

Journal of Educational Technology Development and Exchange (JETDE)

Volume 14 | Issue 2

2021

Characteristics of Initial Posts and Peer Engagement: Density Score Analyses for Social Presence in Online Discussions

Kathy Hanselman
kathrynh@unr.edu

Leping Liu
University of Nevada, Reno, liu@unr.edu

Follow this and additional works at: <https://aquila.usm.edu/jetde>



Part of the [Educational Technology Commons](#)

Recommended Citation

Hanselman, Kathy and Liu, Leping (2021) "Characteristics of Initial Posts and Peer Engagement: Density Score Analyses for Social Presence in Online Discussions," *Journal of Educational Technology Development and Exchange (JETDE)*: Vol. 14 : Iss. 2 , Article 3.

DOI: 10.18785/jetde.1402.03

Available at: <https://aquila.usm.edu/jetde/vol14/iss2/3>

This Article is brought to you for free and open access by The Aquila Digital Community. It has been accepted for inclusion in Journal of Educational Technology Development and Exchange (JETDE) by an authorized editor of The Aquila Digital Community. For more information, please contact Joshua.Cromwell@usm.edu.

Characteristics of Initial Posts and Peer Engagement: Density Score Analyses for Social Presence in Online Discussions

Kathryn Hanselman

University of Nevada

Leping Liu

University of Nevada

Abstract: *A common activity in online courses that allows for student interaction is the asynchronous discussion; however, discussions do not inherently lead to meaningful engagement among students. This study explores how the moves that students make in their initial discussion posts influence the emotional engagement of their peers in response posts. 1500 asynchronous online discussion messages were collected from an undergraduate online course offered at a western state university. 608 online discussion threads were analyzed to determine how the characteristics of initial posts are associated with the engagement in peer responses. Six characteristic variables from initial posts were identified and analyzed. Density scores for social presence categories and indicators were calculated as the measure of the emotional engagement in the response posts. Results suggest that three characteristic variables in initial posts significantly influence the emotional engagement of peers in the response posts.*

Keywords: online discussion, student engagement, social presence, interpersonal communication, open communication, cohesive communication, intrarater reliability

1. Introduction

Online learning is growing rapidly in higher education as institutions try to meet the needs of a student population desiring more options for course and degree offerings (Seaman et al., 2018). While more online classes are being offered, effective course design has grown into a field of research and application in education. Methods of encouraging student engagement have become an important consideration in online course design (Martin, 2019).

Student engagement involves learners actively working within the course and its activities to build understanding of the course content (Hu & Kuh, 2002). Online courses present challenges to engagement, as students may feel isolated from their peers, their instructors, or the course, which may lead them to lose interest or drop the course (Martin, 2019). Student interaction in the online environment can help students to feel involved with the course content and with the learning community, which helps to make learning meaningful (Russell et al., 2009). Creating opportunities for students to connect with one another and course content can lead to meaningful interaction and engagement in pursuit of knowledge building.

The Community of Inquiry (CoI) framework provides a lens to view the ways students engage when they construct knowledge as part of a learning community. In CoI, students within a social environment negotiate meaning and build understanding as a community of learners working together. Garrison et al. (1999) applied this framework to online learning, exploring the ways that a community develops in terms of its members' cognitive presence, social presence, and teaching presence. Each presence contributes to a participant's engagement within the

community. Investigating one or more of these presences can provide a window into the ways that students are engaging with content and one another within the community as they build knowledge in their online courses.

To promote student interaction and engagement, instructors and course designers often include online discussions in which students are asked to connect with course content by answering provided prompts and to engage with peers by reading and responding to the peers' posts. Online discussions make possible student-to-student interaction, providing students a location where they can collaboratively construct knowledge, and "...the opportunity to share ideas, learn from peers and build knowledge collectively, while reading and reflecting on each other's thoughts" (Kent et al., 2016, p. 117). Discussions can be seen as a location where students participate in and contribute to a learning community, which can help to foster engagement and deepen student learning.

Online discussions provide an opportunity for students to demonstrate social presence, and construct understanding of course concepts. Students demonstrate social presence in how they present themselves as people within the community, and how they communicate and build interpersonal relationships with others, which helps them to engage with peers and succeed in the course (Garrison, 2009; Tu & McIssac, 2002). Studying indicators of social presence found in online discussions can provide insight into how students engage with peers and the course.

Social presence can indicate student engagement as students interact as part of a learning community bent on building knowledge as a group. Since these aspects can help point to student engagement, it is beneficial to know what contributes to their

development in online discussions. Previous research has focused on how student actions within their learning communities may demonstrate their own engagement (Garrison, 2009), but less attention has been paid to how student actions may influence the actions of their peers. Within a learning community, no student is an island; members of the community work together to build knowledge, and thus rely on the input of their peers.

As such, looking to the actions of students in their initial discussion posts that influence the responses they receive can shed light on how students promote engagement in their peers, specifically in terms of social presence demonstrated. Characteristics of initial posts that may influence this engagement include when the initial post is made (Zingaro & Oztok, 2012), the length of the initial post (Ho & Swan, 2007), how easy it is to read (Zingaro & Oztok, 2012), the use of first- or second-person pronouns (Carroll, 2007), and the cognitive presence demonstrated (Garrison et al., 1999). Each of these traits present in an initial post may subconsciously or consciously influence the peer reader, which may lead to decisions made as to whether and how to respond. A focus on how the work of one student influences the engagement of another may explain how the ties that bind the community are strengthened, and how students may or may not be contributing to the success of their peers.

The overarching purpose of this correlational study is to determine the relationship between initial discussion posts and response post engagement for undergraduate students in online asynchronous discussions, specifically demonstrated in terms of social presence in response posts. In other words, how do the moves that students make in their initial discussion posts affect the response posts that their peers may provide?

Unlike previous research that focuses on engagement of students within the community of learners in terms of their own participation, this study sought to understand how student engagement is related to the actions of their peers.

2. Review of Literature

This literature review provides background information and pertinent research in the area of online learning, student engagement, the Community of Inquiry theory, online discussions as a location of student interaction and knowledge construction, and the variables present in student discussion posts that may influence interaction and engagement of their peers.

2.1. Online Learning

Online learning has experienced rapid growth in the twenty-first century, as colleges and universities attempt to better serve students' eager for more options of course offerings. Even as face-to-face enrollments have begun to decrease, distance education, which encompasses online learning, had increased for the past fourteen years as of 2018 (Seaman et al., 2018). The percentage of all U.S. post-secondary students who enrolled in at least one online course increased from 31.1% in 2016 to 33.1% in 2017 (Allen & Seaman, 2017). Further, university administrations saw online offerings as critical to their long-term strategy, with over 77% agreeing in both 2014 and 2015 (Allen & Seaman, 2016). This was before the COVID-19 pandemic, which prompted most institutions of higher education to rapidly transition to much larger online course offerings than in previous terms.

This increase in online course offerings has brought about an increase in research

in the field, as instructors, instructional designers, and administrators seek to create online courses committed to the success of students. However, even with carefully planned courses, students were not guaranteed success (Ali & Smith, 2015). Online students dropped out of their classes and programs at higher rates than students enrolled in face-to-face classes and programs (Schaeffer & Konetes, 2010). Schaeffer and Konetes (2010) found that a commonly cited factor for student dissatisfaction in their online courses was a feeling of social isolation from their peers, their teachers, and the course in general. One approach to overcome these feelings of isolation and disconnection that online students may feel is to focus on engagement to draw students into the course and learning community.

2.2. Student Engagement

Engagement is one of the most important factors in student learning. Engagement refers to how students actively work within course components and activities in an effort to develop knowledge and understanding of the course content (Hu & Kuh, 2002). This active involvement can lead to student achievement of learning outcomes and cognitive development (Ma et al., 2015).

The concept of student engagement in learning can be viewed as a multidimensional construct that includes the three areas of behavioral, emotional, and cognitive engagement (Fredricks et al., 2004). Behavioral engagement hinges on a student's level of participation, both academic and social. Emotional engagement is based on students' reactions to the class, their peers, the content, and their instructors, and may include interest, perceived value, boredom, or anger, among other emotions. Cognitive engagement is measured by students' effort in building

knowledge or understanding (Fredricks et al., 2004). All three of these areas of engagement may work in concert to promote overall student engagement in learning.

Online courses present unique challenges related to engagement. Online students may feel isolated from their peers and the course and may lose interest (Martin, 2019). Dumford and Miller (2018) found that online courses promoted more engagement in terms of some activities, but face-to-face courses were more likely to promote collaborative learning and quality interactions with others; they posited that the self-directed nature of many online classes may be isolating and may not provide enough opportunities for collaborative learning. Physical distance between students learning online may lessen the feeling of being part of a learning community and contribute to attrition rates (Rovai, 2002). These challenges may lead to students losing interest or disengaging from their online classes, which may lead to incomplete assignments, lower grades, or withdrawal from the course.

In the online classroom, lack of opportunity to meet in person may make students feel removed or distant from their classmates. By focusing on engaging students, instructors and designers can build courses that help to close this distance. Such engagement can help to contribute to a learning community in which students work together to construct knowledge, feel more involved with the course, and become more connected with their peers. The Community of Inquiry framework introduces a lens through which to view the ways students engage with one another, course content, and their instructors in such learning communities.

2.3. Community of Inquiry

The Community of Inquiry (CoI) framework is a theory of learning that helps

educators better approach the question of how students learn as part of a group. In this framework, learners negotiate meaning and work in a social environment toward understanding, reaching understanding together. This framework was applied to the realm of online learning by Garrison et al. (1999), who further defined the framework as one arising from participants' cognitive presence, social presence, and teaching presence.

Cognitive presence is the manner by which learners within the community create meaning in collaboration with their fellow community members. Social presence measures how members project themselves to identify themselves within the community, and to communicate and develop relationships with members of the group. The third aspect, teaching presence, may fall to peers within the community or, more likely, to the instructor guiding the learning activity; this includes the design and presentation of content and the facilitation of the learning activity (Garrison et al., 1999).

The CoI framework sets the scene for higher order thinking through collaboration. Lipman (2003) put forth a collaborative environment in which students were encouraged to participate in "questioning, reasoning, connecting, deliberating, challenging, and developing problem-solving techniques" (p. 14). These actions enabled students to negotiate their understanding within a group, by detecting problems or errors in understanding and challenging accepted ideas, which has led to deeper learning for participants (Ramsden, 1988). Not only did they provide their own thoughts and ideas, but they would also need to listen to the ideas of their peers, challenging when necessary, and adjusting their own understanding when the explanations or ideas of others were more feasible than their own. They were

not just waiting to provide their own ideas, but interacting with others to build deeper understanding (Gardner, 1999).

Over more than twenty years of research, the CoI framework has provided a foundation for the investigation of how students develop knowledge together, at the intersection of cognitive presence, social presence, and teaching presence. Each dimension plays an important role in how students engage with one another and learn. This study focuses on social presence.

2.4. Social Presence

Social presence dictates how learners feel as part of the CoI which may affect their openness to learning. Garrison (2009) defines social presence as "the ability of participants to identify with the community (e.g., course of study), communicate purposefully in a trusting environment, and develop inter-personal relationships by way of projecting their individual personalities" (p. 352). It is related to the concept of "immediacy" or behaviors that enhance closeness and interaction (Rourke et al., 2001). Many behaviors that enhance closeness, such as facial expression or eye contact, are not available in a computer mediated environment. However, the indicators of social presence are examples of ways participants still project themselves socially and emotionally in an online course. It should also be noted that social presence does not focus entirely on engagement based on social interaction: it instead focuses on the moves that are made to support an environment that welcomes questions and the contribution of ideas from members of the community (Garrison & Aykol, 2010)

The three dimensions of social presence relate to how students identify with the learning community, how they communicate within the community, and how they

develop interpersonal relationships with others in the community (Garrison, 2009). These dimensions of social presence have been used to develop an instrument to code indicators of social presence, based on a literature review undertaken by Garrison et al. (1999). These indicators can be divided into three categories: (a) actions that express interpersonal communication, (b) actions that express open communication, and (c) actions that express group cohesion. Interpersonal communication includes actions that may help initiate a community by promoting emotional connection between participants. This may include use of emoticons or unusual grammar/punctuation (e.g., LOL), self-disclosure of personal information, and use of humor. These moves may indicate students attempting to build trust within the community. Open communication includes actions that contribute to discourse between members of the community, in terms of interaction and communication. Students must use open communication to interact in a way to successfully work together, which may include asking one another questions, prompting input, negotiating the meaning presented in members' contributions, etc.

Cohesive communication includes actions that contribute to students identifying and acting as part of the community. Communication aimed at group cohesion within a discussion board may include posts that address members of the discussion group using inclusive pronouns, or the use of phatics or salutations to help build familiarity within the group (Garrison, 2011).

Each category of social presence describes a way that students demonstrate interaction and engagement in the way they communicate with one another as part of a learning community. Rourke et al. (2001) developed a coding model for use in analyzing social presence in terms of behavioral indices and indicators of social interaction in communicative responses, and Garrison (2011) refined this coding model. The classification model was developed by researchers undertaking a theoretical analysis of previous research and coding discussion transcripts, resulting in three categories of social presence that were further described by between three and five indicators per category (Garrison, 2011). The categories, their indicators, and their definitions are outlined in Table 1.

Table 1. A Coding Model for Assessment of Social Presence Based on Garrison (2011)

Category	Label	Definition
Interpersonal communication	Affective expression	Conventional expressions of emotion, or unconventional expression of emotion, include repetitious punctuations, conspicuous capitalization, emoticons
	Self-disclosure	Presents details of life outside of class, or expresses vulnerability
	Use of humor	Teasing, cajoling, irony, understatements, sarcasm

Open communication	Asking questions	Students ask questions of other students or the moderator
	Referring explicitly to others' messages	Direct references to the contents of others' posts
	Complimenting, expressing appreciation, agreement	Complimenting others or contents of others' messages Expressing agreement with others or content of others' messages
Cohesive communication	Vocatives	Addressing or referring to participants by name
	Addresses or refers to the group using inclusive pronouns	Addresses the group as we, us, our, group
	Phatics or salutations	Communication that serves a purely social function; greetings, closures

Two indicators from the Social Presence model were removed: “continuing threads” and “quoting from others’ messages” are technical aspects of the discussion forum, which do not constitute decisions made by students in terms of their interactions with peers (Kovanović et al., 2014; Lee, 2014).

2.5. Emotional Engagement: Social Presence in Online Discussions

Student engagement serves as one of the goals in all learning, especially for online learning, where students may feel disconnected from peers and the class or learning community as a whole. Fredericks, Blumenfeld, and Paris’ (2004) three-component model for student engagement is based on behavioral, emotional, and cognitive engagement concepts. This study focuses on emotional engagement, as demonstrated by social presence detected in response posts.

Emotional engagement is based on

students’ reactions to the class, their peers, the content, and their instructors, and may include interest, perceived value, boredom, or anger, among other emotions (Fredericks et al., 2004). Further, it can be viewed as identification with or belonging to a learning community (Fredericks et al., 2004). In other words, emotional engagement centers on how students affectively interact with learning, their learning environment, or their learning community. Emotional engagement then can be seen as related to social presence, which has been defined as “the ability of participants to identify with the community (e.g., course of study), communicate purposefully in a trusting environment, and develop inter-personal relationships by way of projecting their individual personalities” (Garrison, 2009, p. 352), or, more succinctly, as “degree to which participants feel affectively connected to one another” (Kozan & Richardson, 2014, p. 69).

Social presence helps students to feel more like members of their learning

community, helping them to engage with their peers and succeed in the course (Tu & McIssac, 2002). Research suggests that student perception of social presence can be related to interaction, satisfaction, and learning, all of which play into a student's feeling of engagement and connectedness with their community of learners. Wei and Chen (2012) noted that student perception of social presence may affect how comfortable they were in interacting with peers. Doo and Bonk (2020) explained the ability of social presence to reduce students' feelings of isolation in online learning, which may lead to a shift in student preference towards more active and engaging interactions within their learning community. Further, Swan and Shih (2005) noted that students who perceived high levels of presence in their peers also themselves projected high levels of social presence, specifically in the areas of self-disclosure, building toward community, and building on their peers' ideas.

Engagement reflected in the higher levels of social presence tends to affect student learning. Cobb (2009) found social presence to be a key component to student perception of quality in an online learning experience. Swan and Shih (2005) found that students reporting high social presence believed they learned more than their peers who perceived lower levels of social presence. Social presence may also indicate a greater level of engagement based on student satisfaction. Bulu (2012) noted that students were more satisfied with their online learning experience when they perceived greater social presence. Swan (2002) found that student perception of social presence accounted for 35% of their overall satisfaction with their course. This engagement may lead to a higher likelihood of students persisting in their online learning courses. Boston et al. (2009) analyzed over

28,000 student records and surveys and found that indicators of social presence accounted for a significant variance in student re-enrollment in future terms within an online program.

Collaborative student interaction in the online classroom can help students to feel engaged with the course content and with the community of learners, which helps to make learning meaningful to students. By creating a place for students to interact and collaborate with peers, instructors can encourage students to emotionally engage with the course in terms of social presence. One popular activity in online learning is the asynchronous online discussion, which invites students to interact with peers as they grapple with course concepts, working to build understating together. These online discussions provide a prime location to view the emotional engagement of students.

2.6. Online Discussions

Online discussions are activities in which students communicate and build knowledge with one another in an online space. The asynchronous nature of discussions means that students are not required to be logged on at the same time, and instead can leave messages for one another to read or view at each student's convenience. The purpose of online discussions is to encourage communication between students, where they can explain their experience or knowledge, and interact with others to construct greater understanding (Garrison, 2011). Vai and Sosulski (2016) describe a common format for online discussions as follows: The instructor posts a prompt for the entire class or smaller groups within the class, then students reply to the instructor's prompt with their answers, and finally students read and respond to their classmates' answers to the prompt.

Students tend to find more structured

discussions to be more meaningful (Jacobi, 2017).

While time between posting and responding is found to be helpful (Jacobi, 2017), the number of responses to peers may cause problems: drawing from real-world discussion forums, if too many posts populate the discussion board, students have been found to simplify their posts, or not post at all (Jones et al., 2004). Instructors must find a balance between encouraging as much interaction as possible, without asking so much of students that they give up. Additionally, while structure is appreciated, getting too specific on characteristics such as post length and required citations might lead to less discussion (Gilbert & Dabbagh, 2005). With student interaction as a goal of online discussion, the instructor must consider not just the quantity and timing of student responses to one another, but also the quality. Many discussions include directions for peer response, disallowing short and somewhat meaningless responses such as “I agree” or “Great post!” Providing specific structure and content guidelines is recommended but finding the right balance of requirements is necessary to promote meaningful student interaction and engagement.

Well-designed discussions can promote learning within the community, as students encounter problems in their knowledge and work with others in their learning environment to solve those problems and construct new understanding (Vai & Sosulski, 2016). This design contributes to the achievement of the overarching goals of online discussions, in support of the development of learning communities.

Students must communicate and interact to successfully work together, which may include asking one another questions, prompting input, negotiating the meaning

presented in members’ contributions, etc. Examining student posts for indicators of social presence can provide some level of understanding of how students are emotionally engaging with the discussion, the class, and the learning community.

2.7. Discussion Post Characteristics

The ways that students act and interact in online discussion posts may provide some insight into peer-promoted engagement. Generally, a student’s initial post is their own individual thoughts or ideas in relation to the instructor-posted prompt; a response post requires that students read the thoughts and ideas of their peers then formulate a response that is in at least some way related to the peer’s post. As such, characteristics of the initial post may influence peer responses. The following traits of initial discussion posts were investigated in this study.

2.7.1. Post Time from Due Date

One important initial post characteristic that may influence responses from classmates is the time when it is posted. Specific due dates may be in place for initial posts separate from response posts: students may be directed to post their initial message earlier in the week, then return to the discussion board later in the week to read and reply to peers’ posts. Students can only read and reply to posts that exist at the point at which they are ready to begin making replies. This means that messages posted earlier in the week, when there are fewer messages available, may generate more response than posts made later in the week, when there are many more messages available (Pena-Schaff & Nicholls, 2004; Zingaro & Oztok, 2012). Hewitt (2003) found that students tended to reply to newer messages rather than older ones, meaning those initial posts that had been made closer

to the time when the student returned to the discussion board to post their own replies. Blanchette (2011) found that discussion interaction generally took place as a shorter period of concentrated activity, which was then followed by little activity or a complete fall off of responses.

2.7.2. Word Count

Another characteristic of an online discussion posting that may influence peer response is its length. While some instructors include a minimum length for discussion posts, there is still much variation in how much students choose to write. Zingaro and Oztok (2012) hypothesized that longer messages included more ideas that peers could respond to than shorter messages. On the other hand, Ho and Swan (2007) argued that long posts may contain too many ideas for a simple response. Ho and Swan (2007) found that students were more likely to respond to shorter posts, possibly due to the time and patience required to get through the message to reply. Abe (2020) noted that word count was a predictor of overall academic success and hypothesized that this may be due to a level of conscientiousness: students who spent more time studying for quizzes and completing other work were also likely to put more effort into their discussion posts.

2.7.3. Reading Ease

Another initial post trait that may influence peer response is how easy it is for the post to be read. A post that is easier for students to read may attract more responses; posts that students perceive as harder to read may be bypassed in favor of those that they believe get the point across more easily. The Flesch Reading Ease Readability score indicates how difficult a piece of writing is to

read in English, based on averages of sentence length and word length. The Flesch Reading Ease Readability provides a text with a score between 1-100: higher scores indicate that the message is easier to read, while lower scores indicate more difficult. Scores between 90-100 indicate that a passage is very easy, scores between 80-89 indicate that it is easy, scores between 70-79 indicate that it is fairly easy, scores between 60-69 indicate that it is standard, scores between 50-59 indicate that it is fairly difficult, scores between 30-49 indicate that it is difficult, and scores between 0-29 indicate that it is very confusing (Readability Formats Website, 2014).

The Flesch Reading Ease Readability score is calculated with sentence length (average of length of sentence in words) and word length (average number of syllables per word). The specific mathematical formula is as shown in Equation 1 (Readability Formats Website, 2014):

$$RE = 206.835 - (1.015 \times ASL) - (84.6 \times ASW)$$

(Equation 1)

Where RE stands for *Readability Ease*, ASL for *Average Sentence Length* (i.e., the number of words divided by the number of sentences), and ASW for *Average number of Syllables per Word* (i.e., the number of syllables divided by the number of words).

Zingaro and Oztok (2012) suggest that messages that are easier to read may more easily facilitate communication and connection between students, making the argument that courses high in social presence may contain more messages that are easier to read.

2.7.4. First- and Second-Person Language

Another characteristic of initial posts that may influence peer responses is the use of first- and second-person language within the

body of the message. First-person language includes “I,” “me”, and “my”. Second-person language includes “you” and “your.” The use of pronouns can indicate the quality of personal relationships, by showing how people are referred to both within and outside of an interaction (Tausczik & Pennebaker, 2010). Furthermore, related to cognitive theory of multimedia learning, personalization of teaching materials can lead to increases in student interest, which then leads to the student putting forth more cognitive effort, which may result in deeper learning (Mayer et al., 2004).

The use of first- and second-person language may be affected by a student’s experience in college-level writing. As the use of pronouns is less common in academic writing, Carroll (2007) hypothesized that a decline in their usage may be impacted by students moving away from the use of personal experiences and instead relying on researched evidence to help make their arguments. Formality of writing in initial posts may affect their peers’ desire to respond, as it may contribute to the feeling of distance between students and respondents (Tu & McIsaac, 2002). Mayer et al. (2004) also suggested that the use of self as a reference point increased the learner’s interest, which then helped to open the learner up to further cognitive processing.

The use of first-person plural pronouns such as “we”, “us”, or “our” may indicate a measure of group identity, and in fact the use of such pronouns is included in the CoI theoretical framework for the social presence category of cohesive communication.

2.7.5. Initial Post Cognitive Presence

Cognitive presence is defined as the extent to which students within a learning community are able to construct knowledge based

on communication with peers within that community (Garrison et al., 1999). Cognitive presence outlines the learner’s meaning making through learning scenarios comprised of four stages. The learning scenario begins with the triggering event, or the point at which learners feel a sense of unease or discomfort regarding an idea or concept. The exploration phase follows, wherein learners seek additional or alternate information on the concept. In the integration phase, learners integrate the information in with their previous understanding into a new concept. Finally, in the resolution stage, learners resolve the issue and overcome the problematic understanding from the first phase (Garrison et al., 1999). At this point, new understanding or knowledge is reached. The level of cognitive presence reached in an initial post may influence peer responses.

How students express themselves in their communications can demonstrate their own social presences as members of their learning community. The characteristics of initial discussion posts explained above may impact those presences, both in their own communiques and in the responses that they prompt in their peers.

3. Methods

This study sought to investigate the impact of specific moves made by students in their initial online discussion posts on their peers’ responses. The following research question was used to guide through the study:

To what extent can the six characteristic variables in initial discussion posts (time from due date, word count, reading ease score, use of first-person pronouns, use of second-person pronouns, and cognitive presence level) be used to predict social presence in response posts?

A nonexperimental correlational research design was used to determine if characteristics present in student initial posts could be used to predict whether the post would stimulate social presence in the responses of peers. This research design was chosen because it attempts to explore possible correlational relationship between independent variables and a dependent variable on an occasion in which the researcher is unable to control the independent variable (Lobmier, 2010).

3.1. Participants and Sampling

For the purpose of this study, a dataset was obtained from two sections of a required undergraduate core course that took place online at a mid-sized public Research 1 university in a western state in 2020-2021. The course chosen for this study (Core Humanities 212: Science, Technology, and Society in the Modern Era) offered a focus on science and technology; participants tended to come from STEM-related majors and were at sophomore level. A total of 59 students were in these two sections.

In this study, individual students did not serve as the cases to be examined; rather, discussion threads served as the cases under analysis. This included each initial post and any subsequent responses to that initial post as a single case. A sample of around 1500 discussion post messages (including initial posts and response posts) was used in this study, representing all students enrolled in the two sections of the online course. 608 discussion threads were formed and served as the sample cases to be examined. In each case, the initial post was coded for six characteristic variables, and response messages to that initial post were coded for the density of social presence indicators. This is a nonrandom sample (or convenience sample), which is often used in educational research when

random sampling is not applicable or not possible (Fowler, 2002; Rovai et al., 2013). In the case of this study, the sample was a convenience sample readily available to the researchers from existing online courses.

3.2. Procedures

3.2.1. Course Setting

Courses were delivered fully online in a completely asynchronous format using the Canvas learning management system. The two sections followed the same course template including the syllabus, reading and lecture content, assignments, assessments, and discussions. Students earned three credit hours for successfully completing the course.

The course for this study was chosen because it is an option for a core curriculum requirement, and thus contains a cross section of the larger student population. With multiple sections following the same online discussion design and using the same discussion prompts, a large sample size was possible. Finally, access to the discussion data was readily available in the learning management system.

3.2.2. Discussion Requirements

Asynchronous discussions were required of all students in ten weeks of the twelve-week semester for both sections. Students were assigned into small discussion groups of four to five students each, with group membership changing each week. Students were provided five to nine options for pre-determined prompts to respond to, of which they had to choose two to answer in two separate discussion posts by 11:59 p.m. on Thursdays. Students then were required to review and reply to at least two classmates' posts in response to different prompts than they had answered themselves, on the discussion board by 11:59 p.m. on Sundays. Discussion

participation was worth 25 points each week (2% of the class total grade), or 250 points overall (25% of the class total grade). Initial posts were worth 7.5 of the points each, and each response post was worth 5 points.

3.2.3. Data Collection

Data collection was approved by the Institutional Review Board (IRB) of the university. Data was collected from online sections after the conclusion of the semesters included in the study, when student grades were already assigned, and students no longer had access to the course sites. All discussion posts were downloaded from the Canvas learning management system using the “Canvas-Discussions-Get_Entries” userscript developed by Dr. Brian Reid (Canvas, 2020). The Tampermonkey for Chrome browser add-on was used to access the userscript. Once installed, the userscript was utilized by entering the course site on Canvas, navigating to the Discussions index page, and clicking the “Userscript: Get Discussion Entries” button at the top or bottom of the page, and selecting “Generate one file with responses to topics” option.

A Microsoft Excel file was created that contained all posts from all discussion forums within the class. The Excel sheet contained information for each discussion post, including the course identification number, the topic (discussion forum) identification number, the topic (discussion forum) title, the discussion type (in this case, all discussions were threaded), entry (or post) identification number, the initial post author, the initial post, the initial post word count, the reply post identification number, the reply post author, the reply post, and the reply post word count. Non-content-based discussions (such as introductions, icebreakers, and extra-credit activities) were removed. Once the Excel spreadsheets were prepared in this manner,

data coding took place.

3.3. Measurement and Instrument

In this study, initial posts and response posts were analyzed and their characteristics measured and coded. Initial discussion posts were coded for the amount of time before a due date a post was made; word count; reading ease; the use of first-person language such as “I”, “me,” and “my”; the use of second person language such as “you” or “your”; and the highest level of cognitive presence achieved in the initial post. Response posts were coded for the density of social presence categories and indicators.

3.3.1. Content Analysis

A content analysis of individual student discussion posts was undertaken to translate discussion data into quantitative form for data analysis. Rourke et al. (2001) recommend choosing a unit of analysis “that multiple coders can identify reliably, and simultaneously, one that exhaustively and exclusively encompasses the sought-after construct” (p. 17). Options for unit of analysis include sentence, paragraph, message, or thematic levels, although the more granular units may make it difficult for a coder to identify the full intent of the statement (Garrison et al., 2009). Individual discussion posts, or the entire message created by a student, were chosen as the unit of analysis for this study because they were easily demarcated from one another, and present the author’s decisions related to content and length (Garrison et al., 2001b).

Each student discussion post underwent content analysis, and the features in the written communication of the post was converted into numeric variables that could be statistically analyzed. The variables were first identified from each post, and then measured and coded

accordingly with the methods as described in the following measurement for each variable.

3.3.2. Density Score Calculation

To quantify and measure the discussion data and characteristic variables, density scores were used. Density scores are calculated by taking the count of instances of each variable within each post, dividing that count by the total number of words in the post, and multiplying by 1000 (Rourke et al., 2001). This yields a unit of incidents per 1,000 words. Using density scores makes it possible to examine or compare the characteristic variables between posts, as the value of the count may have been skewed by the differences in length of posts. Equation 2 details the calculation for density scores.

$$\text{Density score} = \frac{\text{instances}}{\text{post word count}} \times 1000$$

(Equation 2)

For example, for the social presence indicator “Open Communication: Asking questions,” if a student asked three separate questions in a response of 158 words, this would count as three instances, and the calculation would be three divided by 158, multiplied by 1000. The resulted density score would be 18.99.

This study treated a discussion thread as a single case. In a case that includes one initial post and more than one response (i.e., three responses), the calculation of density score for the social presence indicator “Open Communication: Asking questions” from the three responses will undertake two steps: (a) calculating the density score of this indicator for each of the three responses using the method described above, and (b) averaging the three scores. The resulted average score will serve as the value of this indicator for that

particular case.

Density score have been used to calculate social presence in many studies (Lowenthal & Dunlap, 2020; Rourke et al., 2001; Swan & Shih, 2005); they can also be applied to any variable for which individual counts are sought to be compared amongst cases that may present at varying lengths. While there is no evidence of using density scores in analyzing first- or second-person language in the literature, it was employed in this study to measure and examine these two variables presented in posts of varying lengths.

3.3.3. Initial Post Coding

All initial discussion posts were coded for the following variables. A description of each variable and how it was coded is provided below:

Post Time from Due Date. In this study, “post time from due date” is the time measured in minutes counting from the point at which a student posts a message to the time of the required due date time. When an initial post was made in relation to the stated due date may influence responses from peers. If a post is made early in the week, well before the due date, then it has more time to be viewed by other students. Earlier in the week, there are fewer posts for students to choose from to respond to, meaning those initial posts may get more replies than initial posts made closer to the due date (Pena-Schaff & Nicholls, 2004; Zingaro & Oztok, 2012). It may also be possible that students prefer to respond to more recently made posts (Blanchette, 2011; Hewitt, 2003).

Discussion post times were tracked by the learning management system. The post time from due date was calculated using the post

time obtained using the “Canvas-Discussions-Get_Entries” userscript. Each discussion forum has specific deadlines for initial posts (Thursdays by 11:59 p.m.) and response posts (Sundays by 11:59 p.m.). The time that a post was made was subtracted from the specific deadline for that week’s discussion and converted into minutes. Posts made after the deadline were given a negative value. The “post time from due date” variable is continuous data.

Word Count. In this study, “word count” is the number of words in a discussion post. The number of words in an initial post may influence peer responses. The length of the initial post may draw more or fewer responses, based on how peers react to the initial post: if it is long, peers may decide it will take too much time to read and respond to; on the flipside, they may be attracted to a longer post that has more content or ideas that they may base their response on (Zingaro & Oztok, 2012). A shorter post may be more attractive because it can be read and reflected upon more quickly (Ho & Swan, 2007), or it may not contain enough information for the responding student to connect with. The word count of the initial post was included as a predictor variable in this study because it may provide information on how the length of a message influences its responses: this may provide instructors with evidence to set a minimum or maximum word count for discussion assignments.

The word count for each initial post was tabulated by Canvas and accessed using the “Canvas-Discussions-Get_Entries” userscript. The word unit is counted as a group of letters with spaces on either side of it. This may lead to minor errors, if students combine words without a space between them, as might be found with a typo or misspelling. The “word

count” variable is continuous data.

Reading Ease. In this study, “reading ease” is the Fleshe Reading Ease score for the discussion post, which is calculated using sentence length and word length (see Equation 1) to indicate how difficult the reading may be. How easy an initial post is to read and comprehend may influence peer responses. Students are required to read through and reflect on the points made in their peers’ initial discussion posts; as such, the readability of initial posts may influence who chooses to read and respond to them. Those that use a simpler writing structure may communicate their ideas in a way that readers can more easily digest, which may in turn promote further interactive communication (Zingaro & Oztok, 2012).

The value for the reading ease variable was calculated using the built-in spelling and grammar checking tool in Microsoft Word. The discussion post was copied to the clipboard, then pasted into a blank Microsoft Word document. Under the Review menu, the Spelling and Grammar tool was opened. Once all spelling and grammar errors were reviewed, a “Readability Statistics” window opened that provided information on counts, averages, and readability scores for the open document. The Flesche Reading Ease score was found here and entered into the data tracking spreadsheet. The “reading ease” variable is continuous data.

First- or Second-Person Language. In this study, “first-person language” and “second-person language” are separate variables that were measured and coded with density scores calculated from the count of each variable within a discussion message. The use of first- or second-person language in an initial post may also influence the

responses made by peers. First-person singular language includes “I”, “me,” or “my”, and second-person language were calculated by counting the instances of the words “you,” or “your” within the discussion post. The use of personal pronouns can help place the writer as a person within the community, as well as demonstrate how the student views or relates to peers. Further, the use of personal pronouns may lead to more informal writing, which may influence the social distance felt by peers (Tu & McIsaac, 2002). In the courses under investigation here, no directions were provided as to the use of personal pronouns, so students could choose for themselves whether to use them or to remain in the more impersonal third-person.

First-person language was calculated separately from second-person language. First-person language was calculated by counting the instances of the words “I”, “me,” or “my” within the discussion post; second-person language was calculated by counting “you” and “your”. Values for first-person and second-person language were calculated as density scores, using the density score calculation explained in Equation 2. This created a percentage that is more easily comparable between posts, as the value of the count may have been skewed by the differences in length of posts. The “first-person language” and “second-person language” variables were continuous data.

Cognitive Presence. In this study, “cognitive presence” was measured and coded according to the highest level of cognitive presence a message reached. Each initial post was reviewed and coded for cognitive presence. Cognitive presence outlines the learner’s meaning making through learning scenarios comprised of four stages. The

learning scenario begins with the triggering event, or the point at which learners feel a sense of unease or discomfort regarding an idea or concept. The exploration phase follows, wherein learners seek additional or alternate information on the concept. In the integration phase, learners integrate the information in with their previous understanding into a new concept. Finally, in the resolution stage, learners resolve the issue and overcome the problematic understanding from the first phase (Garrison et al., 1999). At this point, new understanding or knowledge is reached. The level of cognitive presence reached in an initial post may influence peer responses.

Content analysis of the initial discussion posts was undertaken using the four categories of cognitive presence (triggering, exploration, integration, and resolution). The instrument for coding indicators of cognitive presence was developed using Garrison et al.’s (2001b) Indicators of Cognitive Presence. The coding model is divided into four categories (one for each level of cognitive presence), and provides information about the sociocognitive processes behind each indicator label.

Some messages may include evidence of more than one cognitive presence event, as learners move through the phases of cognitive processing. For the purpose of coding, only one phase was counted for each case. As outlined by Garrison et al. (2001b), each case was “coded up” to the highest level of cognitive presence achieved in the message if multiple levels are found to be present in a single message. This means that only one score for cognitive presence was given per post, and that score described the highest level achieved.

Table 2. Coding Model for Assessment of Cognitive Presence

Phase	Code
No cognitive presence detected	0
Triggering event	1
Exploration event	2
Integration	3
Resolution	4

The “cognitive presence” variable is categorical data. Table 2 shows the coding for each phrase of cognitive presence total.

3.3.4. Response Post Coding: Social Presence

In response posts, “social presence” was measured and coded with density scores calculated from the count of each social presence category, indicator, and a social presence total within a discussion message. Each response post was reviewed and coded for social presence. Social presence dictates how learners feel and act as part of the CoI and may affect their openness to learning. Indicators of social presence can be divided into three categories: actions that express interpersonal communication, actions that express open communication, and actions that express cohesive communication. Actions that express interpersonal communication are those that promote emotional connection between members of the community. Actions that express open communication are those that may affect student relation and trust of one another, and by extension trust of one another’s contributions. Actions that help promote group cohesion are those that may affect group member’s commitment to

meaning making and the unified acceptance of understanding (Garrison, 2011).

Content analysis of the response discussion posts was undertaken using the three categories of social presence. The instrument for coding indicators of social presence was developed using Garrison’s (2011) Indicators of Social Presence, which is displayed in Table 1, and provides a coding model for assessing social presence for this study. The coding model is divided into three categories containing three indicator labels each and provides definitions for each indicator label.

Each discussion response post was coded by the count of nine social presence indicators, from the three social presence categories. A count was given for each indicator, then the counts for each indicator within a category were added together to determine the category count. Finally, all category counts were combined to determine the social presence total. Values for each indicator, category, and the social presence total were calculated using the density score calculation (see Equation 2).

Table 3 shows the coding for each indicator, category, and social presence total.

Table 3. Social Presence Variables and Measurement

Social Presence Event
Interpersonal Communication: Expression of emotions
Interpersonal Communication: Use of humor
Interpersonal Communication: Self-disclosure
Open Communication: Referring explicitly to others' messages
Open Communication: Asking questions
Open Communication: Complimenting, expressing appreciation, expressing agreement
Cohesive Communication: Vocatives
Cohesive Communication: Addresses or refers to the group using inclusive pronouns
Cohesive Communication: Phatics or salutations
Interpersonal Communication total
Open Communication total
Cohesive Communication total
Social Presence total
Measurement
All indicators were coded according to the density score per message.

Each response post's coded information related to its correspondent initial post and was categorized as a characteristic of the initial post for the sake of data organization. If an initial post had more than one response post, the social presence density scores for *each* indicator, category, and the social presence total for all responses to that initial post were first calculated individually for each post, then averaged. The initial post data included one score that encompassed the average of all of its responses.

4. Data Analyses

Data analyses of this study were undertaken in three phases. First, intrarater reliability analysis was performed during the

period of data coding. Second, to prepare for the main multiple regression analyses, routine pre-analysis data screening took place including screening data for missing data and outliers, and assumptions for multiple regression analysis were also conducted. Finally, a set of multiple regression analyses to answer the research question were performed. IBM SPSS Statistics, version 27 for Mac was used for all statistical calculations. For all analyses in this study, an α level of .05 was used to indicate the threshold probability that was acceptable for this study.

4.1. Intrarater Reliability Analysis

The coding for the initial post cognitive presence and response post social presence

indicators was first completed using all data collected. After two months, the same researcher revisited the data to check the reliability of the initial round of coding. For both the initial posts and response posts, 15% of the total number of posts were randomly selected and recoded (Landis & Koch, 1977). This resulted in 155 initial posts and 147 response posts being recoded.

An intrarater reliability analysis using the Kappa statistic was conducted to determine the agreement of the coding results for the

variables between the two coding periods. The reliability test was done with the initial post cognitive presence and response post social presence indicators. Table 4 shows the intrarater reliabilities for the coding periods regarding these variables. Values of Kappa between .40 and .59 are considered moderate, between .60 and .79 are considered substantial, and above .80 are considered outstanding (Landis & Koch, 1977). Based on this guideline, the levels of agreement between the two coding periods regarding the variables were generally very good.

Table 4. Levels of Agreement between Coding Periods

Variable	Kappa Coefficient	p-value
Initial post cognitive presence	.692	<.001
Interpersonal: Expression of emotions	.783	<.001
Interpersonal: Use of humor	.588	<.001
Interpersonal: Self-disclosure	.689	<.001
Open: Referring explicitly to others' messages	.756	<.001
Open: Asking questions	.865	<.001
Open: Complimenting, expressing appreciation, expressing agreement	.742	<.001
Cohesive: Vocatives	.950	<.001
Cohesive: Addresses or refers to the group using inclusive pronouns	.854	<.001
Cohesive: Phatics or salutations	.981	<.001

4.2. Preliminary Data Screening and Assumption Checking

Routine pre-analysis data screening took place before the regression analyses began, including screening data for missing data and outliers. All data was reviewed for missing data; since the data was compiled from the available original data source, any variables missing data were tracked down and entered.

The original data set included 630 cases.

After screening for missing data, this data set was screened for outliers. Extreme values on the predictor variables were examined for outliers, which were then deleted from the sample. Any cases with a .458 Mahalanobis value exceeded the chi-square criterion ($\chi^2_3 = 22.458$ at $p = .001$) were eliminated from the sample; 22 cases were removed from the data set, leaving 608 cases to be analyzed.

Assumptions for multiple regression analysis were also evaluated. The regression

coefficients for each variable of each multiple regression analysis indicate that multicollinearity was not violated because tolerance statistics for all six independent variables were greater than .1. Linearity, homoscedasticity, and normality were also investigated, but were minorly violated. Moderate violations of these assumptions do not invalidate the regression, but may weaken the analysis (Tabachnick & Fidell, 2007). Additionally, moderate violations of normality will not negatively affect the regression (Tate, 1992 as cited in Vanatta Reinhart & Mertler, 2016).

4.3. Multiple Regression Analyses

To answer the research question, multiple regression analyses were conducted where the six variables coded from the initial posts served as the independent variable or predictor variables (X_n), and the social presence variables including the social presence indicators, categories, and social presence total that were coded from the response posts served as the dependent variable or criterion variable (\hat{Y}_n). Again, the predictor variables (X_{1-6}) are:

X_1 = Word count (exact count of words in initial posting)

X_2 = Time from due date (exact count of minutes from due date of initial posting)

X_3 = Reading ease score

X_4 = First-person language (density score)

X_5 = Second-person language (density score)

X_6 = Cognitive presence (dummy variable)

The criterion variables are the social presence total (\hat{Y}_1), the three categories ($\hat{Y}_{2,3,4}$) of social presence, and social presence indicators under each category:

\hat{Y}_1 = Social presence communication total

\hat{Y}_2 = Interpersonal communication total

\hat{Y}_{2a} = *Interpersonal: Expression of emotions*

\hat{Y}_{2b} = *Interpersonal: Use of humor*

\hat{Y}_{2c} = *Interpersonal: Self-disclosure*

\hat{Y}_3 = Open communication total

\hat{Y}_{3a} = *Open: Referring explicitly to others' messages*

\hat{Y}_{3b} = *Open: Asking questions*

\hat{Y}_{3c} = *Open: Complimenting, expressing appreciation or agreement*

\hat{Y}_4 = Cohesive communication total

\hat{Y}_{4a} = *Cohesive: Vocatives*

\hat{Y}_{4b} = *Cohesive: Addresses or refers to the group using inclusive pronouns*

\hat{Y}_{4c} = *Cohesive: Phatics or salutations*

In a set of multiple regression analyses, the six predictor variables were used to predict each of the criterion variables, and the results are reported in the next section.

5. Results

In the multiple regression analyses, the dependent variable (criterion variable) is the social presence density detected in response posts and the independent variables (predictor variables) are the initial post characteristics including time from due date, word count, reading ease score, first-person language density score, second-person language density score, and initial post cognitive presence score. In this section, we report the results from the multiple regression analyses undertaken for (a) social presence overall, (b) each of the three social presence categories, and (c) those social presence indicators that obtained significant models and predictors. In total, seven results are presented.

5.1. Social Presence Total

Result 1. Multiple regression was

conducted to determine how well each or any of the six predictor variables was associated with the Social Presence Total (\hat{Y}_1) detected from response posts. The results showed that the linear model was not significant ($R^2 = .007$, $F_{(1,607)} = .662$, $p = .680$), indicating that the six characteristic variables measured in initial posts were not significantly contributing to the variance of the overall social presence measured in the response posts. Further analyses were undertaken on each of the social presence categories.

5.2. Interpersonal Communication Total and Indicators

Result 2. Multiple regression was conducted to determine the accuracy of the predictor variables influencing the overall Interpersonal Communication (\hat{Y}_2) detected in response posts. Regression results showed that the linear model was significant ($R^2 = .016$, $F_{(1,607)} = 9.944$, $p = .002$), indicating that 1.6% of variance in Interpersonal Communication in response posts was associated with the linear model.

Table 5. Coefficients for Final Model Predicting Interpersonal Communication (\hat{Y}_2)

Variable	β	t	p	Partial r
Post time from due date	.065	1.617	.106	.066
Word count	-.011	-.272	.785	-.011
Reading ease	-.005	-.117	.907	-.005
First-person language	.073	3.153	.002	.127
Second-person language	.007	.180	.857	.007
Initial post cognitive presence	.026	.639	.523	.026
Constant	3.455	6.919	<.001	

A summary of regression coefficients is presented in Table 5. One (first-person language) of the six variables significantly contributes to the model. First-person language in initial posts significantly contributes to the interpersonal communication density score in response posts ($\beta = .073, t = 3.153, p = .002$). Interpersonal Communication density score in responses increases .073 units for each unit of increase in first-person language density in initial posts. Partial r (.111) is the partial correlation between first-person language and Interpersonal Communication, while controlling for the influence of the other independent variables. The 95% confidence interval for β is between .028 and .119.

Result 3. As the regression model significantly predicted the Interpersonal Communication category overall, further multiple regression analyses were undertaken for each of its indicators to determine if any of the individual indicator presence in the responses could be predicted from a linear combination of the initial post characteristics. Of the regression models, only the model for

the indicator Interpersonal Communication: Self-disclosure (\hat{Y}_{2c}) measured in response posts was significant.

Regression results showed that the overall model was significant ($R^2 = .015, F_{(1,607)} = 9.403, p = .002$), and 1.5% of variance in Interpersonal Communication: Self-disclosure in response posts was associated with the linear model. A summary of regression coefficients is presented in Table 6. One (first-person language) of the six variables significantly contributes to the model. First-person language in initial posts significantly contributes to the Interpersonal Communication: Self-disclosure score in response posts ($\beta = .050, t = 3.066, p = .002$). Interpersonal Communication: Self-disclosure density score increases .050 units for each unit of increase in first-person language density. Partial r (.124) is the partial correlation between first-person language and Interpersonal Communication: Self-disclosure, while controlling for the influence of the other independent variables. The 95% confidence interval for β is between .018 and .081.

Table 6. Coefficients for Final Model Predicting Interpersonal Communication: Self-disclosure (\hat{Y}_{2c})

Variable	β	t	p	Partial r
Post time from due date	.034	.840	.401	.034
Word count	.015	.373	.709	.015
Reading ease	-.049	-1.164	.245	-.047
First-person language	.050	3.066	.002	.124
Second-person language	.016	.389	.697	.016
Initial post cognitive presence	.034	.841	.401	.034
Constant	1.860	5.360	<.001	

5.3. Open Communication Total and Indicators

Result 4. Multiple regression was conducted to determine the accuracy of the

independent variables predicting the Open Communication (\hat{Y}_3) detected in response posts. Regression results showed that the linear model was not significant ($R^2 = .004, F_{(1,607)} = .419, p = .867$), indicating that the

six characteristic variables measured in initial posts were not significantly associated with the Open Communication Total measured in the response posts.

Further multiple regression analyses were undertaken to explore whether or which of the Open Communication category's indicators measured in response posts are associated with a linear combination of the six characteristic variables from the initial posts.

Result 5. Of the regression models for the three Open Communication indicators, one was found significantly associated with the indicator Open Communication: Referring explicitly to others' messages (\hat{Y}_{3a}). The linear model was significant ($R^2 = .009$, $F_{(1,607)} = 5.238$, $p = .022$), and .9% of variance in Open Communication: Referring explicitly to others' messages in response posts was

associated with the linear model. A summary of regression coefficients is presented in Table 7. One (reading ease) of the six variables significantly contributes to the model. Reading ease in initial posts significantly contributes to the Open Communication: Referring explicitly to others' messages score in response posts ($\beta = -.041$, $t = -2.2892$, $p = .022$). Open Communication: Referring explicitly to others' messages density score decreases .041 units for each unit of increase in reading ease score. That is, the more difficult to read the initial post is, the less engagement in the response post for referring explicitly to others' messages. Partial r (-.092) is the partial correlation between reading ease and Open Communication: Referring explicitly to others' messages, while controlling for the influence of the other independent variables. The 95% confidence interval for β is between -.077 and -.006.

Table 7. Coefficients for Final Model predicting Open Communication: Referring Explicitly to Others' Messages (\hat{Y}_{3a})

Variable	β	t	p	Partial r
Post time from due date	.004	.087	.931	.004
Word count	.023	.581	.561	.024
Reading ease	-.041	-2.289	.022	-.092
First-person language	-.041	-.971	.332	-.039
Second-person language	-.001	-.017	.986	.001
Initial post cognitive presence	.041	1.102	.312	.041
Constant	6.380	6.512	<.001	

5.4. Cohesive Communication Total and Indicators

Result 6. Multiple regression was conducted to determine the accuracy of the independent variables predicting the Cohesive Communication (\hat{Y}_4) detected in response

posts. Regression results showed that the overall model was significant ($R^2 = .012$, $F_{(1,607)} = 7.213$, $p = .007$), and 1.2% of variance in Cohesive Communication in response posts was associated with the linear model. A summary of regression coefficients is presented in Table 8. One (word count) of the

six characteristic variables from initial posts significantly contributes to the model. Word count in initial posts significantly contributes to the Cohesive Communication score in response posts ($\beta = -.013$, $t = -2.686$, $p = .007$). Cohesive Communication density score decreases .013 units for each unit of increase

in word count. Partial r (-.108) is the partial correlation between word count and Cohesive Communication, while controlling for the influence of the other independent variables. The 95% confidence interval for β is between -.022 and -.003.

Table 8. Coefficients for Final Model Predicting Cohesive Communication (\hat{Y}_4)

Variable	β	t	p	Partial r
Post time from due date	-.051	-1.242	.215	-.050
Word count	-.013	-2.686	.007	-.108
Reading ease	-.037	-.909	.363	-.037
First-person language	-.013	-.329	.742	-.013
Second-person language	.024	.588	.557	.024
Initial post cognitive presence	-.020	-.478	.633	-.019
Constant	21.620	20.017	<.001	

As the regression model significantly predicted the Cohesive Communication category overall, further multiple regression analyses were undertaken for each of its indicators to determine if the individual indicator presences in the responses could be predicted from a linear combination of the initial post characteristics.

Result 7. Of the regression models for the three Cohesive Communication indicators, one was found significantly associated with the presence of the indicator Cohesive Communication: Phatics or salutations (\hat{Y}_{4c}). Regression results showed that the overall model was significant ($R^2 = .018$, $F_{(1,607)} = 11.048$, $p = <.001$), and 1.8% of variance in Cohesive Communication: Phatics or

salutations in response posts was associated with the linear model. A summary of regression coefficients is presented in Table 9. One (word count) of the six variables significantly contributes to the model. Word count in initial posts significantly contributes to the Cohesive Communication: Phatics or salutations score in response posts ($\beta = -.009$, $t = 3.324$, $p < .001$). Cohesive Communication: Phatics or salutations density score decreases .009 units for each unit of increase in word count. Partial r (-.134) is the partial correlation between word count and Cohesive Communication: Phatics or salutations, while controlling for the influence of the other independent variables. The 95% confidence interval for β is between -.014 and -.003.

Table 9. Coefficients for Final Model Predicting Cohesive Communication: Phatics or salutations (\hat{Y}_{4c})

Variable	β	t	p	Partial r
Post time from due date	-.037	-.924	.356	-.037
Word count	-.009	-3.324	<.001	-.134
Reading ease	-.044	-1.093	.275	-.044
First-person language	-.023	-.580	.562	-.024
Second-person language	-.010	-.253	.800	-.010
Initial post cognitive presence	-.004	-.099	.921	-.004
Constant	11.201	19.172	<.001	

Results from the seven multiple regression analyses have revealed five significant association between three characteristic variables (first-person language, reading ease, and word count) measured in online discussion initial posts and four social presence

category/indicator variables (Interpersonal communication total, Interpersonal: self-disclosure, Open: Referring Explicitly, Cohesive communication total, and Cohesive: Phatics or salutations) measured from the response posts as summarized in Figure 1.

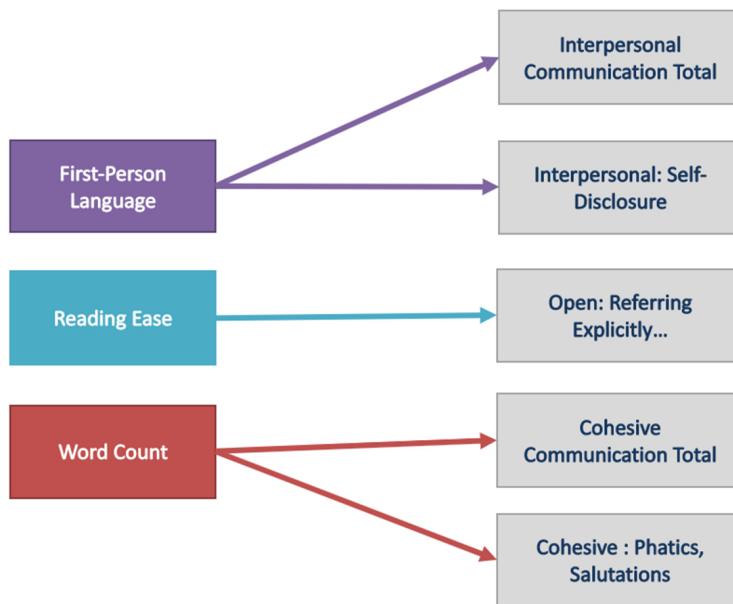


Figure 1. Summary of Significant Results

6. Discussion and Conclusion

The purpose of this research was to investigate the influence of student initial post features on the emotional engagement of their peers, which was demonstrated by the social presence detected in the responses to that initial post. Multiple linear regression analyses were conducted to develop prediction models that could predict the total social presence displayed in a response, as well as the totals for each of the social presence categories and their indicators.

6.1. Social Presence Total

The first multiple regression analysis sought to find an equation that could predict the overall social presence density of response posts based on initial post characteristics. Garrison (2009) defines social presence as “the ability of participants to identify with the community (e.g., course of study), communicate purposefully in a trusting environment, and develop inter-personal relationships by way of projecting their individual personalities” (p. 352). This concept includes the categories interpersonal communication, open communication, and cohesive communication that served as the dependent variables in further multiple regression analyses conducted for this research question.

No combination of the six independent variables could significantly predict the amount of social presence demonstrated in the correspondent response posts. Social presence as a whole dictates how learners feel as part of the CoI and may influence their openness to learning, and demonstrates ways participants project themselves socially and emotionally in an online course. As the three categories of social presence and their

individual indicators vary widely in the specific manner of demonstrating presence, it is possible that the overarching concept of social presence as a whole was too large to be influenced significantly by specific actions taken in initial posts. Little research exists that investigates how certain aspects of discussion posts made by one student may influence the social presence demonstrated by peers, so comparisons with the present results are difficult. It may be surprising that none of the independent variables were found to significantly predict social presence as a whole, but specific characteristics of initial posts were found to significantly predict the more defined categories and specific indicators that make up the social presence total. This may mean that the overall concept of social presence is too broad to be directly related to specific traits in initial discussion posts.

6.2. Interpersonal Communication

Of the six independent variables analyzed, only one initial post characteristic significantly contributed to the model predicting the amount of interpersonal communication present in a response post: first-person language. Interpersonal communication includes actions that may help initiate a community by promoting emotional connection between participants and may indicate students attempting to build trust within the community. This includes expressing emotion, sharing details of students’ life outside of class, and using humor to relate to peers. Peers may respond to the use of first-person pronouns in initial posts as an invitation to share more of themselves, in terms of emotion, self-disclosure, and humor. The use of personal pronouns can denote the quality of relationships built between people, as they are used to refer to people within and outside of the interaction (Tausczik & Pennebaker, 2010).

By using first-person language in initial posts, students may be sharing more of themselves with their peers; peers may respond to this sharing by reciprocating the same in their responses.

Deeper investigation into each of the three indicators of interpersonal communication found that first-person language also contributed significantly to a model predicting the indicator Interpersonal: Self-disclosure. Initial posts may include personal histories, examples, or opinions that peers will respond to in kind in their response posts; this would necessitate the use of first-person language in the discussion of these personal messages. Additionally, Mayer et al. (2004) note that “using the self as a reference point increases the learner’s interest,” which may draw peers in to want to share their own experiences. Finally, the idea that a student is opening themselves up in an initial post by presenting themselves using first-person language may encourage their peers to reciprocate by disclosing their own experiences as a way to relate or build relationships.

The other five initial post variables were not found to significantly contribute to a model predicting interpersonal communication in response posts. This could be due to various reasons for each variable. An initial post’s time from due date is temporal and does not focus on the content of the message. Interpersonal communication may stem more from how students present themselves in that initial message, rather than when they do the presenting. The word count of an initial post, too, may not be clearly related to how students are presenting themselves in their initial posts. Further, a post’s length could have varying effects: shorter posts may not give peers much to grab on to as they respond, with fewer opportunities to relate to one another on an interpersonal level. Longer posts may provide

too much information, making it harder for the peer to find a way to relate interpersonally. It is surprising that reading ease of an initial post had no influence on the interpersonal communication in a response; Zingaro and Oztok (2012) suggest that a post that is easier to read might promote connection more easily as meaning is made clear and students can effortlessly understand one another. In terms of second-person language, this could be seen as distancing on the part of the initial poster: rather than relating ideas or experiences to the self, the student may use “you” as a stand-in. This may lead peers to feel less connected to the initial post, and lead to less interpersonal communication in their responses. Finally, in terms of the variable initial post cognitive presence, it’s possible that the level of cognitive presence focuses more on the content of the post in terms of understanding and solving an issue and does not relate to the ways in which students speak to one another.

6.3. Open Communication

No combination of the six independent variables could significantly predict the total amount of open communication demonstrated in the correspondent response posts. Open communication demonstrates ways that students contribute to discourse between members of the community, in terms of interaction and communication, and may include asking questions of one another, asking for input, and negotiating meaning of the writing of their peers.

Deeper investigation into each of the three indicators of open communication found that a higher score in reading ease contributed significantly to a model predicting Open Communication: Referring explicitly to others’ messages. A higher score in reading ease means that the message was easier to read, which Zingaro and Oztok (2012) suggest

may better facilitate communication. By referring explicitly to parts of the initial post, a peer might be trying to point to specific areas they'd like to investigate further. Being able to clearly understand the idea set forth in the initial post makes it easier for a peer to be confident as they build on it in their response.

There are many possible explanations for why the initial post characteristics investigated were not found to significantly predict open communication. A post's time from due date focuses on when the post was made, not its content. Posts made earlier in the week or later in the week might still contain the same ideas, which may or may not encourage peers to respond in ways that contribute to discourse. For example, if a peer asks a question of the initial poster, this question will need to be asked regardless of when the post was made. A post's word count could influence responses in multiple ways: shorter posts may require peers to respond with questions for clarification, while longer posts may provide many points that the peer needs to negotiate meaning on. The use of personal pronouns may have had little influence here, as the positioning of the initial poster within their message mattered less than the content of the message itself in terms of garnering responses contributing to discourse.

6.4. Cohesive Communication

Of the six independent variables analyzed, only one initial post characteristic significantly contributed to the model predicting the amount of cohesive communication present in a response post: word count. Cohesive communication is that which contributes to students identifying and acting as part of the learning community, and includes using inclusive pronouns when discussing the group, addressing peers by name, or using phatics or salutations to help develop familiarity within

the group (Garrison, 2011). Word count in initial posts contributed to a model predicting the density of cohesive communication in responses, in that a lower word count predicted a higher level of cohesive communication. Zingaro and Oztok (2012) hypothesize that longer messages include more ideas that peers could respond to than shorter messages; it is possible that fewer words in the initial post provided peer less to respond to, and so responses used more "friendly" wording rather than more content-focused wording.

Deeper investigation into each of the three indicators of cohesive communication found that word count contributed significantly to a model predicting Cohesive Communication: Phatics or salutations. This category includes communication that serves a purely social function such as greetings or closures. This model predicts that initial posts with lower word counts will generate responses that have higher use of phatics or salutations. This could be grounded in the idea that a shorter initial post provides fewer opportunities for the response to speak to, and the responding peer may try to bulk up their post by including a greeting or sign off.

As to the five independent variables that did not significantly contribute to the model predicting cohesive communication, there are many possible explanations. An initial post that is made well before the due date may attract the response of other students working ahead of schedule, who may be more likely to try to build cohesive relationships within the group; however, students who post later in the week may also generate responses that contain cohesive communication as peers try to draw those late posters in and make them feel a part of the community. It is also possible that the concept of cohesive communication is tied more specifically to communication actions that have less to do with the content of

the post, and more to do with the community itself. Reading ease, first- and second person language and cognitive presence may have no bearing on whether a peer chooses to use greetings or peer names, as these actions are not tied to the content itself but rather to building that community.

6.5. Conclusion

This study aimed to investigate the influence that characteristics of student initial discussion posts may have on the engagement of their peers as demonstrated in peer response posts. The results indicated that some initial post characteristics can influence peer emotional engagement in terms of social presence demonstrated in discussion response posts. These findings show that the work of students can influence the engagement of their peers in online discussions. For the emotional dimension of engagement, how a student initial post influenced peer engagement was investigated in terms of peer social presence demonstrated. First-person language, reading ease, and word count each influenced at least one aspect of peer social presence in responses. These results indicate that the decisions made by a student in their initial post influence their peers, specifically in terms of how those peers participate as members of the learning community, communicate with one another, and build relationships (Garrison, 2009). By making decisions related to what to include in their post, students are affecting the decisions made by peers to demonstrate social presence, and how they present themselves as a part of the learning community. In this way, one student may influence how a peer not only learns but develops as a contributing member of a community of learners.

In summary, the findings from this study contribute to the overall understanding that students can and do influence the engagement

of their peers. While the focus of this study was student engagement and how it is demonstrated, that engagement is also directly related to how students learn. Not only is engagement tied to active involvement in a course that can lead to student achievement of learning outcomes and cognitive development (Ma et al., 2015), but it also influences how a student chooses to interact with course material. Being engaged with a certain discussion thread begun by a peer directs a student's trajectory of thought by forcing a response to a specific idea presented by another. This means that the student is not just thinking for themselves in terms of the course content, but is reacting to the ideas or understanding put forth by another person, which requires additional consideration to relate to, agree with, or diverge from. In this way, the work of one student can influence not only the engagement of peers, but their potential for learning as well.

6.6. Limitations

There are several limitations of this study. The first is that the specific design of the discussions in the course that data was drawn from may have had a larger influence on the responses generated than anticipated, or that could be generalized to a larger or different population. Another limitation is the discussion prompts that students were able to choose from: certain prompts may have led to some responses automatically requiring more social presence than others. The third limitation deals with the reliability of the data included in the analysis: only one researcher coded the data in two separate coding periods in order to test coding reliability. For some variables, the reliability between coding periods was not as robust as it could have been.

6.7. Recommendations

Recommendations for Practice. Based on the findings of this study, recommendations for instructors and designers may be made as they design discussion posts aiming to encourage emotional engagement among students. Instructors or designers may consider discussing the concepts of social presence, and its effect on engagement and success, with their students prior to assigning online discussions. Instructors could also encourage demonstration of social presence by adjusting their discussion design: asking for personal opinions, prompting students to ask questions of one another, or asking students to include specific reference to points made by their peers. Finally, instructors or designers could include example discussion posts that demonstrate to students what is appropriate within course discussions. For example, by demonstrating the use of salutations and expressions of appreciation or agreement with peers, instructors could model behavior that students should themselves follow as they interact with their peers.

Recommendations for Research. Beyond recommendations for practical applications in terms of the results of this study, some recommendations for future research in this area may be made. Future studies might find more conclusive results by working with a course with different design considerations, such as one that uses whole-class discussions or includes fewer discussion prompts per week. Further, studies could employ data from multiple completely different courses to draw data that could be generalized across a larger swath of undergraduate or graduate student populations. Future studies might also find success in designing an experimental study, rather than a correlational study. By including an intervention to be tested with one portion of the sample, such as including a module on social and cognitive presence prior to the first discussion or encouraging questions or

more personal details in initial posts, then comparing the results of the test and control groups, it may be possible to see differences in results in terms of specific variables.

6.8. Summary

No student is an island when it comes to learning, although it may feel that way to some students in online classes. Students must actively interact with the content, their instructor, and their peers to build their own understanding and knowledge. How students engage with the course and their peers impacts their learning as part of a community of learners; investigating the variables that may influence that engagement provides a starting point to better understand how to encourage or build deeper engagement in discussions and the course overall. This study focused on how students influenced the emotional engagement of their peers as demonstrated in online discussion posts. This knowledge can help instructors and designers to better understand how students influence one another, and how to best structure online discussions to encourage traits that set in motion such engagement.

References

- Abe, J. A. A. (2020). Big five, linguistic styles, and successful online learning. *The Internet and Higher Education*, 45, 100724. <https://doi.org/10.1016/j.iheduc.2019.100724>
- Ali, A., & Smith, D. (2015). Comparing social isolation effects on students attrition in online versus face-to-face courses in computer literacy. *Issues in Informing Science and Information Technology*, 12, 11–20. Retrieved from <http://iisit.org/Vol12/IISITv12p011-020Ali1784.pdf>
- Allen, I.E. & Seaman. J. (2016). Online report card: Tacking online education in the United States. Babson Survey Research Group, Newburyport, MA.
- Allen, I. E., & Seaman, J. (2017). Digital learning compass: Distance education enrollment report 2017. Babson Park, MA: Babson Survey Research Group.
- Blanchette, J. (2011). Participant interaction in asynchronous learning environments: Evaluating interaction analysis methods. *Linguistics and Education*, 23(1), 77-87. <https://doi.org/10.1016/j.linged.2011.02.007>
- Bulu, S. T. (2012). Place presence, social presence, co-presence, and satisfaction in virtual worlds. *Computers & Education*, 58, 154-161. <https://doi.org/10.1016/j.compedu.2011.08.024>
- Canvas. (2020). Harvesting discussion data. Available at: <https://learn.canvas.net/courses/1176/pages/5-dot-1-2-reading-harvesting-discussion-data>
- Carroll, D. W. (2007). Patterns of student writing in a critical thinking course: A quantitative analysis. *Assessing Writing*, 12(3), 213-227. <https://doi.org/10.1016/j.asw.2008.02.001>
- Cobb, S. C. (2009). Social presence and online learning: A current view from a research perspective. *Journal of Interactive Online Learning*, 8(3), 241–254
- Doo, M. Y., & Bonk, C. J. (2020). The effects of self-efficacy, self-regulation and social presence on learning engagement in a large university class using flipped learning. *Journal of Computer Assisted Learning*, 36(6), 997-1010. <https://doi.org/10.1111/jcal.12455>
- Dumford, A. D., & Miller, A. L. (2018). Online learning in higher education: Exploring advantages and disadvantages for engagement. *Journal of Computing in Higher Education*, 30(3), 452-465. <https://doi.org/10.1007/s12528-018-9179-z>
- Fowler, F. J. (2002). *Survey Research Methods*. India: SAGE Publications.
- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. (2004). School engagement: Potential of the concept: State of the evidence. *Review of Educational Research*, 74(1), 59–119. <https://doi.org/10.3102/00346543074001059>
- Gardner, H. (1999). *Intelligence reframed: Multiple intelligences for the 21st century*. Basic Books.
- Garrison, D.R. (2009). *Communities of inquiry in online learning*. In P.L. Rogers, et al. (Eds.), *Encyclopedia of distance learning* (2nd ed.), IGI Global, Hershey, PA, 352-355.
- 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), 87-105. [https://doi.org/10.1016/S1096-7516\(00\)00016-6](https://doi.org/10.1016/S1096-7516(00)00016-6)
- Garrison, R. D., Anderson, T., & Archer, W. (2001a). Critical thinking and computer conferencing: A model and tool to assess cognitive presence. *American Journal of Distance Education*, 15(1), 1-18.
- Garrison, D. R., Anderson, T., & Archer, W. (2001b). Critical thinking, cognitive presence, and computer conferencing in distance education. *American Journal of Distance Education*, 15(1), 7-23. [https://doi.org/10.1016/S1096-7516\(00\)00016-6](https://doi.org/10.1016/S1096-7516(00)00016-6)

- doi.org/10.1080/08923640109527071
- Garrison, D.R. & Aykol, Z. (2010). The Community of Inquiry Theoretical Framework. *Handbook of Distance Education*, edited by Michael Grahame Moore, Taylor & Francis Group.
- Garrison, D. R. (2011). *E-learning in the 21st century: A framework for research and practice*. New York: Taylor & Francis.
- Garrison, D. R., Cleveland-Innes, M., Koole, M., & Kappelman, J. (2006). Revisiting methodological issues in transcript analysis: Negotiated coding and reliability. *The Internet and Higher Education*, 9(1), 1-8. <https://doi.org/10.1016/j.iheduc.2005.11.001>
- Gilbert, P. K., & Dabbagh, N. (2005). How to structure online discussions for meaningful discourse: a case study. *British Journal of Educational Technology*, 36(1), 5–18. <https://doi.org/10.1111/j.1467-8535.2005.00434.x>
- Hewitt, J. (2003). How habitual online practices affect the development of asynchronous discussion threads. *Journal of Educational Computing Research*, 28(1), 31-45. <https://doi.org/10.2190/PMG8-A05J-CUH1-DK14>
- Ho, C., & Swan, K. (2007). Evaluating online conversation in an asynchronous learning environment: An application of Grice's cooperative principle. *The Internet and Higher Education*, 10(1), 3-14. <https://doi.org/10.1016/j.iheduc.2006.11.002>
- Hu, S., & Kuh, G. D. (2002). Being (dis) engaged in educationally purposeful activities: The influences of student and institutional characteristics. *Research in Higher Education*, 43(5), 555–575. <https://doi.org/10.1023/A:1020114231387>
- Jacobi, L. (2017). The Structure of Discussions in an Online Communication Course: What Do Students Find Most Effective? *Journal of University Teaching & Learning Practice*, 14(1), 45–61. <https://doi.org/10.53761/1.14.1.4>
- Jones, Q., Ravid, G., & Rafaeli, S. (2004). Information Overload and the Message Dynamics of Online Interaction Spaces: A Theoretical Model and Empirical Exploration. *Information Systems Research*, 15(2), 194–210. <https://doi.org/10.1287/isre.1040.0023>
- Kent, C., Laslo, E., & Rafaeli, S. (2016). Interactivity in online discussions and learning outcomes. *Computers and Education*, 97, 116–128. <https://doi.org/10.1016/j.compedu.2016.03.002>
- Kovanović, V., Gašević, D., Joksimović, S., Hatala, M., & Adesope, O. (2015). Analytics of communities of inquiry: Effects of learning technology use on cognitive presence in asynchronous online discussions. *The Internet and Higher Education*, 27, 74–89. <https://doi.org/10.1016/j.iheduc.2015.06.002>
- Kozan, K., & Richardson, J. C. (2014). Interrelationships between and among social, teaching, and cognitive presence. *The Internet and Higher Education*, 21, 68-73. <https://doi.org/10.1016/j.iheduc.2013.10.007>
- Landis, J. R., Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159-174.
- Lee, S. (2014). The relationships between higher order thinking skills, cognitive density, and social presence in online learning. *The Internet and Higher Education*, 21, 41-52. <https://doi.org/10.1016/j.iheduc.2013.12.002>
- Lipman, M. (2003). *Thinking in education* (Second edition.). Cambridge University Press.
- Lobmier, J. H. (2010). Nonexperimental designs. In Salkind (Ed.), *Encyclopedia of research design*. Thousand Oaks, CA: SAGE Publications, Inc. doi: 10.4135/9781412961288
- 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>
- Ma, J., Han, X., Yang, J., & Cheng, J. (2015). Examining the necessary condition for engagement in an online learning environment based on learning analytics approach: The role of the instructor. *The Internet and Higher Education*, 24, 26–34. <https://doi.org/10.1016/j.iheduc.2014.09.005>
- Mayer, R. E., Fennell, S., Farmer, L., & Campbell, J. (2004). A personalization effect in multimedia learning: Students learn better when words are in conversational style rather than formal style. *Journal of Educational Psychology*, 96(2), 389–395. <https://doi.org/10.1037/0022-0663.96.2.389>
- Martin, J. (2019). Building relationships and increasing engagement in the virtual classroom: Practical tools for the online instructor. *The Journal of Educators Online*, 16(1).
- Pena-Schaff, J. B., & Nicholls, C. (2004). Analyzing student interactions and meaning construction in computer bulletin board discussions. *Computers and Education*, 42(3), 243-265. <https://doi.org/10.1016/j.compedu.2003.08.003>
- Ramsden, P. (1988). Context and strategy: Situational influences on learning. In R. R. Schmeck (Ed.) *Learning strategies and learning styles*, 159-184. New York: Plenum
- Readability Formats Website. (2014). The Flesch reading ease readability formula. Retrieved from <http://www.readabilityformulas.com/flesch-reading-ease-readability-formula.php>.
- Rourke, L., Anderson, T. Garrison, D. R., & Archer, W. (2001). Assessing social presence in asynchronous, text-based computer conferencing. *Journal of Distance Education*, 14(3), 51-70.
- Rovai, A. P. (2002). Building sense of community at a distance. *The International Review of Research in Open and Distance Learning*, 3(1), 1–16. <https://doi.org/10.19173/irrodl.v3i1.79>
- Rovai, A. P., Ponton, M. K., Baker, J. D. (2013). *Social science research design and statistics: A practitioner's guide to research methods and IBM SPSS analysis*. United States: Watertree Press.
- Russell, M., Kleiman, G., Carey, R., & Douglas, J. (2009). Comparing self-paced and cohort-based online courses for teachers. *Journal of Research on Technology in Education*, 41(4), 443–466. <https://doi.org/10.1080/15391523.2009.10782538>
- Seaman, J.E., Allen, I. E., & Seaman, J. (2018). *Grade Increase: Tracking online education in the United States*. Babson Park, MA: Babson Survey Research Group.
- Schaeffer, C. E., & Konetes, G. D. (2010). Impact of learner engagement on attrition rates and student success in online learning. *International Journal of Instructional Technology & Distance Learning*, 7(5), 3-9.
- Swan, K. (2002). Building learning communities in online courses: The importance of interaction. *Education, Communication and Information*, 2(1), 23-49. <https://doi.org/10.1080/1463631022000005016>
- Swan, K., & Shih, L. F. (2005). On the nature and development of social presence in online course discussions. *Journal of Asynchronous Learning Networks*, 9(3), 115-136.
- Tate, R. (1992). General linear model applications. Unpublished manuscript, Florida State University.
- Tabachnick, B.G. & Fidell, L. Sl. (2007). *Using multivariate statistics* (5th ed.).

Boston, MA: Allyn & Bacon.

- 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>
- Tu, C., & McIsaac, M. (2002). The relationship of social presence and interaction in online classes. *The American Journal of Distance Education*, 16(3), 131-150. https://doi.org/10.1207/S15389286AJDE1603_2
- Vannatta Reinhart, R., & Mertler, C. A. (2016). *Advanced and multivariate statistical methods: Practical application and interpretation*. Taylor and Francis.
- Vai, M. & Sosulski, K. (2016) *Essentials of online course design: A standards-based guide*. New York, N.Y.: Routledge Press.
- Wei, C., & Chen, N. (2012). A model for social presence in online classrooms. *Educational Technology Research and Development*, 60(3), 529-545. <https://doi.org/10.1007/s11423-012-9234-9>
- Zingaro, D., & Oztok, M. (2012). Interaction in an asynchronous online course: A synthesis of quantitative predictors. *Online Learning*. 16(4), 71. <https://doi.org/10.24059/olj.v16i4.265>

Contact the Authors

Kathryn Hanselman

Ph.D., Assistant Director,
Office of Digital Learning
University of Nevada, Reno, USA
Email: kathrynh@unr.edu

 <https://orcid.org/0000-0003-2238-7149>

Leping Liu

Ph.D., Professor of Information Technology
and Statistics
University of Nevada, Reno, USA
Email: liu@unr.edu

 <https://orcid.org/0000-0001-5859-8189>