

MEASURING THE COMMUNITY IN ONLINE CLASSES

Beth Rubin

DePaul University

Ronald Fernandes

DePaul University

ABSTRACT

This paper proposes both theoretical and empirical approaches to differentiating the community of inquiry (CoI) in online classes at individual and group levels. Following the example of research on organizational climate, it assesses the strength of shared perceptions of teaching presence, social presence, and cognitive presence. The paper develops a theory of composition that relates isomorphic constructs of these presences at the individual and class levels. Hypotheses are made about the agreement among students on the survey that is used to measure individual perceptions of the three presences of the CoI. These are tested using a set of statistics designed to measure the extent to which shared perceptions of these presences exist within classes. Strong evidence of shared perceptions is found, and potential practical and theoretical applications are described.

KEYWORDS

Community of inquiry, teaching presence, social presence, cognitive presence, online learning, learning effectiveness, distance learning, multilevel research, composition theory, organizational climate, class climate, within-group agreement

I. INTRODUCTION

The concept of a community of inquiry (CoI) was developed to describe phenomena that occur among all the members of online classes—in other words, to describe their interactions as a community. According to this view, “Deep and meaningful learning ... takes place in a community of inquiry composed of instructors and learners as the key participants in the educational process. The model assumes that in this community, learning occurs through the interaction of three core components: cognitive presence, teaching presence, and social presence” [1, p. 51–52]. Research on the CoI measures the presences both at the group level and at the individual level. Years of research, conducted by coding and counting discussion posts from entire classes, have validated the theory [1, 2, 3, 4]. Recently, there has been an explosion of research on the CoI that measures individual perceptions of teaching, social, and cognitive presences in online classes using surveys [5, 6]. Measuring the CoI through individual survey responses creates some degree of confusion about the level of analysis of the concept: Is the CoI framework founded on shared perceptions about a learning community, or does it refer to the individual perceptions and behaviors of students who happen to be taking a class? If CoI truly represents a shared learning community, should the model also incorporate the extent to which these perceptions or behaviors are shared?

A multilevel approach can help to address some recent critiques of the CoI framework. Rourke and Kanuka argued that there was little evidence that social and teaching presence lead to “deep, meaningful learning” [7]. Annand further questioned the role of social presence, particularly the aspects that reflect interaction among students; he argued that “social presence does not impact cognitive presence in a meaningful way” [8, p. 52] and held that individual differences between students are ignored in the CoI paradigm. Xin approached the model from the opposite perspective, holding that *all* online

communication is social in nature, and described the interconnectedness of social, teaching, and cognitive presence in determining the learning environment [9]. These critiques focus attention on the social aspect of the CoI, inviting examination of the level of analysis: How much of online learning is individual (and therefore should the CoI describe what individual students perceive and do), and how much is a group phenomenon (and therefore should the CoI describe what all the members of the group jointly perceive and do)? The reality may involve two parallel constructs, one focusing on the student and one on the learning community or class.

This study examines the CoI from the perspective of level of analysis. The research on organizational climate—in which individual perceptions of organizations, if shared, are aggregated to reflect group or organizational-level phenomena—is summarized. Both the theoretical and data analytic approaches to organizational climate are used as templates; a theory of composition is proposed in which social presence, teaching presence, and cognitive presences are described at both the level of the individual student and the class, and it is argued that shared perceptions of the class environment are a critical feature of a true CoI in online classes. A set of analytical techniques used to determine the strength of group shared perceptions are adapted to test hypotheses about the strength of class-level community; this approach can be used to measure group-level CoI by calculating the agreement among individual students on responses to the CoI survey.

II. ORGANIZATIONAL CLIMATE AND COMPOSITION THEORY

Organizational researchers have studied climate in organizations for more than fifty years [10, 11, 12, 13, 14]. Two isomorphic concepts have been identified and examined at different levels of analysis: psychological climate and organizational climate. Psychological climate is “the individual’s cognitive representation” of salient aspects of his/her situation and circumstances that mediates his or her perception of the environment, focusing attention and guiding interpretations of what occurs [15]. This representation is affected by the nature of the environment as well as the individual’s prior experiences, beliefs, needs, goals, and competencies [13, 15]. The psychological climate creates a frame through which individuals interpret the environment, and this frame affects what actions they take [13, 15, 16], which in turn affects the environment in a reciprocal process [15].

Organizational climate is the *shared* perception of what the work environment is like [17], or “how we do things around here” [18]. There are many objects, or referents, of the perceptions that make up the psychological and organizational climate. Some researchers have identified a set of categories that all employees use to frame the experience of work, such as the four factors of leader support and facilitation; role stress and lack of harmony; job challenge and autonomy; and workgroup cooperation, warmth and friendliness [13, 19]; or climate for learning, performance, and avoiding failure [20]. Other researchers focus on climate for a specific type of behavior or concern, such climate for service [17, 18], technical updating [21], safety [14, 22, 23], procedural justice [24], and ethical behavior and policies [25, 26].

Regardless of the referent, organizational climate affects how group members perceive their work and environment, and the climate affects member behaviors, such as following safety rules, updating technical skills, and providing high (or low) levels of service [18, 21, 23]. Organizational climate also affects outcomes such as job satisfaction and absenteeism [27]. Organizational climate can be examined at different levels of aggregation, such as work teams, departments, or an entire organization, and can differ across levels, such that a team can have one climate but the department in which it exists can have another [22, 28].

Organizational climate is affected both by structural factors of the organization, such as the size of groups and human resource practices, and by the cognitions and actions of members [13, 18, 27, 29, 30]. The formation of organizational climate is “interactive and reciprocal. That is, similar individuals are ... exposed to similar features within contexts, and share their interpretations with others in the setting. Over time this process tends to result in consensus on climate perceptions” [31, pp. 546–547]. Several studies have found that groups that have greater degrees of interaction and greater interdependence on work tasks have stronger agreement on the climate [24, 32]. Newcomers are socialized into the shared climate,

although individuals also affect the organizational climate by their behaviors, which reflect their experiences, values and perceptions [24, 29].

Many researchers have found that group leaders play a critical role in developing both psychological and organizational climate [14, 18, 20, 27, 31]. According to Dragoni, “[l]eaders model the behavior they deem appropriate, provide direct and indirect feedback on whether group members have met expectations, and reward individuals who exhibit expected behaviors” [20, p. 1086]. The closeness of the relationship employees have with leaders affects the climate; employees who have closer relationships with their leaders have more positive views of the climate and greater consensus about it [31]. Similarly, groups that have transformational leaders, who are proactive and engaged closely with subordinates, have more consensus on climate than do groups with nontransformational leaders [14].

The construct of climate in organizations has much in common with the CoI in classrooms. Both describe perceptions of the environment shared among members and leaders; both describe psychological spaces that are created by the interaction between structural factors and participants, and in which group members and leaders (i.e., teachers) reciprocally affect and are affected by the shared space. Both refer to what happens in the group; in the case of the CoI, the context is a group created for the purpose of learning.

A. Composition Theories

The concepts of individual and organizational climate are differentiated from each other in composition theories, which guide understanding of related phenomena that exist at different levels. These kinds of theories “specify the functional relationships among phenomena or constructs at different levels of analysis (e.g., individual level, team level, organizational level) that reference essentially the same content but that are qualitatively different at different levels” [33, p. 234]. According to Chan’s [33] widely cited analysis, there are five different kinds of composition theories: additive, direct consensus, referent-shift consensus, dispersion, and process theories. In additive theories, the group-level concept reflects the sum or average of the concept at lower levels, such as the research productivity of a university department. In direct consensus theories, the group-level concept exists only when members at lower levels of aggregation demonstrate consensus, or agree with one another. For example, if employees of a bank branch share a common understanding of how they should treat customers, there can be an identifiable *climate for service* within the branch. That common understanding held at the group level is different from the perceptions of each individual, because it creates a norm that affects new employees, who change their behavior to match it. In order to aggregate individual responses in a direct consensus theory, there must be demonstrable agreement among the members. Once that is established, the individual perceptions can be summed or averaged to create a measure of the group’s climate.

Referent-shift consensus theories are similar in that they require consensus among group members, but the referent concept changes from the individual to group level; for example, perceptions of self-efficacy refer to individuals, while perceptions of team efficacy reflect shared beliefs about a team’s ability to accomplish a task. Dispersion theories link concepts at different levels by viewing the degree of agreement or disagreement among members of the group as the concept of interest. Last, process composition theories describe how phenomena at different levels develop over time, rather than describing a static state. Organizational climate has been studied using direct consensus, referent-shift, and dispersion theories; all have implications for measurement.

B. Measuring Organizational Climate

There is a significant degree of agreement about how to measure organizational climate: individuals are asked to describe the climate, and those responses are averaged, summed, or otherwise combined for a group or larger unit if sufficient agreement exists [11, 27, 33, 34, 35]. In a study guided by a direct consensus theory, organizational climate is examined using questions that refer to individual behaviors, beliefs, and values (e.g., “The supervisor to whom I report is cold to me”). The individual responses are then integrated into a shared perception of climate [14, 17, 22, 23, 30, 32, 33, 36, 37].

Several different analytical indices have been recommended to assess the degree of agreement among group members, and different approaches have been recommended to determine what constitutes a sufficient level of agreement. Details of the different statistics, and their advantages and disadvantages, are presented in a Technical Appendix. The statistics r_{WG} and $r_{WG(j)}$ [35], Average Deviation (AD), ICC(1) and ICC(2) are very commonly used to measure within-group homogeneity and between-group variance, and in recent years many researchers have included multiple indices within their studies of workplace climate [14, 17, 22, 23, 32, 36, 37]; this practice is recommended for determining the extent to which group perceptions are shared [34, 39].

The next section will develop a direct consensus composition theory for the CoI, linking individual and group-level teaching, social, and cognitive presences. The CoI bears many similarities to the construct of organizational climate in that both describe a shared set of perceptions of the environment and behavioral expectations in a group (i.e., class section) that is structured and guided by a leader.

III. A COMPOSITION THEORY OF THE COMMUNITY OF INQUIRY

The CoI model takes a “collaborative constructivist perspective on the teaching and learning transaction ... [where] collaboration is seen as an essential aspect of cognitive development since cognition cannot be separated from the social context” [40, p. 92]. According to this perspective, knowledge is created through communication and interaction. It is intrinsically tied to the social context of the class; learning in online courses occurs through the interaction of the faculty, student, and course materials, mediated by technology. Deep, higher-order learning occurs in, and as a result of, discourse that takes place among the members of a class or other learning community [41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52]. The community of inquiry model that describes learning processes in online courses is intrinsically a group phenomenon, generally existing within an online class and comprised of the students and teacher(s) of that class.

There have been two main approaches to measuring the CoI elements: (1) coding and counting transcripts of asynchronous discussions, which focus on the group level, and (2) conducting surveys of students, which focus on the individual level. In most studies that use transcript coding, each sentence or message is coded, and the codes are aggregated across all members of the group who participated in the discussion by summing, averaging, or other using some other calculation [1, 4, 49, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65]. Most of the research aggregates posts to examine the community of inquiry that exists within a class, although some look at larger levels, such as disciplinary areas [66] and some at learning groups within a class [55].

Coding and counting discussion threads is painstaking and labor-intensive, so an alternative approach to examining the CoI was developed: surveys filled out by students. The Community of Inquiry Survey has been extensively validated in a wide range of universities with very large samples in two countries [5, 6, 46, 47, 56, 67, 68, 69, 70, 71]. The instrument consists of 34 items measuring social presence, cognitive presence, and teaching presence. Individual students respond to questions describing their perception of multiple indicators of each presence.

Survey responses have generally been analyzed using the student as the level of analysis. For example, research has examined the relationships among different presences, and their correlation with individual attitudes, such as course satisfaction [46, 47, 51, 52, 56, 65, 67, 72, 73]. Other research examined relationships between course factors (such as design, discipline, or class size) and individual student scores on CoI surveys [67, 68]. In a few cases, individual survey responses have been aggregated into means to create class-level variables [e.g., 53, 54, 63]; however, no assessment of the agreement among class members was undertaken. Overall, the complexities of the level of analysis have not been addressed.

The following sections describe each element of the CoI framework at both the individual and class section (group) levels.

A. Social Presence

Social presence describes the strength of the social relationships and emotional connection among the members of a class or learning community. It has three elements: identifying with a learning community, communicating openly in a trusting environment, and developing interpersonal relationships [44, 74]. This construct has evolved over the past twelve years, moving from an emphasis on immediacy, projecting a sense of “self” into the digital space, and developing emotional connection and open communication [1, 62, 72, 75] to focus more on educational goals and behaviors.

Some of the changes in the framing of these concepts over time may reflect the transition from the original methodology of measuring group phenomena through aggregated discussion posts [1, 4] to the more recent approach of measuring the CoI through individual survey responses [e.g., 5, 6]. For example, the original element of *affective expression* became framed as *developing interpersonal relationships*; the total amount of affect expressed in a group discussion forum is now viewed as a set of actions that individuals take to develop relationships. Similarly, the facet of *group cohesion* became *identifying with the community*; groups have cohesiveness, while individuals identify with groups.

Social presence reflects public behaviors on the part of individual students and faculty that are perceived by others and jointly form a collective perception of people in the group and of how “things are done” in the learning community. Student actions to relate with others in the learning community create a cohesive group when most students engage in those actions. In this formulation, individual-level social presence refers to the feelings of an individual and the actions that an individual takes to become a part of a community of inquiry, including personally identifying with the community, communicating in a trusting and comfortable environment, and developing interpersonal relationships with others in the class. Community-level social presence refers to the aggregate or shared sense of belonging that the students experience, the overall comfort and trust experience of the learning group, and the interpersonal relationships that exist among all members.

Many student actions are contributions to discourse made in public posts or in live sessions, available to be viewed by others. This visibility should lead to agreement among class members about the level of social presence in the class. However, students vary in their interaction with learning groups due to personal factors, such as self-efficacy [59, 71], amount of prior online experience [67], and circumstances such as family and work demands, which can limit time for active participation [76]. Students range widely in their engagement with their peers through reading or responding; those who are not involved in creating a shared community tend to perceive less connection with others. Swan and Shih found that individual behavior is related to perceptions of social presence; students who perceived the most social presence displayed nearly 50% more indicators of social presence in their own discussion posts than did students who perceived the least social presence [65].

This phenomenon occurs in workplace climate as well, where interaction with others affects the levels of climate perceived even when there is a high degree of agreement [27]. Although everyone does not participate equally in a community, it is nonetheless a community. While individual levels of social presence may vary, there is a measurable community-level social presence in online classes that are designed and taught to support it; as Arbaugh and Hwang note, “a community will not be created unless participants actually engage each other” [77, p. 12].

Teachers affect the degree of social presence by the way that they design assignments, such as using group activities, as well as by teaching activities such as creating informal discussion areas, rewarding students for having discussions with one another, modeling openness and encouraging it among students, and through many other teaching behaviors [44, 50, 56, 65, 78, 79]. Therefore, this theory also holds that there are differences between classes in the degree of community-level social presence that exists in each class. Because course design, instructor behavior, and student characteristics all may affect social presence, sections of the same course are likely to differ from one another. Therefore the focus on agreement and disagreement is placed at the level of a class section rather than at the level of a unique course.

This is a direct consensus composition theory in that it proposes agreement among class members regarding answers to questions about each person's personal perceptions, feelings, and experiences in the online class. If there is strong agreement among the students about these perceptions, feelings, and experiences, a community with social presence would exist. There would also be an expectation that classes would differ from each other in the extent to which they exhibit community-level social presence. There would then be a justification for considering the effects of community-level social presence in online learning research.

Hypothesis 1a: It is hypothesized that there will be strong agreement among members of a class section on perceptions of social presence.

Hypothesis 1b: It is hypothesized that class sections will differ in their community-level social presence.

B. Teaching Presence

Teaching presence is the extent to which the course instructor designs learning experiences, guides and directs student work, and facilitates interaction to support deep learning [4, 44, 46, 47, 50, 73, 77]. Akyol and Garrison referred to it as the provision of "leadership throughout the course of study" [54, p. 235]. This concept has expanded from its original focus on the design and support of discourse in asynchronous discussions to include the design of all learning activities and the provision of feedback [47, 77, 79, 80].

While instructor behavior is critical, student perceptions of that behavior are just as important in developing a community of inquiry. For example, an instructor who intends to provide clear guidance but who is perceived as condescending or overly directive will not be effective in creating an environment that supports exploration and deep learning. Similarly, if instructors provide information that students cannot access (perhaps due to difficulty locating it), it will not help students learn. Therefore, teaching presence should be measured through the perceptions of students.

Teaching behavior is akin to group leadership; just as organizational leaders set goals, prioritize tasks, give feedback, and manage day-to-day work for employees, instructors set learning goals, structure activities, solve problems, facilitate work, provide information, and give feedback. Teachers lead the class in the achievement of its goal, which is learning. As with organizational leaders, most actions taken by teachers occur in public venues; these actions include the design of course elements, learning activities, and discussion prompts; the communication of evaluative criteria and rubrics; public announcements to guide movement through the course and address administrative issues; and posting in the discussions to support extended learning. The only actions that are not shared throughout the class are individual e-mails and private feedback on work. Therefore, perceptions of teaching presence are likely to be widely shared.

In this theory of composition, individual-level teaching presence refers to each student's perception of the instructor's actions to support teaching by designing learning experiences, guiding learning, and providing feedback and direct instruction. Community-level teaching presence, on the other hand, is the shared perception of the instructor's actions to support teaching by designing learning materials and activities and organizing the class's progression through them; facilitating discourse among the members in a range of modalities (in discussions, groups, webinars, or other settings); and providing feedback and direct instruction for the entire group.

If there is a strong shared agreement among the students on these perceptions, a community with teaching presence would exist. Teaching presence is affected by the individual instructor, so class sections would also differ in their community-level teaching presence scores. This would justify considering the effects of community-level teaching presence in online learning research.

Hypothesis 2a: It is hypothesized that there will be strong agreement between members of a class section on perceptions of teaching presence.

Hypothesis 2b: It is hypothesized that class sections will differ in their community-level teaching presence.

C. Cognitive Presence

Cognitive presence has been defined as “the extent to which the participants in ... a community of inquiry are able to construct meaning through sustained communication” [40, p. 89]. It reflects the extent to which students engage in all four stages of the practical inquiry process [81], including addressing triggering problems, exploring the concepts and issues that underlie the problems, integrating information to identify solutions, and evaluating those solutions and considering further applications [3].

One might imagine that cognitive presence is more of an individual than a shared phenomenon; because learning occurs within each student’s mind, cognitive presence might be viewed as an individual and unique procession through the stages of deep learning. However, in a community of inquiry these stages result from students interacting with the learning activities created by the faculty (teaching presence), under faculty guidance (teaching presence), and are supported by interaction with other students (social presence) [82]. Several studies have found that teaching and social presence together predict cognitive presence [46, 56, 71, 83]. Thus, like social and teaching presence, cognitive presence has both an individual and a community component. Individual-level cognitive presence is viewed as the extent to which a specific student progresses through the four stages of inquiry, while community-level cognitive presence is the extent to which the entire group makes this progression.

In a learning community, both faculty and other students support engagement, correction, deeper understanding, and learning. The selection and design of course activities creates more or fewer opportunities for higher-order thinking, resolution, and integration; for example, Kanuka and her colleagues found that debates and WebQuests produced higher levels of cognitive presence than did reflective deliberation, invited experts, or nominal group techniques [58]. Both individual and group-level phenomena affect learning; Kanuka and Garrison identified three external processes that exist among members of a class that support cognitive presence—discourse, collaboration, and management—and three internal processes that a student must personally engage in—reflection, monitoring, and the construction of knowledge [84]. They concluded that “critical thinking is not solely a reflective, personal, and internal process. Rather, it is an iterative relationship between the personal and shared worlds” [84, p. 25].

Clearly, cognitive presence is also affected by individual differences in student motivation, the time students dedicate to their classwork, self-efficacy, and self-regulation [59, 71]. There are also situational factors that support or hinder learning, particularly for adult students [76]. Students must participate, whether publicly or privately, in order to achieve deep learning; they must read resources and posts, ask and answer questions, reflect on ideas, explain concepts, evaluate options, and apply concepts to other situations. They must choose to take the time to read or listen to others’ contributions, and create their own. When teachers structure and support practical inquiry and deep learning, and students engage in discourse in support of learning, there will be collective-level cognitive presence. The more members of a learning community participate, the more a community is created in which deeper learning is possible and is shared among the members. As with social presence, cognitive presence is affected by students, teachers, and course design; therefore it should be examined at the level of the class section.

Hypothesis 3a: It is hypothesized that there will be strong agreement between members of a class section on perceptions of cognitive presence.

Hypothesis 3b: It is hypothesized that class sections will differ in their community-level cognitive presence.

IV. METHODOLOGY

D. Sample

The hypotheses were tested in a large private midwestern university. Respondents were 875 students enrolled in 126 sections of 44 unique courses in five different colleges. Classes were taught by 55 different faculty over the course of three years, in three graduate programs in the Schools of Public

Service (5% of responses representing five class sections with three unique courses), Commerce (2.5% of responses, representing one class), and Education (8% of responses representing six class sections with two unique courses), and two undergraduate programs in Computer Science (1.8% of responses, with one class) and an interdisciplinary school for adult students (82.6% of responses, reflecting 113 sections with 37 unique courses). The interdisciplinary courses included liberal arts, humanities, sciences, and professional studies. Most (85.5%) were undergraduate students. Students ranged from 19–69 years old, with an average age of 35.5 years old. Nearly two-thirds (64.7%) were female, and 35.3% were male. The vast majority (83.3%) had taken prior online courses, with a mean response of between four and five prior courses.

E. Procedures and Measures

Students were offered the opportunity to participate in the study by their instructors, some of whom offered extra credit for participation. Participants' names were sent to researchers after classes had ended so that credit could be allocated. All classes but two were designed to create a community of inquiry, with the assistance of an instructional designer. All included active engagement in discussion, guided by the teachers; some involved weekly synchronous webinar meetings. Most were focused not only on developing competence but also on active application of knowledge to the real world. Courses taught more than once were “locked down” with prestructured syllabi, learning materials, activities, discussions, and assessments. Instructors were trained in the learning management system, and most were also trained in how to support active engagement in online courses.

Faculty were invited to participate, and those who did sent information to their students. Some, although not all, offered extra credit for participation, which is a common practice in this university. Student participants were sent links to an online survey, which included the Community of Inquiry questionnaire [5, 6, 46, 69, 70, 85] along with questions about the technology used to teach the classes, and other outcomes, such as satisfaction. Only the community of inquiry responses are reported in this study. Response rates averaged 35.5%, which is consistent with response rates to online surveys [86]; class sizes ranged from 5–28 students, and the average number of participants per class was 6.9. The three scales in the survey had high internal consistency; the coefficient alpha was .95 for teaching presence, was .89 for social presence, and was .93 for cognitive presence.

Following the recommendations of Lüdtke et al. [87], LeBreton and Senter [34], and Biemann, Cole, and Voelpel [39] to include multiple indices of within-group agreement and between-group variance, multiple statistics are used to assess each. To test all of the hypotheses about the degree of agreement within groups (Hypotheses 1a, 2a, and 3a), three statistics are examined: $r_{WG(j)}$, ICC(2), and $AD_{M(j)}$. To assess the hypotheses about the variance among groups (Hypotheses 1b, 2b, and 3b), two statistics are used: F-ratio and ICC(1). Only groups with three or more respondents were used. While some researchers have recommended using groups of 10 or more [88], Maloney and her colleagues conducted a Monte Carlo study of missing data in small group composition research; they found that both $r_{WG(j)}$ and ICC provide more accurate estimates of consensus and reliability when all groups are included, even those with only a single respondent [89]. However, $r_{WG(j)}$ can only be calculated on groups of two or larger, and the possibility of response bias from missing data/measurement error from a nonrepresentative sample led to a cutoff of three students per course. Given that the scales include between nine and fourteen items, and that there is a reciprocal relationship between the number of items and raters [34], inclusion of groups with three or more respondents was deemed appropriate.

In order to calculate $r_{WG(j)}$ accurately, the distribution of the null hypothesis must be determined; this is the set of ratings that would be made by group members if there were no agreement whatsoever, raters were responding randomly, and the only factor influencing scores was response bias, such as central tendency, leniency, or severity. The likely bias must be established on theoretical grounds, considering the nature of the questions that are posed [34, 39].

The Community of Inquiry survey measures faculty behaviors in designing, developing, and leading online courses; student behaviors and beliefs about communication with one another to accomplish course

tasks; and student engagement in intellectual activities in the course. At least one large online university, American Public University, uses the CoI instrument as its course evaluation. Therefore, one would expect the distribution of responses to reflect patterns of course evaluation responses, particularly for teaching presence. Prior research has found a slight positive skew in teaching evaluations. For example, one study of more than 60,000 Australian university responses found average ratings between 2.92 and 3.17 on eight items about course design and instruction [90]; the response scale of 1–4 had a numerical mean of 2.5, and the average deviation from the mean was .61, or 15% higher. Another study of course evaluations in a California university found overall averages of 3.9 on a 1–5 scale, with an average deviation of .9 or 18% above the arithmetic mean of 3 [91]. We therefore hold the null hypothesis for teaching presence to reflect mild skew.

The questions in the social and cognitive presence scales overlap with many aspects of the National Student Survey of Engagement (NSSE) [92]. In particular, the NSSE "deep thinking" scale is comprised of three components: higher order thinking; integrative learning; and reflective learning. These include items that reflect working with others, and analyzing and applying information as found in the social and cognitive presence scales. One national study found that first-year online students averaged between 58, 69, and 62 on these three scales, which were 8%, 12%, and 19% above the arithmetic mean, respectively [93]. Another national study found that seniors reported average scores of 69, 57, and 60 on the scales, which were 7%–19% above the arithmetic mean [94]. Items indicative of social presence (labeled as "curricular peer interaction") were found to be skewed for older students, with seniors reporting a range of 35%–71% and an average of 57%, which was 7% above the mean [95]. Overall, there is reason to assume a slight positive skew in the distributions of social and cognitive presence as well.

Two programs were used to calculate statistics. First, SPSS syntax was created, following the description in LeBreton and Senter [34], to calculate the $r_{WG(j)}$, the $AD_{M(j)}$, and the $AD_{Med(j)}$ for every section. The $r_{WG(j)}$ statistics were calculated with a uniform null distribution as well as a slight skew. Second, the Excel tool developed by Biemann et al. [39] was used to calculate $r_{WG(j)}$ for each class as well as the ICC(1), ICC(2) and F-ratios. This tool automatically replaces values lower than 0.0 and greater than 1.0 with 0.0, because such values reflect complete lack of consensus and, typically, a bimodal distribution [34]. The distributions of a sample of such groups were plotted and examined, and all were found to be bimodal. The $r_{WG(j)}$ statistics of a sample of groups were compared across the two approaches to calculation and found to be identical for all indices greater than 0.0 and equal to or less than 1.0.

V. RESULTS

A. Social Presence

Hypotheses 1a and 1b addressed social presence. Hypothesis 1a held that there would be agreement among students in each class about the degree of social presence in the online community, and Hypothesis 1b held that there would be differences between class sections. As described previously, the first of these hypotheses is tested with four analytical methods; Table 2 presents the first of these, $r_{WG(j)}$. Perceptions of students in 126 class sections were examined, using both a uniform null distribution and a slightly skewed distribution. Following the procedure recommended by LeBreton and Senter [34], the proportion of groups with either very strong or strong agreement ($>.70$) are examined. Using a uniform null distribution, nearly all of the groups (124 out of 126 groups, or 98.4%) had either very strong or strong agreement; using a slightly skewed null distribution ($\sigma_e^2 = 1.34$), the proportion is smaller but still substantial, with 81.0% (102) of the groups showing strong or very strong agreement. The averages are .93 (uniform null) and .82 (slightly skewed null), which are well above the 95% confidence levels of .61 (uniform) and .74 (slightly skewed) reported by Cohen, Doveh, and Nahum-Shani for groups of 7 on scales with high internal consistency among items [96].

Social presence agreement	Size	$r_{WG(J)}$ Uniform null distribution	$r_{WG(J)}$ Slightly skewed null distribution
very strong	>.9	103	60
strong	.71-.9	21	42
moderate	.51-.7	2	8
weak	.31-.5	0	1
none	<.3	0	0
	out of range	0	8
	mean $r_{WG(J)}$	0.93	0.82
	SD of mean $r_{WG(J)}$	0.06	0.24

Table 1: $r_{WG(J)}$ statistics for the social presence scale

The second index, the $AD_{M(J)}$, was 0.59, and the $AD_{Med(J)}$ was 0.35. According to Burke and Dunlap [97], these are well below the cutoff of 0.80 for a 5-point scale and therefore indicate acceptable levels of agreement; 0.0 indicates no deviation, so a smaller score indicates greater within-group agreement. Both indices are also within the .05 significance level of the uniform distribution found by Cohen et al. for groups of seven on scales with high internal consistency reliability, (<0.99), and fall within the bounds for slightly skewed distributions (<0.70) [96, p. 159].

The ICC(2) was 0.82; traditionally, this is assessed as a reliability coefficient, although it measures both consistency and consensus and therefore may be lower than indices that measure either consistency or consensus alone [34, 39]. In this case the result exceeds the minimum cutoff of 0.70.

A large majority of class sections had strong agreement on $r_{WG(J)}$ falling within the range needed to meet the standard of $p < .05$ probability; the average levels fell within this standard as well, as did the ICC(2) and the $AD_{M(J)}$, with a uniform distribution. With all indices of agreement falling within cutoff levels at a probability of $p < .05$, Hypothesis 1a is supported.

Hypothesis 1b holds that class sections differ on social presence. Analysis of variance produced an F-ratio of 5.44 ($p < 0.000$), and an ICC(1) of 0.39, which is far above the minimum effect size of 0.05 [31, 36]. Hypothesis 1b is also supported.

B. Teaching Presence

Hypotheses 2a and 2b addressed teaching presence, with Hypothesis 2a holding that there would be agreement among students in each class section about the degree of teaching presence, and Hypothesis 2b holding that there would be differences between class sections. Table 2 presents the first set of analyses for Hypothesis 2a, $r_{WG(J)}$. With a uniform null distribution, 91.3% of sections (115 out of 126) had strong or very strong agreement, while 79.4% of sections (100) did so with a slightly skewed null distribution. The mean $r_{WG(J)}$ were .88 (uniform null) and .78 (slightly skewed null), both of which exceed the 95% confidence level minimums of .61 (uniform) and .74 (slightly skewed) [96].

Teaching presence agreement	Size	$r_{WG(J)}$ Uniform null distribution	$r_{WG(J)}$ Slightly skewed null distribution
very strong	>.9	103	81
strong	.71-.9	12	19
moderate	.51-.7	1	0
weak	.31-.5	0	2
none	<.3	1	3
	out of range	8	17

	mean $r_{WG(J)}$	0.88	0.78
	SD of mean $r_{WG(J)}$	0.25	0.34

Table 2: $r_{WG(J)}$ statistics for the teaching presence scale

The second index, the $AD_{M(J)}$, was 0.62 for teaching presence, and the $AD_{Med(J)}$ was 0.33. These fall below the 0.80 cutoff [95] as well as the estimates of Cohen et al. [96] for both uniform (<0.99) and skewed (<0.70) distributions. The ICC(2) was 0.82, which well exceeds the minimum reliability requirement of 0.70.

With all indices of agreement meeting the minimum levels needed to reject the null hypothesis with a probability of $p < .05$, Hypothesis 2a is supported.

Hypothesis 2b holds that class sections differ on their perceptions of teaching presence. The F-ratio was 5.60 ($p < .000$), and the ICC(1) was 0.40, far exceeding the minimum effect size of 0.05. Therefore, Hypothesis 2b is also supported.

C. Cognitive Presence

Hypothesis 3a holds that there will be agreement within class sections on cognitive presence, and differences between class sections. Table 3 presents the results of the $r_{WG(J)}$ calculations. Using a uniform null distribution, 97.6% of all sections (123 out of 126) had either very strong or strong agreement, and 87.3% of the sections (110) had very strong or strong agreement using a slightly skewed null distribution. The mean $r_{WG(J)}$ statistics were .94 (uniform) and .87 (slightly skewed), far exceeding the minimums of .61 (uniform) and .74 (slightly skewed) established by Cohen et al. [96].

Cognitive presence agreement	Size	$r_{WG(J)}$ Uniform null distribution	$r_{WG(J)}$ Slightly skewed null distribution
very strong	>.9	116	84
strong	.71-.9	7	26
moderate	.51-.7	1	1
weak	.31-.5	0	1
none	< .3	0	2
	out of range	2	6
	mean $r_{WG(J)}$	0.94	0.87
	SD of mean $r_{WG(J)}$	0.13	0.23

Table 3: $r_{WG(J)}$ statistics for the cognitive presence scale

The $AD_{M(J)}$ was 0.54, and the $AD_{Med(J)}$ was 0.28, which demonstrate acceptable levels of agreement at the $p < .05$ level, according to both Burke and Dunlap (<0.80) [97] and Cohen et al. [96] for both uniform (<0.99) and skewed (<0.70) distributions. The ICC(2) was 0.82, which again exceeds the 0.70 cutoff. Thus, all indices support Hypothesis 3a.

Hypothesis 3b holds that class sections differ on cognitive presence. This is supported by both the F-Ratio of 5.45 ($p < .000$) and the ICC(1) of 0.39, which exceeds the 0.05 minimum.

VI. DISCUSSION

Fourteen years of research have found that online classes are more successful in supporting deep learning when they are characterized by a community of inquiry [1, 3, 4, 44, 74]. An important advance was the development and validation of a survey instrument, which enables broader examination of the community of inquiry and factors that aid in its development [5, 6, 70, 85]. As this valuable instrument became widely used, attention focused on causal analysis and relationships among individual-level variables such as perceptions and student competences [46, 59, 71, 83, 98]. While both individual and group-level behaviors and perceptions have been part of the CoI framework, the question of level of analysis has not often been directly addressed. The concept of organizational climate provides a template for examining

levels in the CoI, because both individual and group-level climate exist, and both levels affect one another reciprocally [11, 15, 29]. Both climates are typically measured with the same surveys, and perceptions of individuals may be aggregated if there is sufficient agreement among the members of a group to support the existence of a shared group perception [38, 88].

This study posits that both individual and class-level CoI are relevant to understanding how online learning occurs. The perceptions and actions of individual students and teacher(s) are critical; ultimately, every student experiences learning as a result of his or her actions. However, in a true community of inquiry, the actions of many members together create synergy and, if carefully guided, can lead to both trusting relationships and a deep level of learning [56, 58, 99, 100]. In this work, individual-level CoI was formally distinguished from class, group, or community-level CoI.

This study examined agreement among perceptions of students in a class section on teaching, social, and cognitive presence. After describing the individual-level and group-level presences, it examined three indices of within-class agreement ($r_{WG(j)}$, $AD_{M(j)}$, and $ICC(2)$) and two indices of between-class differences (F-ratio and $ICC(1)$) on the CoI survey instrument [19, 30, 34, 38, 88, 89, 96, 97, 101, 102, 103], considering various possibilities for underlying response bias and examining deviation from means and medians. More than 125 classes were analyzed, and results strongly supported all hypotheses. All indices of agreement within class sections on social, teaching, and cognitive presences were significant, providing very strong evidence of community-level shared perceptions about CoI presences. All indices of differentiation across multiple class sections were also significant, indicating that sections differed from one another in their teaching, social, and cognitive presences at the community level.

Overall, the results found very strong support for the existence of a community of inquiry that is shared within each class section, and that differs across sections. The survey items that measure individual perceptions of teaching, social, and cognitive presences can also be used, when aggregated, to measure class-level teaching, social, and cognitive presences when agreement among students indicates that a learning community exists. This study provided statistical techniques to determine whether such a learning community exists, and evidence that it often does when built into the design of courses and expectations of teachers. It also provides evidence that the CoI exists at multiple levels; individual students perceive their own actions, the actions of others, and the online class environment. To a large extent those perceptions are shared with others, creating a genuine community that exists among the members.

This technique can be used by online teachers and administrators to examine the ability of teachers to create a strongly shared community within online classes. The individual levels of each presence, the average group levels, and the amount of agreement may be calculated to help instructors strengthen their teaching. For example, if low agreement indicates that subgroups have different levels of social presence in online courses, instructors may redesign courses to create group activities that engage all students, or modify grade distributions to reward interaction. Low agreement on cognitive presence may require redesigning learning activities or providing additional support for some students to enable them to fully integrate material. Low agreement about teaching presence may reflect that instructors are engaging more with some students than others, which can be interpreted as “playing favorites.” Examination of individual feedback would not readily reveal such issues, while examining agreement among class members would identify them easily.

A few potential limitations of the study should be noted. Between 5% and 13% of groups had $r_{WG(j)}$ scores that were out of range, either because they were smaller than 0.0 or greater than 1.0, indicating bimodal distributions. Teaching presence had the most groups with bimodal distributions and out-of-range scores. It is possible that this reflects two subgroups within the class that may be produced by sampling error. A number of students participated because class credit was given, which may have created two subgroups of respondents: those who needed the credit and were not engaged with their teacher or the course, and those who felt closely connected with their teachers and wanted to honor their request. The $r_{WG(j)}$ for the out-of-range (bimodal) classes was set to 0.0, which may have increased the likelihood of a Type 2 error. However, the significant findings indicate that, if response bias occurred, it

did not substantively alter the

results or conclusions. It would nonetheless be useful to examine agreement in situations in which all students complete the CoI instrument to eliminate any possibility of response bias.

This study is also limited by the fact that all data were collected in one university, and the overwhelming majority of the students were adults. In addition, the majority of the classes were taught in one school that adopts a social constructivist andragogy in its online courses. Most of these courses involve scaffolded learning, regular reflection on personal experience, real-world application, and intensive use of discussion for collaborative problem-solving [104, 105]. These approaches are likely to strengthen social and cognitive presences and enhance communal interactions, and therefore are likely to create stronger agreement among class members. These features increase the likelihood of a shared community of inquiry in the classes, but they are not representative of all online courses. It may be that in classes with more traditional pedagogies, focused on knowledge acquisition and multiple-choice testing, students interact differently with the teacher and content, creating less social presence and more variation among the students on teaching and cognitive presence. Further research in other contexts is needed to support the external validity of the findings.

On the other hand, this study introduces a methodology for considering levels of analysis and aggregation in online courses. It may be applied to other concepts that exist at different levels, such as class or university leadership, pedagogy, and other factors. Climate itself might be addressed in online classes; whereas the organizational climate literature has examined the climate for safety, technological updating, and justice, the topics of interest in online learning might include the climate for academic integrity, for welcoming diverse ideas, or maximizing learning.

While academic research has examined the relationships among individual responses to the CoI presences [46, 56, 71, 83], the relationships among the group scores should also be examined. Both the average class level and the degree of agreement among class members may have important effects. For example, it may be that class-level social presence affects cognitive presence more strongly than individual-level social presence, which may explain some findings that have led to critiques of social presence [7, 8]. The methodology introduced in this study could be used and extended to examine these possibilities.

It is also possible that courses in different subject areas have different levels of agreement. Prior research has found differences in the levels of teaching, social, and cognitive presences across disciplines, with different levels in “hard” vs. “soft” disciplines as well as “pure” vs. “applied” ones [66, 68]. As the pedagogies, student expectations, and instructor expectations all differ, both the level and agreement about the CoI will likely differ as well.

Another potentially fruitful area to explore is the effect of the strength of the agreement among members on outcomes. There is some evidence that the strength of agreement on perceptions about climate affects group outcomes; it is not just the average level, but how closely everyone agrees with one another that determines what groups do, because greater consensus creates a stronger force on individuals [17, 39, 106]. The same may occur in online classes; future research should examine the effect of class agreement on outcomes such as learning or satisfaction.

In sum, the analytical approach presented here enables practitioners and researchers to measure both individual and class-level teaching presence, social presence, and cognitive presence quickly and easily, using the validated and widely used CoI survey. This may help improve course instruction, facilitate research at the group level, and clarify conceptual issues regarding the role of community in online courses.

VII. REFERENCES

1. **Rourke, L., Anderson, T., Garrison, D.R., and Archer, W.** Assessing Social Presence in Asynchronous Text-Based Computer Conferencing. *Journal of Distance Education* 14(2): 50–71(1999).

2. **Anderson, T., Garrison, D.R., Archer, W., and Rourke, L.** Methodological Issues in the Content Analysis of Computer Conference Transcripts. *International Journal of Artificial Intelligence in Education* 12(1): 8–22 (2001).
3. **Garrison, D.R., Anderson, T., and Archer, W.** Critical Thinking, Cognitive Presence, and Computer Conferencing in Distance Education. *American Journal of Distance Education* 15(1): 7–23 (2001).
4. **Anderson, T., Rourke, L., Garrison, D.R., and Archer, W.** Assessing Teaching Presence in a Computer Conferencing Context. *Journal of Asynchronous Learning Networks* 5(2): 1–17 (2001).
5. **Swan, K., Richardson, J.C., Ice, P., Garrison, D.R., Cleveland-Innes, M., and Arbaugh, J.B.** Validating a Measurement Tool of Presence in Online Communities of Inquiry. *E-Mentor* 2(24) (2008).
6. **Arbaugh, J.B., Cleveland-Innes, M., Diaz, S.R., Garrison, D.R., Ice, P., Richardson, J.C., and Swan, K.** Developing a Community of Inquiry Instrument: Testing a Measure of the Community of Inquiry Framework Using a Multi-Institutional Sample. *The Internet and Higher Education* 11(3–4): 133–136 (2008).
7. **Rourke, L. and Kanuka, H.** Learning in Communities of Inquiry: A Review of the Literature. *Journal of Distance Education* 23(1): 19–48 (2009).
8. **Annand, D.** Social Presence within the Community of Inquiry Framework. *The International Review of Research in Open and Distance Learning* 12(5): 40–56 (2011).
9. **Xin, C.** A Critique of the Community of Inquiry Framework. *The Journal of Distance Education* 12(1) (2012).
10. **Forehand, G.A. and Von Haller, G.** Environmental Variation in Studies of Organizational Behavior. *Psychological Bulletin* 62(6): 361–382 (1964).
11. **James, L.R. and Jones, A.P.** Organizational Climate: A Review of Theory and Research. *Psychological Bulletin* 81(12): 1096–1112 (1974).
12. **Argyris, C.** Some Problems in Conceptualizing Organizational Climate: A Case Study of a Bank. *Administrative Science Quarterly* 2(4): 501–520 (1958).
13. **James, L.R., Choi, C.C., Ko, C.-H.E., McNeil, P.K., Minton, M.K., Wright, M.A., and Kim, K.** Organizational and Psychological Climate: A Review of Theory and Research. *European Journal of Work and Organizational Psychology* 17(1): 5–32 (2008).
14. **Zohar, D. and Luria, G.** Group Leaders as Gatekeepers: Testing Safety Climate Variations across Levels of Analysis. *Applied Psychology: An International Review* 59(4): 647–673 (2010).
15. **James, L.R., Hater, J.J., Gent, M.J., and Bruni, J.R.** Psychological Climate: Implications from Cognitive Social Learning Theory and Interactional Psychology. *Personnel Psychology* 31(4): 783–813 (1978).
16. **Brown, S.P. and Leigh, T.W.** A New Look at Psychological Climate and Its Relationship to Job Involvement, Effort, and Performance. *Journal of Applied Psychology* 81(4): 358–368 (1996).
17. **Schneider, B., Salvaggio, A.N., and Subirats, M.** Climate Strength: A New Direction for Climate Research. *Journal of Applied Psychology* 87(2): 220–229 (2002).
18. **Schneider, B., Macey, W.H., and Young, S.A.** The Climate for Service: A Review of the Construct with Implications for Achieving CLV Goals. *Journal of Relationship Marketing* 5(2/3): 111–132 (2006).
19. **James, L.A. and James, L.R.** Integrating Work Environment Perceptions: Explorations into the Measurement of Meaning. *Journal of Applied Psychology* 74(5): 739–751 (1989).
20. **Dragoni, L.** Understanding the Emergence of State Goal Orientation in Organizational Work Groups: The Role of Leadership and Multilevel Climate Perceptions. *The Journal of Applied Psychology* 90(6): 1084–95 (2005).
21. **Kozlowski, S.W. and Hults, B.M.** An Exploration of Climates for Technical Updating and Performance. *Personnel Psychology* 40(3): 539–563 (1987).

22. **Zohar, D. and Luria, G.** A Multilevel Model of Safety Climate: Cross-Level Relationships between Organization and Group-Level Climates. *Journal of Applied Psychology* 90(4): 616–628 (2005).
23. **Zohar, D.** A Group-Level Model of Safety Climate: Testing the Effect of Group Climate on Microaccidents in Manufacturing Jobs. *Journal of Applied Psychology* 85(4): 587–596 (2000).
24. **Roberson, Q.M. and Williamson, I.O.** Justice in Self-Managing Teams: The Role of Social Networks in the Emergence of Procedural Justice Climates. *Academy of Management Journal* 55(3): 685–701 (2012).
25. **Wimbush, J.C. and Shepard, J.M.** Toward an Understanding of Ethical Climate: Its Relationship to Ethical Behavior and Supervisory Influence. *Journal of Business Ethics* 13(8): 637–647 (1994).
26. **Wimbush, J.C., Shepard, J.M., and Markham, S.E.** An Empirical Examination of the Relationship between Ethical Climate and Ethical Behavior from Multiple Levels of Analysis. *Journal of Business Ethics* 16(16): 1705–1716 (1997).
27. **Lindell, M.K. and Brandt, C.J.** Climate Quality and Climate Consensus as Mediators of the Relationship between Organizational Antecedents and Outcomes. *Journal of Applied Psychology* 85(3): 331–348 (2000).
28. **Schneider, B., Ehrhart, M.G., and Macey, W.H.** Organizational Climate and Culture. *Annual Review of Psychology* 64(1): 361–388 (2013).
29. **Schneider, B. and Reichers, A.E.** On the Etiology of Climates. *Personnel Psychology* 36(1): 19–39 (1983).
30. **Glisson, C. and James, L.R.** The Cross-Level Effects of Culture and Climate in Human Service Teams. *Journal of Organizational Behavior* 23(6): 767–794 (2002).
31. **Kozlowski, S.W. and Doherty, M.L.** Integration of Climate and Leadership: Examination of a Neglected Issue. *Journal of Applied Psychology* 74(4): 546–553 (1989).
32. **Klein, K.J., Conn, A.B., Smith, D.B., and Sorra, J.S.** Is Everyone in Agreement? An Exploration of within-Group Agreement in Employee Perceptions of the Work Environment. *Journal of Applied Psychology* 86(1): 3–16 (2001).
33. **Chan, D.** Functional Relations among Constructs in the Same Content Domain at Different Levels of Analysis: A Typology of Composition Models. *Journal of Applied Psychology* 83(2): 234–246 (1998).
34. **LeBreton, J.M. and Senter, J.L.** Answers to 20 Questions About Interrater Reliability and Interrater Agreement. *Organizational Research Methods* 11(4): 815–852 (2008).
35. **James, L.R., Choi, C.C., Ko, C.-H.E., McNeil, P.K., Minton, M.K., Wright, M.A., and Kim, K.-I.** Organizational and Psychological Climate: A Review of Theory and Research. *European Journal of Work and Organizational Psychology* 17(1): 5–32 (2008).
36. **Schneider, B., White, S.S., and Paul, M.C.** Linking Service Climate and Customer Perceptions of Service Quality: Tests of a Causal Model. *Journal of Applied Psychology* 83(2): 150–163 (1998).
37. **Cho, J. and Dansereau, F.** Are Transformational Leaders Fair? A Multi-Level Study of Transformational Leadership, Justice Perceptions, and Organizational Citizenship Behaviors. *The Leadership Quarterly* 21(3): 409–421 (2010).
38. **James, L.R., Demaree, R.G., and Wolf, G.** Estimating within-Group Interrater Reliability with and without Response Bias. *Journal of Applied Psychology* 69(1): 85–98 (1984).
39. **Biemann, T., Cole, M.S., and Voelpel, S.** Within-Group Agreement: On the Use (and Misuse) of r_{WG} and $r_{WG(j)}$ in Leadership Research and Some Best Practice Guidelines. *The Leadership Quarterly* 23(1): 66–80 (2012).
40. **Garrison, D.R., Anderson, T., and Archer, W.** Critical Inquiry in a Text-Based Environment: Computer Conferencing in Higher Education. *The Internet and Higher Education* 2(2–3): 87–105 (2000).
41. **De Wever, B., Schellens, T., Valcke, M., and Van Keer, H.** Content Analysis Schemes to Analyze Transcripts of Online Asynchronous Discussion Groups: A Review. *Computers &*

- Education* 46(1): 6–28 (2006).
42. **Akyol, Z., Arbaugh, J.B., Cleveland-Innes, M., Garrison, D.R., Ice, P., Richardson, J.C., and Swan, K.** A Response to the Review of the Community of Inquiry Framework. *Journal of Distance Education* 23(2): 123–136 (2009).
 43. **Garrison, D.R.** Implications of Online Learning for the Conceptual Development and Practice of Distance Education. *Journal of Distance Education* 23(2): 93–103 (2009).
 44. **Garrison, D.R. and Arbaugh, J.B.** Researching the Community of Inquiry Framework: Review, Issues, and Future Directions. *The Internet and Higher Education* 10(3): 157–172 (2007).
 45. **Rovai, A.P.** Sense of Community, Perceived Cognitive Learning, and Persistence in Asynchronous Learning Networks. *The Internet and Higher Education* 5(4): 319–332 (2002).
 46. **Shea, P. and Bidjerano, T.** Community of Inquiry as a Theoretical Framework to Foster “Epistemic Engagement” and “Cognitive Presence” in Online Education. *Computers & Education* 52(3): 543–553 (2009).
 47. **Shea, P. and Bidjerano, T.** Cognitive Presence and Online Learner Engagement: A Cluster Analysis of the Community of Inquiry Framework. *Journal of Computing in Higher Education* 21(3): 199–217 (2009).
 48. **Shea, P.** A Study of Students’ Sense of Learning Community in Online Environments. *Journal of Asynchronous Learning Networks* 10(1): 35–44 (2006).
 49. **Schrire, S.** Interaction and Cognition in Asynchronous Computer Conferencing. *Instructional Science* 32(6): 475–502 (2004).
 50. **Rovai, A.P.** Facilitating Online Discussions Effectively. *The Internet and Higher Education* 10(1): 77–88 (2007).
 51. **Shea, P., Li, C.S., and Pickett, A.M.** A Study of Teaching Presence and Student Sense of Learning Community in Fully Online and Web-Enhanced College Courses. *The Internet and Higher Education* 9(3): 175–190 (2006).
 52. **Shea, P., Li, C.S., Swan, K., and Pickett, A.** Developing Learning Community in Online Asynchronous College Courses: The Role of Teaching Presence. *Journal of Asynchronous Learning Networks*, 2005. 9(4): p. 59-82.
 53. **Akyol, Z., Garrison, D.R., and Ozden, M.Y.** Online and Blended Communities of Inquiry: Exploring the Developmental and Perceptual Differences. *International Review of Research in Open & Distance Learning* 10(6): 65–83 (2009).
 54. **Akyol, Z. and Garrison, D.R.** Understanding Cognitive Presence in an Online and Blended Community of Inquiry: Assessing Outcomes and Processes for Deep Approaches to Learning. *British Journal of Educational Technology* 42(2): 233–250 (2011).
 55. **Bangert, A.** The Influence of Social Presence and Teaching Presence on the Quality of Online Critical Inquiry. *Journal of Computing in Higher Education* 20(1): 34–61 (2008).
 56. **Garrison, D.R., Cleveland-Innes, M., and Fung, T.S.** Exploring Causal Relationships among Teaching, Cognitive and Social Presence: Student Perceptions of the Community of Inquiry Framework. *The Internet and Higher Education* 13(1–2): 31–36 (2010).
 57. **Garrison, D.R., Cleveland-Innes, M., Koole, M., and Kappelman, J.** Revisiting Methodological Issues in Transcript Analysis: Negotiated Coding and Reliability. *The Internet and Higher Education* 9(1): 1–8 (2006).
 58. **Kanuka, H., Rourke, L., and Laflamme, E.** The Influence of Instructional Methods on the Quality of Online Discussion. *British Journal of Educational Technology* 38(2): 260-271 (2007).
 59. **Shea, P. and Bidjerano, T.** Learning Presence: Towards a Theory of Self-Efficacy, Self-Regulation, and the Development of a Communities of Inquiry in Online and Blended Learning Environments. *Computers & Education* 55(4): 1721–1731 (2010).
 60. **Shea, P., Hayes, S., Vickers, J., Gozza-Cohen, M., Uzuner, S., Mehta, R., Valchova, A., and Rangan, P.** A Re-Examination of the Community of Inquiry Framework: Social Network and Content Analysis. *The Internet and Higher Education* 13(1–2): 10-21 (2010).
 61. **Vaughan, N. and Garrison, D.R.** Creating Cognitive Presence in a Blended Faculty Development

- Community. *The Internet and Higher Education* 8(1): 1–12 (2005).
62. **Swan, K.** Building Learning Communities in Online Courses: The Importance of Interaction. *Education, Communication & Information* 2(1): 23–49 (2002).
 63. **Akyol, Z. and Garrison, D.R.** The Development of a Community of Inquiry over Time in an Online Course: Understanding the Progression and Integration of Social, Cognitive and Teaching Presence. *Journal of Asynchronous Learning Networks* 12: 3–22 (2008).
 64. **Ke, F.** Examining Online Teaching, Cognitive, and Social Presence for Adult Students. *Computers & Education* 55(2): 808–820 (2010).
 65. **Swan, K. and Shih, L.F.** On the Nature and Development of Social Presence in Online Course Discussions. *Journal of Asynchronous Learning Networks* 9(3): 115–136 (2005).
 66. **Gorsky, P., Caspi, A., Antonovsky, A., Blau, I., and Mansur, A.** The Relationship between Academic Discipline and Dialogic Behavior in Open University Course Forums. *International Review of Research in Open & Distance Learning* 11(2): 49–72 (2010).
 67. **Arbaugh, J.B.** Does the Community of Inquiry Framework Predict Outcomes in Online MBA Courses? *International Review of Research in Open and Distance Learning* 9(2): 1–21 (2008).
 68. **Arbaugh, J.B., Bangert, A., and Cleveland-Innes, M.** Subject Matter Effects and the Community of Inquiry (CoI) Framework: An Exploratory Study. *The Internet and Higher Education* 13(1–2): 37–44 (2010).
 69. **Bangert, A.** Building a Validity Argument for the Community of Inquiry Survey Instrument. *The Internet and Higher Education* 12(2): 104–111 (2009).
 70. **Díaz, S.R., Swan, K., Ice, P., and Kupczynski, L.** Student Ratings of the Importance of Survey Items, Multiplicative Factor Analysis, and the Validity of the Community of Inquiry Survey. *The Internet and Higher Education* 13(1–2): 22–30 (2010).
 71. **Shea, P. and Bidjerano, T.** Learning Presence as a Moderator in the Community of Inquiry Model. *Computers & Education* 59(2): 316–326 (2012).
 72. **Richardson, J.C. and Swan, K.** Examining Social Presence in Online Courses in Relation to Students' Perceived Learning and Satisfaction. *Journal of Asynchronous Learning Networks* 7(1): 68–88 (2003).
 73. **Shea, P., Pickett, A.M., and Pelz, W.E.** A Follow-up Investigation of “Teaching Presence” in the SUNY Learning Network. *Journal of Asynchronous Learning Networks* 7(2): 61–80 (2003).
 74. **Garrison, D.R., Anderson, T., and Archer, W.** The First Decade of the Community of Inquiry Framework: A Retrospective. *The Internet and Higher Education* 13(1–2): 5–9 (2010).
 75. **Gunawardena, C.N. and Zittle, F.J.** Social Presence as a Predictor of Satisfaction within a Computer Mediated Conferencing Environment. *American Journal of Distance Education* 11(3): 8–26 (1997).
 76. **Morris, T.A.** Exploring Community College Student Perceptions of Online Learning. *International Journal of Instructional Technology and Distance Learning* 8(6): 31–43 (2011).
 77. **Arbaugh, J.B. and Hwang, A.** Does “Teaching Presence” Exist in Online MBA Courses? *The Internet and Higher Education* 9(1): 9–21 (2006).
 78. **Rovai, A.P.** In Search of Higher Persistence Rates in Distance Education Online Programs. *The Internet and Higher Education* 6(1): 1–16 (2003).
 79. **Swan, K.** Relationships between Interactions and Learning in Online Environments. The Sloan Consortium, 2004.
 80. **Kupczynski, L., Ice, P., Weisenmayer, R., and McCluskey, F.** Student Perceptions of the Relationship between Indicators of Teaching Presence and Success in Online Courses. *Journal of Interactive Online Learning* 9(1): 23–43 (2010).
 81. **Dewey, J.** How We Think: A Restatement of the Relation of Reflective Thinking to the Educative Process. New York: D.C. Heath and Co., 1933.
 82. **Garrison, D.R. and Cleveland-Innes, M.** Facilitating Cognitive Presence in Online Learning: Interaction Is Not Enough. *The American Journal of Distance Education* 19(3): 133–148 (2005).

83. **Joo, Y.J., Lim, K.Y., and Kim, E.K.** Online University Students' Satisfaction and Persistence: Examining Perceived Level of Presence, Usefulness and Ease of Use as Predictors in a Structural Model. *Computers & Education* 57(2): 1654–1664 (2011).
84. **Kanuka, H. and Garrison, D.R.** Cognitive Presence in Online Learning. *Journal of Computing in Higher Education* 15(2): 21–39 (2004).
85. **Arbaugh, J.B.** An Empirical Verification of the Community of Inquiry Framework. *Journal of Asynchronous Learning Networks* 11(1): 73–84 (2007).
86. **Shih, T.-H. and Xitao Fan.** Comparing Response Rates from Web and Mail Surveys: A Meta-Analysis. *Field Methods* 20(3): 249–271 (2008).
87. **Lüdtke, O., Trautwein, U., Kunter, M., and Baumert, J.** Reliability and Agreement of Student Ratings of the Classroom Environment: A Reanalysis of TIMSS Data. *Learning Environments Research* 9(3): 215–230 (2006).
88. **Kozlowski, S.W. and Hatstrup, K.** A Disagreement About Within-Group Agreement: Disentangling Issues of Consistency Versus Consensus. *Journal of Applied Psychology* 77(2): 161–167 (1992).
89. **Maloney, M.M., Johnson, S.G., and Zellmer-Bruhn, M.E.** Assessing Group-Level Constructs under Missing Data Conditions: A Monte Carlo Simulation. *Small Group Research* 41(3): 281–307 (2010).
90. **Denson, N., Loveday, T., and Dalton, H.** Student Evaluation of Courses: What Predicts Satisfaction? *Higher Education Research & Development* 29(4): 339–356 (2010).
91. **Bedard, K. and Kuhn, P.** Where Class Size Really Matters: Class Size and Student Ratings of Instructor Effectiveness. *Economics of Education Review* 27(3): 253–265 (2008).
92. **NSSE.** *National Survey of Student Engagement Survey Instrument*, 2013. http://nsse.iub.edu/html/survey_instruments.cfm
93. **NSSE.** *Promoting Engagement for All Students: The Imperative to Look Within—2008 Results*. National Survey of Student Engagement Annual Results, 2008.
94. **NSSE.** *Assessment for Improvement: Tracking Student Engagement Over Time*. National Survey of Student Engagement Annual Results, 2009.
95. **NSSE.** *Major Differences: Examining Student Engagement by Field of Study*. National Survey of Student Engagement Annual Results, 2010.
96. **Cohen, A., Doveh, E., and Nahum-Shani, I.** Testing Agreement for Multi-Item Scales with the Indices rWG(J) and ADM(J). *Organizational Research Methods* 12(1): 148–164 (2009).
97. **Burke, M.J. and Dunlap, W.P.** Estimating Interrater Agreement with the Average Deviation Index: A User's Guide. *Organizational Research Methods* 5(2): 159–172 (2002).
98. **Kim, J., Kwon, Y., and Cho, D.** Investigating Factors That Influence Social Presence and Learning Outcomes in Distance Higher Education. *Computers & Education* 57(2): 1512–1520 (2011).
99. **Rourke, L. and Kanuka, H.** Barriers to Online Critical Discourse. *International Journal of Computer-Supported Collaborative Learning* 2(1): 105–126 (2007).
100. **Kanuka, H.** Interaction and the Online Distance Classroom: Do Instructional Methods Effect the Quality of Interaction? *Journal of Computing in Higher Education* 23(2): 143–156 (2011).
101. **Dunlap, W.P., Burke, M.J., and Smith-Crowe, K.** Accurate Tests of Statistical Significance for $r_{(WG)}$ and Average Deviation Interrater Agreement Indexes. *Journal of Applied Psychology* 88(2): 356–262 (2003).
102. **LeBreton, J.M., James, L.R., and Lindell, M.K.** Recent Issues Regarding rWG, rWG, rWG(J), and rWG(J). *Organizational Research Methods* 8(1): 128–138 (2005).
103. **Roberson, Q.M., Sturman, M.C., and Simons, T.L.** Does the Measure of Dispersion Matter in Multilevel Research? A Comparison of the Relative Performance of Dispersion Indexes. *Organizational Research Methods* 10(4): 564–588 (2007).

104. **Bonk, C.J. and Cunningham, D.J.** Searching for Learner-Centered, Constructivist, and Sociocultural Components of Collaborative Educational Learning Tools. In: Bonk, C.J. and King, K.S. (Eds.) *Electronic Collaborators: Learner-Centered Technologies for Literacy, Apprenticeship, and Discourse* Mahwah, NJ: Lawrence Erlbaum, 25–50, 1998.
105. **Hirumi, A.** Student-Centered, Technology-Rich Learning Environments (Scentrle): Operationalizing Constructivist Approaches to Teaching and Learning. *Journal of Technology and Teacher Education* 10(4): 497–537 (2002).
106. **Yammarino, F.J., Dionne, S.D., Schriesheim, C.A., and Dansereau, F.** Authentic Leadership and Positive Organizational Behavior: A Meso, Multi-Level Perspective. *The Leadership Quarterly* 19(6): 693–707 (2008).
107. **Lindell, M.K., Brandt, C.J., and Whitney, D.J.** A Revised Index of Interrater Agreement for Multi-Item Ratings of a Single Target. *Applied Psychological Measurement* 23(2): 127–135 (1999).
108. **Schmidt, F.L. and Hunter, J.E.** Interrater Reliability Coefficients Cannot Be Computed When Only One Stimulus Is Rated. *Journal of Applied Psychology* 74(2): 368–370 (1989).
109. **Burke, M.J., Finkelstein, L.M., and Dusig, M.S.** On Average Deviation Indices for Estimating Interrater Agreement. *Organizational Research Methods* 2(1): 49–68 (1999).
110. **Bliese, P.** Within-Group Agreement, Non-Independence, and Reliability: Implications for Data Aggregation and Analysis. In: Klein, K.J. and Kozlowski (Eds.) *Multilevel Theory, Research, and Methods in Organizations: Foundations, Extensions, and New Directions*. Jossey-Bass: San Francisco, CA, 349–381, 2000.
111. **Newman, D.A. and Sin, H.-P.** How Do Missing Data Bias Estimates of Within-Group Agreement? Sensitivity of SD_{WG} , CV_{WG} , $R_{WG(j)}$, $R_{WG(j)}^*$, and ICC to Systematic Nonresponse. *Organizational Research Methods* 12(1): 113–147 (2009).

VIII. TECHNICAL APPENDIX

The most popular statistic for measuring organizational climate was proffered by James, Demaree, and Wolf [38], who developed r_{WG} , in which within-group variance is divided by a hypothetical variance that would result if there were no consensus among raters (the null hypothesis), and therefore all variance were due to measurement error. The formula for $r_{WG(j)}$ is:

$$r_{WG(j)} = \frac{J \left(1 - \frac{\bar{S}_{x_j}^2}{\sigma_E^2} \right)}{J \left(1 - \frac{\bar{S}_{x_j}^2}{\sigma_E^2} \right) + \frac{\bar{S}_{x_j}^2}{\sigma_E^2}}$$

where J is the number of items in a scale, $\bar{S}_{x_j}^2$ is the mean observed variance for J items, and σ_E^2 is the variance expected when there is a complete lack of agreement, or error variance. In some situations, this hypothetical error variance may be uniform across all response options; that is, evaluators using a 1–5 scale are as likely to select the number 1 as they are 2, 3, 4, and 5. For scales with multiple items, James, Demaree, and Wolf recommended the $r_{WG(j)}$ statistic, which combines the variances of each item in the scale. If there is reason to believe that systematic response bias, such as leniency or severity error, would affect ratings, then the hypothetical error variance would differ from a uniform distribution, and can be calculated. They calculated null variances to use for common skewed distributions for multi-item measures, and LeBreton and Senter added additional variances for a wider range of scale types [34].

This statistical index is widely favored because it has an easily interpretable range, from 0 to +1.0, where

0 indicates no agreement between raters and 1.0 indicates perfect agreement. This index measures true consensus, or agreement both on a relative and absolute basis, rather than consistency or reliability, which refer to agreement only on a relative basis [34, 87, 88, 107]. For example, if one group member responds to six items with scores of 5, 4, 5, 4, 4, and another responds to the same items with scores of 2, 1, 2, 1, 1, there is complete consistency in their use of the scale; therefore, the inter-rater reliability as measured by correlation would be 1.0. However, the ratings are nearly opposites, and there is little actual consensus between the raters; the $r_{WG(j)}$ would be very low.

Despite its popularity, this statistic has been critiqued on several grounds, including the notion that it does not assess inter-rater reliability as claimed [108]. It also requires that the hypothetical null distribution be specified, based on knowledge of likely survey response bias [34, 38, 103]. In addition, it can produce uninterpretable numbers below 0.0 and above 1.0 when the observed variance is larger than the hypothetical null distribution; this is particularly likely if a strong skew is used for the null distribution, if the scores are bimodal, or if there are many items in a scale [27, 107]. Adding items to a scale increases the internal consistency of $r_{WG(j)}$ because it corrects for attenuation [34].

The recommended way to deal with uninterpretable values—greater than 1.0 or less than 0.0—is to replace them with the value 0.0, which indicates no consensus among the group members [34, 38, 39]. In an elegant simulation, Roberson and her colleagues found that nearly all indices of consensus or dispersion are strongly intercorrelated for single-item measures, with correlations between .93 and .98 [103].

Many researchers have used r_{WG} or $r_{WG(j)}$ to determine whether perceptions of different people about organizational climate are similar enough to aggregate [14, 17, 21, 22, 23, 32, 36, 37]. Some provide analyses using multiple comparison null distributions, including both uniform and skewed distributions; skewed distributions tend to produce lower $r_{WG(j)}$ scores, but are conceptually more accurate because response bias is such a common phenomenon with survey data [34]. There is no clear agreement on the optimal distribution for a null hypothesis, nor a smallest sample size that should be used to assess a group score, nor a means of dealing with missing data. Recent analyses have recommended a graduated standard for acceptable levels, rather than a uniform cutoff: .00–.30 = lack of agreement; .31–.50 = weak agreement; .51–.70 = moderate agreement; .71–.90 = strong agreement; and .91–1.00 very strong agreement [34, 39].

Another means of assessing agreement is average deviation (AD), which is the absolute value of the difference between each item and the mean or the median; either can be calculated [34, 103]. Because this index is not squared, it is not unduly sensitive to outliers. However, it reflects the numbering in the scale, and measures dispersion so larger numbers indicate more disagreement within groups [103]. The formula for AD is:

$$AD_{M(j)} = \frac{\sum_{k=1}^K |X_{jk} - \bar{X}_j|}{K}$$

where X is the score on item j , \bar{X}_j is the average score on item j , K is the number of judges, and $k=1$ to K judges. $AD_{M(j)}$ is the mean of the $AD_{M(j)}$ of all items in the scale. The $AD_{M(j)}$ has been used to measure climate strength [103] and to assess the reliability of climate perceptions, both in organizations [14] and in schools [87]. The $AD_{Med(j)}$ is calculated based on medians rather than means, and is more robust as it is less influenced by outliers [34, 109].

In addition to measuring consensus, research on organizational climate also assesses the relative consistency of perceptions by examining the overall within-group and between-group variation [34]. Two

common indices are based on analysis of variance: intraclass correlations 1 (ICC(1)) and 2 (ICC(2)). The ICC(1) measures the proportion of overall variance that reflects between-group differences; it can be considered the effect size attributable to the group alone, or the extent to which ratings are affected by the group to which one belongs [34]. The formula is:

$$ICC(1) = \frac{\tau^2}{\tau^2 + \sigma^2}$$

where τ^2 is the variance between groups, and σ^2 is the variance within groups [87, 110]. Values of .10 to .24 are considered medium, and of .25 or higher are considered large; even small numbers can indicate that group membership affects scores [34].

ICC(2) measures the reliability of group mean scores [87, 110], with high scores indicating that mean (aggregated) scores of the groups are relatively free from variation among the members [111]. The formula is:

$$ICC(2) = \frac{K \times ICC(1)}{1 + (K - 1) \times ICC(1)}$$

where K is the number of judges; the average is used if groups have different sizes. A graduated set of reliability standards is recommended for evaluating it, although the traditional reliability cutoff of .70 is often used [34]. Several scholars have used the ICC(1) and ICC(2) to validate the presence of group-level climate in organization [14, 17, 22, 23, 30, 36]; their use has been recommended as a matter of routine in research on classroom agreement [87].

Others use the simple F-statistic in one-way analysis of variance to determine whether individual climate perceptions can be aggregated into group-level measures, which compares between-group variance with within-group variance. This statistic may be problematic if there is restriction of range or when comparing groups across different organizations [27, 38]. It has frequently been used to provide evidence of between-group differences on climate, indicating that climate is specific to each group [14, 22, 23, 26, 30, 32, 37].

IX. ACKNOWLEDGEMENTS

The authors would like to thank James Moore, Maria Avgerinou, Rebecca Valeriano-Flores, and Christoph Bakke for their invaluable contributions to the project.

X. ABOUT THE AUTHORS

Beth Rubin is an assistant professor and the Director of the online program at the School for New Learning (SNL) at DePaul University. She holds an MA and PhD in Industrial/Organizational Psychology from Michigan State University and has been an instructor, course developer, and administrator of online learning programs since 2000. She has led online instruction, curriculum, and operations at both for-profit and nonprofit universities, and her research focuses on the effect of learning technologies and organizational practices that support effective online learning.

Ron Fernandes is a faculty member and Assistant Director of the School of Public Service (SPS) at DePaul University. He is the founder and Director, since 2007, of the SPS Online program and has taught online since 2003. He has a PhD in Public Policy and Management from Carnegie Mellon University. His research interests include understanding the critical features of online courses that enable learning and the attributes of quality online learning experiences. He is a Master Reviewer for Quality Matters: an interinstitutional, faculty-centered, peer review process to certify the quality of online and blended courses.

