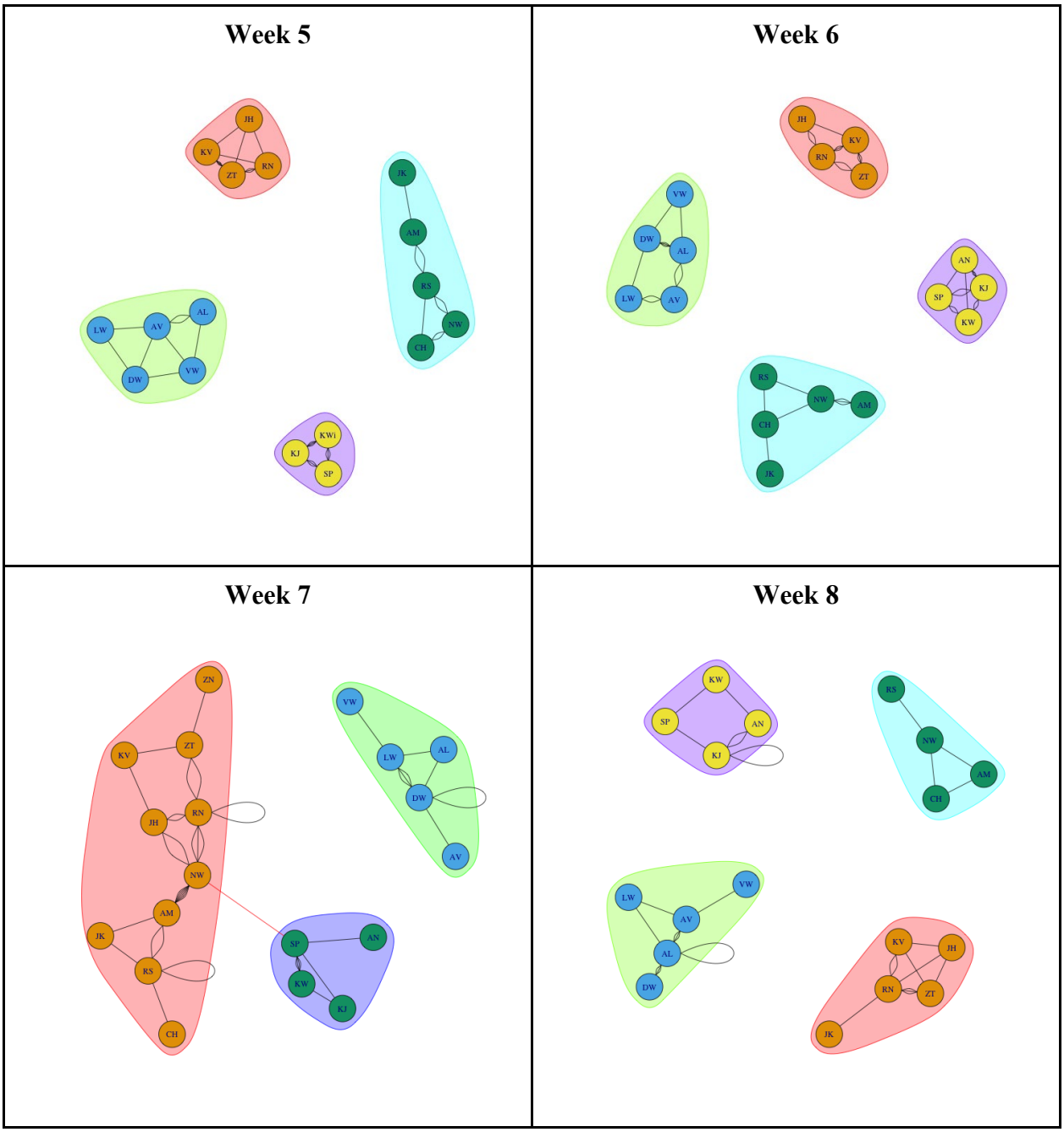
RQ1. Structure of Social Interactions

The analysis of the network structure over 8 weeks showed a shift in interaction dynamics as the class transitioned from a whole-class model (weeks 1–4) to small-group interactions (weeks 5–8), leading to compact and well-defined clusters (Figure 1). In weeks 1–4, most students' interactions were spread out across a wide range of peers, with several isolated students with peripheral engagement. However, their interactions became tighter and more interconnected in clearly bounded small groups over the remaining weeks. One notable exception occurred in week 7, when student NW responded to posts in their assigned small group as well as a post in another group. NW's crossover activity served as a bridge between those two small groups, causing the network diagram to depict those two groups as one.

Figure 1

Network Structure Across Eight Weeks (Weeks 1–4: Whole Class, Weeks 5–8: Small Group)





Note. Letters in the circles represent students in the course. Lines between the circles indicate students responding to other students in the weekly discussion. A line that loops back to where it begins indicates a student who responded to their own discussion post. The colorful clouds around the circles indicate computationally derived clusters in which students are more likely to have responded to other students within the cluster than those outside it.

SNA metrics (Table 2) provide further insight into the differences between whole-class and small-group discussions evident in Figure 1. For instance, the initial, whole-class phase when class discussions exhibited a more loosely connected and distributed network structure was supported by the calculation of the median network diameter (6.0 compared to 3.5), indicating that whole-class discussions spanned many students but were not uniformly connected. In addition, although the response rate was similar between whole-class

and small-group discussions (median of 69.8% compared to 66.4%), the whole-class format exhibited much more variability, with a response rate ranging from 59.6%–75.3% across the 4 weeks, whereas the small-group response rate only ranged from 63.0%–69.5%. This suggests that students were just as likely to receive responses in either format; however, students' interactions were unevenly distributed, with a concentration around a few individuals in the whole-class format.

After the transition to a small-group format in weeks 5–8, we observed notable changes in reciprocity and transitivity. Median reciprocity scores increased from 35.2% to 56.0%, and median transitivity rose from 21.9% to 63.3%; these metrics suggest stronger mutual engagement and more cohesive clustering within the smaller groups. Additionally, the decreased network diameter signifies shorter paths and more focused interactions within groups. Similarly, the median closeness centrality score rose from 0.4 to 0.8, indicating more direct and prompt interactions among discussion participants. However, despite these marks of increased interactivity, the overall volume of contributions decreased slightly from a median of 59.5 posts during whole-class weeks to 58.0 posts during small-group weeks. A larger difference was seen between the maximum weekly posts—the most in a whole-class discussion week was 73 posts, compared to a maximum of 59 posts during small-group discussions. Despite fewer initial posts in small groups, students still received a similar number of responses from classmates, as shown by an identical median in-degree of 2.0.

Table 2

Descriptive Statistics for Social Network Analysis Metrics

Variable	Whole class				Small groups			
	M	*SD*	Range	*Mdn*	*M*	*SD*	Range	*Mdn*
Weekly posts	59.8	10.6	47–73	59.5	55.2	6.2	46–59	58.0
Response rate	68.6%	6.7%	59.6–75.3%	69.8%	66.3%	2.9%	63.0–69.5%	66.4%
Reciprocity	34.5%	10.6%	21.4–46.2%	35.2%	52.0%	11.3%	35.7–60.0%	56.0%
Transitivity	22.5%	5.2%	17.5–28.6%	21.9%	58.3%	20.8%	29.0–77.4%	63.3%
Diameter	5.8	1.3	4–7	6.0	3.8	1.0	3–5	3.5
Density	14.6%	2.7%	11.7–18.0%	14.4%	12.4%	2.0%	9.5–13.6%	13.2%
In-degree	2.4	2.3	0–10	2.0	2.1	1.7	0–8	2.0
Closeness	0.5	0.2	0.2–1.0	0.4	0.7	0.2	0.3–1.0	0.8

Overall, the analysis of the structure of social interactions seems to indicate fewer, but more meaningful, exchanges during small-group class discussions. That is, students in small-group discussions received marginally fewer responses to their posts but were much more likely to have mutual exchanges (i.e., higher reciprocity).

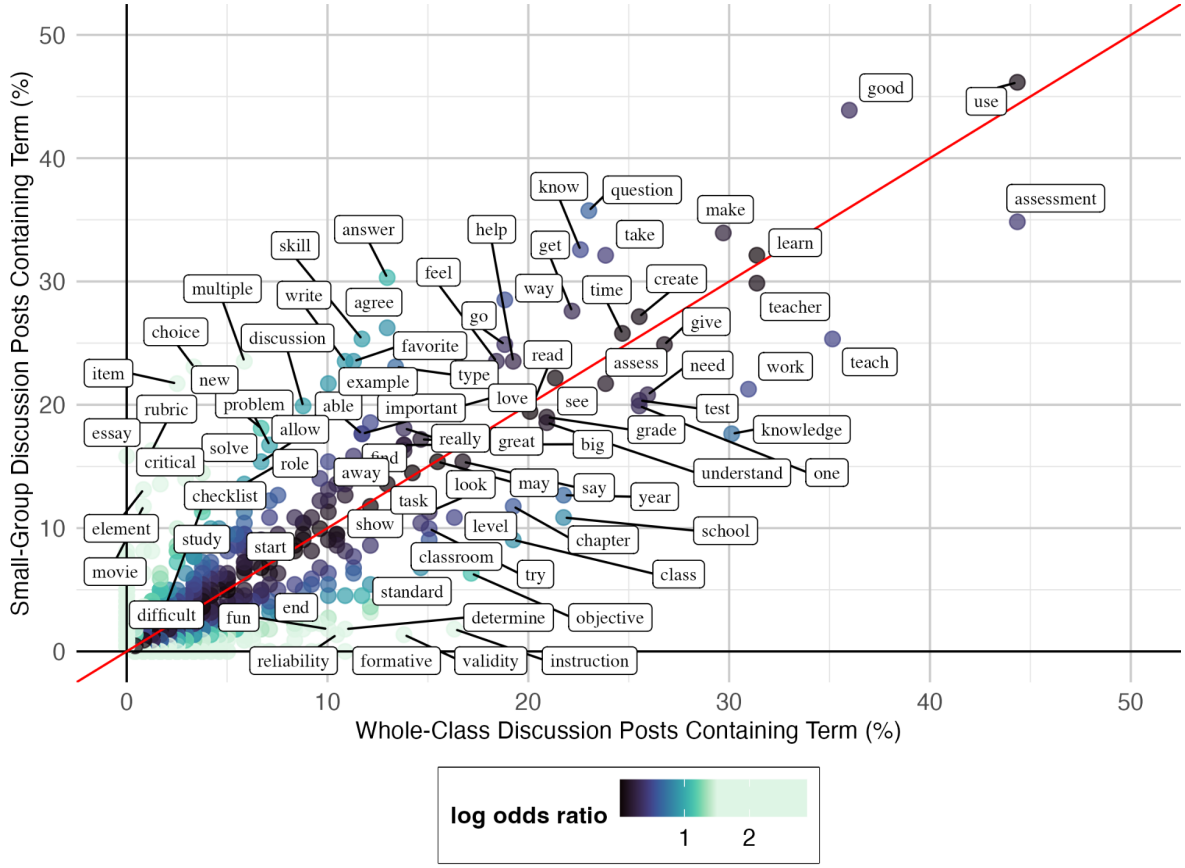
RQ2. Content of Social Interactions

To answer the second research question, we analyzed the linguistic differences of discussions in whole-class versus small-group format. First, we investigated the frequency of words used in whole-class compared to small-group discussions (Figure 2). In the 460 discussion posts and responses across 8 weeks, students

used 3,069 distinct words. We observed a higher likelihood of analytical and content-focused terms in whole-class discussions, such as *assessment*, *knowledge*, *objective*, *validity*, and *reliability*, as indicated by these terms appearing below the diagonal red line (i.e., terms plotted on the red line are equally likely to be used in both discussion formats). In contrast, in small groups, there was a higher likelihood of words that indicate social engagement, such as *question, answer, agree, favorite*, *discussion*, *skill*, *problem*, and *solve*.

Figure 2

Comparison of Terms Used in Whole-Class vs. Small-Group Discussions



Note: Proximity to the red line indicates greater likelihood that the term would have been used similarly in both types of discussion formats.

Second, we examined the differences between whole-class versus small-group discussions by using LIWC to calculate the median scores of linguistic features, and we evaluated the statistical significance of these differences (Table 3). Although we observe more analytical terms in whole-class discussions in Figure 2, the LIWC scores for analytical thinking were not statistically significantly different between discussion formats (p = .355). LIWC analysis also did not show differences across discussion formats between clout, cognition, culture, lifestyle, physical, or perception. Still, some differences were evident, as whole-class discussions exhibited statistically significantly higher LIWC measures of big words and authentic; small-group

discussions showed statistically significantly higher LIWC measures of word count, words per sentence, dictionary words, emotional tone, linguistic, drives, affect, social processes, and conversational. These results align with the observations of word frequencies in Figure 2, particularly that small-group discussions seem to be characterized by more social engagement (as indicated by social processes), positivity (as indicated by emotional tone), and emotional processes (as indicated by affect).

Table 3

Median Scores of LIWC Features and Results of Mann-Whitney U Tests (Wilcoxon Rank Sum)

Feature	Whole-class *Mdn*	Small-group *Mdn*	*W*	*p*
Word count	64	75	29,359.5	.038
Words per sentence	15.62	16.50	29,793.0	.018
Big words	23.53	21.92	20,916.5	< .001
Dictionary words	88.96	90.99	32,693.0	< .001
Analytical thinking	49.68	47.98	25,090.5	.355
Clout	40.06	42.97	29,036.0	.065
Authentic	68.87	65.87	23,323.5	.030
Emotional tone	46.50	61.50	31,877.5	< .001
Linguistic	70.47	72.28	31,362.5	< .001
Drives	3.90	4.85	32,304.5	< .001
Cognition	17.91	17.62	23,807.0	.068
Affect	3.04	4.00	33,291.0	< .001
Social processes	11.30	12.18	29,900.0	.014
Culture	0.00	0.00	27,664.5	.269
Lifestyle	7.14	6.45	26,403.5	.997
Physical	0.00	0.00	26,158.0	.825
Perception	5.88	6.38	28,104.5	.234
Conversational	0.00	0.00	28,945.0	.044

Discussion

The purpose of this study was to investigate how the format of asynchronous online class discussions (whole-class vs. small-group) affected the structure and content of student interactions. Although previous research has suggested that small-group discussions lead to increased student perceptions of social presence and sociability (Akcaoglu & Lee, 2016; Michaelsen et al., 2023), it was unclear whether different discussion formats would also lead to differences in the structure of social interactions (as measured by SNA) or the content of social interactions (as measured by LIWC assessment of linguistic features).

Findings from this study align with previous research that has shown the benefits of asynchronous online discussions for students, regardless of format (e.g., Yang et al., 2020). Still, our results do show observable

differences between whole-class and small-group formats in the structure of social interactions and statistically significant differences in the content of social interactions. Specifically, we found that whole-class discussions seemed to prioritize more formal and analytical communication, which is consistent with past research showing that whole-class discussions are better suited for tasks that require critical thinking and the synthesis of complex ideas (Cui & Teo, 2023). We also found that whole-class discussions had a higher median and notably higher maximum number of weekly posts, as well as a higher response rate. These differences may suggest an enthusiasm generated by a larger audience, which aligns with past research demonstrating the creative and divergent thinking afforded by large discussion groups (Yang et al., 2020) and participants' intent to appeal to a variety of perspectives, as well as, perhaps, attempts to impress their peers (Cui & Teo, 2023).

The differences in the structure and content of social interactions we observed in small-group discussions suggest a complementary set of benefits to those in whole-class discussions. For instance, higher reciprocity, transitivity, and closeness centrality scores as well as LIWC measures of social processes and positive emotional tone highlight the role that smaller conversation clusters can play in social presence: fostering closer connections, decreasing concerns about group members not contributing, and reducing feelings of isolation in online learning (Luo et al., 2023; Zhan et al., 2022). These benefits may help build trust, social skills, and a sense of community, which are critical for keeping learners engaged in asynchronous environments (Lim, 2023; Richardson et al., 2017). Our findings align with past research suggesting that small-group discussions encourage participants to express their ideas in more detail, promoting deeper engagement and thoughtful discussion (Akcaoglu & Lee, 2016; Qui & McDougall, 2015; Yang et al., 2020; Zhan et al., 2022) while offering a more comfortable social space for shy and reserved students to participate in discussions (Qiu et al., 2014).

Implications for Practice

The findings in this study demonstrate the complementary strengths of whole-class and small-group formats for online discussions, highlighting the importance of instructional design for shaping interaction patterns and optimizing collaborative learning environments (Cheung et al., 2021). Whole-class discussions can generate enthusiasm, fostering creativity (Yang et al., 2020) and critical thinking (Cui & Teo, 2023). Meanwhile, small groups can compensate for some of the perceived weaknesses of online teaching, such as limited social presence and students' feelings of isolation, leading to higher dropout rates and lower satisfaction compared to traditional learning contexts (Phirangee & Malec, 2017; Sung & Mayer, 2012).

Instructional designers and online instructors should consider incorporating both small-group and whole-class discussions into their courses, while writing learning objectives that make clear their expectations for social interactions and align with the strengths of the associated discussion format. For example, whole-class discussions encourage deeper creativity and critical thinking, and small-group discussions foster more reciprocity and support. Dividing students into small groups for discussion could create space for students to focus on mutual understanding and completing shared goals; this, in turn, could disrupt unhelpful discussion behaviors such as completing original posts and comments at the last minute. Meanwhile, whole-class discussions could be designed to stimulate and explore complex ideas.

Educators could apply the computational methods from this study to track and update social interactions in online courses. Building on Kreijns et al.'s (2022) conceptualization of social presence as a key factor for social interactions, we used SNA to investigate the structure of social interactions (i.e., social space) and LIWC to assess the content of social interactions (i.e., sociality). A learning analytics dashboard could automatically and continuously report SNA and LIWC metrics for students' interactions, providing instructors with opportunities to make adjustments during the course as needed and potentially increasing students' engagement as well (cf. Ramaswami et al., 2023). For instance, instructors could make adjustments to the discussion format if they notice trends in the content of social interactions, such as shifts in LIWC scores for social processes, emotional tone, drives, or cognition. In addition, SNA metrics would help instructors gain information on students' interaction patterns and be prepared to intervene, such as identifying isolated students and inviting them into small-group discussions or encouraging other students to respond to their posts.

Limitations and Future Research

Although this study extends previous research (Akcaoglu & Lee, 2016), there remain limitations. First, the dataset is limited to a single course and is dated, as the course discussions were collected in fall 2014. Although the instructor (a co-author on this paper) has maintained the overall design and discussion prompts for this course across the past decade, it is possible that student engagement and interactions may have changed over the years. Future research should follow a similar method for a more current comparison and offer perspectives from a different course at a different type of university, in a different national context for higher education outside the United States, or at the undergraduate level. The methods from the present study, specifically SNA and LIWC metrics, would scale well for analysis of much larger courses or multiple contexts. However, these computational methods are limited by collecting data only through observations; future studies would benefit from validating observational data with self-reported data from students and instructors.

Second, our approach only evaluates social interaction in terms of structure and content. Discussions could have been shaped by how the activity prompts were presented and facilitated. Thus, our findings may be affected by course settings and other contextual factors in addition to the discussion format. It was beyond the scope of the current study to control for confounding variables; however, future studies should investigate the order of offering discussion formats (e.g., small groups before whole-class discussions) as well as other factors, such as course subjects and national context. Similarly, participants' personality features may have had an effect. For instance, shy students might have talked more in small-group discussions, feeling comfortable with fewer people. Future research should consider controlling for personality characteristics and psychological state during statistical analysis.

Conclusion

More students than ever are taking online courses, despite critiques of social presence in online learning. The findings in this study align with past research, providing evidence of the positive benefits of online asynchronous discussions. Here, we found different but complementary strengths of whole-class and small-group discussion formats. Extending Kreijns et al. (2022), we measured social interactions as social presence in conjunction with social space and sociality, using computational methods of SNA to analyze

interaction structure (social space) and LIWC to assess interaction content (sociality). Our study serves as a proof-of-concept for a quantitative approach to computational text analysis that can be updated and scaled up to investigate very large courses or compare multiple contexts. As opportunities for online learning expand and the affordances of educational technologies increase, the structure and content of social interactions will remain key considerations for understanding the various factors that influence success in online learning.

References

Adnan, M., & Anwar, K. (2020). Online learning amid the COVID-19 pandemic: Students' perspectives. *Journal of Pedagogical Sociology and Psychology, 2*(1), 45–51. https://doi.org/10.33902/JPSP.2020261309

Akcaoglu, M., & Lee, E. (2016). Increasing social presence in online learning through small group discussions. *The International Review of Research in Open and Distributed Learning*, *17*(3), 1–17. https://doi.org/10.19173/irrodl.v17i3.2293

Boyd, R. L., Ashokkumar, A., Seraj, S., & Pennebaker, J. W. (2022). *The development and psychometric properties of LIWC-22*. University of Texas at Austin. https://www.liwc.app/static/documents/LIWC-22%20Manual%20-%20Development%20and%20Psychometrics.pdf

Cheung, S. K. S., Kwok, L. F., Phusavat, K., & Yang, H. H. (2021). Shaping the future learning environments with smart elements: Challenges and opportunities. *International Journal of Educational Technology in Higher Education, 18*, Article 16. https://doi.org/10.1186/s41239-021-00254-1

Cui, R., & Teo, P. (2023). Thinking through talk: Using dialogue to develop students' critical thinking. *Teaching and Teacher Education, 125*, Article 104068. https://doi.org/10.1016/j.tate.2023.104068

Dado, M., & Bodemer, D. (2017). A review of methodological applications of social network analysis in computer-supported collaborative learning. *Educational Research Review, 22*, 159–180. https://doi.org/10.1016/j.edurev.2017.08.005

Elaish, M. M., Shuib, L., Hwang, G.-J., Ghani, N. A., Yadegaridehkordi, E., & Zainuddin, S. Z. (2023). Mobile English language learning: A systematic review of group size, duration, and assessment methods. *Computer Assisted Language Learning, 36*(3), 430–456. https://doi.org/10.1080/09588221.2021.1931341

Gao, J., Wei, S., & Dubé, A. K. (2024). Evaluating learners' online learning experience of informal learning environments: A LIWC analysis. *The Journal of Applied Instructional Design*, *13*(2), 63–67. https://doi.org/10.59668/1269.15691

Garrison, R. (2000). Theoretical challenges for distance education in the 21st century: A shift from structural to transactional issues. *The International Review of Research in Open and Distributed Learning*, *1*(1), 1–17. https://doi.org/10.19173/irrodl.v1i1.2

Gasell, C., Lowenthal, P. R., Uribe-Flórez, L. J., & Ching, Y.-H. (2022). Interaction in asynchronous discussion boards: A campus-wide analysis to better understand regular and substantive

interaction. *Education and Information Technologies*, *27*(3), 3421–3445. https://doi.org/10.1007/s10639-021-10745-3

Gunawardena, C. N. (1995). Social presence theory and implications for interaction and collaborative learning in computer conferences. *International Journal of Educational Telecommunications, 1*(2), 147–166. https://www.learntechlib.org/primary/p/15156/

Ho, H. C. Y., Poon, K.-T., Chan, K. K. S., Cheung, S. K., Datu, J. A. D., & Tse, C. Y. A. (2023). Promoting preservice teachers' psychological and pedagogical competencies for online learning and teaching: The TEACH program. *Computers & Education, 195*, Article 104725. https://doi.org/10.1016/j.compedu.2023.104725

Hollander, M., Wolfe, D. A., & Chicken, E. (2014). *Nonparametric statistical methods* (3rd ed.). Wiley.

Joksimović, S., Gašević, D., Kovanovíc, 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

Kim, J., Kwon, Y., & Cho, D. (2011). Investigating factors that influence social presence and learning outcomes in distance higher education. *Computers & Education, 57*(2), 1512–1520. https://doi.org/10.1016/j.compedu.2011.02.005

Kreijns, K., Xu, K., & Weidlich, J. (2022). Social presence: Conceptualization and measurement. *Educational Psychology Review, 34*(1), 139–170. https://doi.org/10.1007/s10648-021-09623-8

Lim, J. (2023). Exploring the relationships between interaction measures and learning outcomes through social network analysis: The mediating role of social presence. *International Journal of Educational Technology in Higher Education, 20,* Article 14. https://doi.org/10.1186/s41239-023-00384-8

Lowenthal, P. R., & Dunlap, J. C. (2020). Social presence and online discussions: A mixed method investigation. *Distance Education, 41*(4), 490–514. https://doi.org/10.1080/01587919.2020.1821603

Lowenthal, P. R., Greear, K., Humphrey, M., Lowenthal, A., Conley, Q., Giacumo, L. A., & Dunlap, J. C. (2020). Creating accessible and inclusive online learning: Moving beyond compliance and broadening the discussion. *Quarterly Review of Distance Education, 21*(2). https://www.infoagepub.com/products/Quarterly-Review-of-Distance-Education-21-2

Luo, H., Chen, Y., Chen, T., Koszalka, T. A., & Feng, Q. (2023). Impact of role assignment and group size on asynchronous online discussion: An experimental study. *Computers & Education, 192*, Article 104658. https://doi.org/10.1016/j.compedu.2022.104658

Michaelsen, L. K., Knight, A. B., & Fink, L. D. (2023). *Team-based learning: A transformative use of small groups in college teaching*. Taylor & Francis.

National Center for Education Statistics. (2015, December). *Number and percentage of students enrolled in degree-granting postsecondary institutions, by distance education participation, location of student, level of enrollment, and control and level of institution: Fall 2013 and fall 2014*. Digest of Education Statistics. https://nces.ed.gov/programs/digest/d15/tables/dt15_311.15.asp

National Center for Education Statistics. (2023, December). *Number and percentage of students enrolled in degree-granting postsecondary institutions, by distance education participation, location of student, level of enrollment, and control and level of institution: Fall 2021 and fall 2022*. Digest of Education Statistics. https://nces.ed.gov/programs/digest/d23/tables/dt23_311.15.asp

Norz, L.-M., Dornauer, V., Hackl, W. O., & Ammenwerth, E. (2023). Measuring social presence in online-based learning: An exploratory path analysis using log data and social network analysis. *The Internet and Higher Education, 56*, Article 100894. https://doi.org/10.1016/j.iheduc.2022.100894

Oh, H., Liu, Z., & Xing, W. (2025). Do actions speak louder than words? Unveiling linguistic patterns in online learning communities using cross recurrence quantification analysis. In A. Martínez Monés, C. Mills, J. Jovanovic, & X. Ochoa (Chairs), *The 15th International Conference on Learning Analytics and Knowledge* (pp. 992–998). ACM. https://doi.org/10.1145/3706468.3706569

Palau, R., Fuentes, M., Mogas, J., & Cebrián, G. (2021). Analysis of the implementation of teaching and learning processes at Catalan schools during the COVID-19 lockdown. *Technology, Pedagogy and Education, 30*(1), 183–199. https://doi.org/10.1080/1475939X.2020.1863855

Pennebaker, J. W., Chung, C. K., Ireland, M., Gonzales, A., & Booth, R. J. (2007). *The development and psychometric properties of LIWC2007*. University of Texas at Austin. https://liwc.app/static/documents/LIWC2007%20Manual%20-%20Development%20and%20Psychometrics.pdf

Pennebaker, J. W., Boyd, R. L., Booth, R. J., Ashakkumar, A. & Francis, M. E. (2022). *Linguistic inquiry and word count: LIWC-22* [Computer software]. Pennebaker Conglomerates. https://www.liwc.app

Phirangee, K., & Malec, A. (2017). Othering in online learning: An examination of social presence, identity, and sense of community. *Distance Education, 38*, 160–172. https://doi.org/10.1080/01587919.2017.1322457

Qiu, M., Hewitt, J., & Brett, C. (2014). Influence of group configuration on online discourse writing. *Computers & Education, 71*, 289–302. https://doi.org/10.1016/j.compedu.2013.09.010

Qiu, M., & McDougall, D. (2015). Influence of group configuration on online discourse reading. *Computers & Education, 87*, 151–165. https://doi.org/10.1016/j.compedu.2015.04.006

Ramaswami, G., Susnjak, T., & Mathrani, A. (2023). Effectiveness of a learning analytics dashboard for increasing student engagement levels. *Journal of Learning Analytics, 10*(3), 115–134. https://doi.org/10.18608/jla.2023.7935

Richardson, J. C., Maeda, Y., Lv, J., & Caskurlu, S. (2017). Social presence in relation to students' satisfaction and learning in the online environment: A meta-analysis. *Computers in Human Behavior, 71*, 402–417. https://doi.org/10.1016/j.chb.2017.02.001

Rourke, L., Anderson, T., Garrison, D. R., & Archer, W. (1999). Assessing social presence in asynchronous text-based computer conferencing. *Journal of Distance Education, 14*(2), 50–71. https://www.learntechlib.org/p/92000/

Salter, N. P., & Conneely, M. R. (2015). Structured and unstructured discussion forums as tools for student engagement. *Computers in Human Behavior, 46*, 18–25. https://doi.org/10.1016/j.chb.2014.12.037

Satar, H. M., & Akcan, S. (2018). Pre-service EFL teachers' online participation, interaction, and social presence. *Language Learning & Technology, 22*(1), 157–183. https://doi.org/10.64152/10125/44586

Scholand, A. J., Tausczik, Y. R., & Pennebaker, J. W. (2010, February). Social language network analysis. In K. Inkpen & C. Gutwin (Chairs), *The 2010 ACM conference on computer-supported cooperative work* (pp. 23–26). ACM. https://doi.org/10.1145/1718918.1718925

Short, J., Williams, E., & Christie, B. (1976). *The social psychology of telecommunications*. John Wiley & Sons.

Staudt Willet, K. B., & Carpenter, J. P. (2020). Teachers on Reddit? Exploring contributions and interactions in four teaching-related subreddits. *Journal of Research on Technology in Education, 52*(2), 216–233. https://doi.org/10.1080/15391523.2020.1722978

Steponavičius, M., Gress-Wright, C., & Linzarini, A. (2023). *Social and emotional skills: Latest evidence on teachability and impact on life outcomes* (OECD Education Working Paper No. 304). OECD. https://doi.org/10.1787/ba34f086-en

Sung, E., & Mayer, R. E. (2012). Five facets of social presence in online distance education. *Computers in Human Behavior*, *28*(5), 1738–1747. https://doi.org/10.1016/j.chb.2012.04.014

Tabassum, S., Pereira, F. S., Fernandes, S., & Gama, J. (2018). Social network analysis: An overview. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8*(5), Article e1256. https://doi.org/10.1002/widm.1256

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology, 29*(1), 24–54. https://doi.org/10.1177/0261927X09351676

Van Wart, M., Ni, A., Medina, P., Canelon, J., Kordrostami, M., Zhang, J., & Liu, Y. (2020). Integrating students' perspectives about online learning: A hierarchy of factors. *International Journal of Educational Technology in Higher Education, 17*, Article 53. https://doi.org/10.1186/s41239-020-00229-8

Wickersham, L. E., & Dooley, K. E. (2006). A content analysis of critical thinking skills as an indicator of quality of online discussion in virtual learning communities. *Quarterly Review of Distance Education*, *7*(2), 185–193. https://doi.org/10.1108/QRDE-06-2006-0007

Yang, T., Luo, H., & Sun, D. (2022). Investigating the combined effects of group size and group composition in online discussion. *Active Learning in Higher Education, 23*(2), 115–128. https://doi.org/10.1177/1469787420938524

Ye, D., & Pennisi, S. (2022). Analysing interactions in online discussions through social network analysis. *Journal of Computer Assisted Learning, 38*(3), 784–796. https://doi.org/10.1111/jcal.12648

Zhan, Z., He, G., Li, T., He, L., & Xiang, S. (2022). Effect of groups size on students' learning achievement, motivation, cognitive load, collaborative problem-solving quality, and in-class interaction in an introductory AI course. *Journal of Computer Assisted Learning, 38*(6), 1807–1818. https://doi.org/10.1111/jcal.12722