Volume 16, Number 1, Winter 2018
ISSN: 1541-4914

A Comparison of Independent and Collaborative Instructional Models in a Blended Graduate Teacher Education Program

Jason T. Abbitt, Sarah J. Watt, and William J. Boone
Miami University

Abstract

This evaluation research study investigated the impact of differences in independent and collaborative teaching models in a blended graduate teacher education program. Using the Community of Inquiry model as a framework to evaluate the student experience in this program, data were collected over a two-year period regarding student perceptions of cognitive, social, and teaching presence. One unique aspect of this study is the Rasch measurement techniques that produced scaled measurements of cognitive, social, and teaching presence. These measurements were subsequently used to compare courses in which both independent (single instructor) and collaborative teaching models were present. A discussion of the results of these analyses is provided that highlight models that are promising when examined across the entire program. Also, specific forms of collaborative teaching are discussed that were found to impact student ratings of their experience in online courses throughout the academic program.

When examining online learning from a broad perspective, there is a distinct trend of growth as the number of enrollments in online and hybrid learning in higher education has steadily risen. A 2018 annual report on distance education noted that enrollments in distance learning have risen consistently for the past 14 years, the majority are undergraduate students, many of whom also enroll in on-campus courses (Seaman, Allen, & Seaman, 2018). According to the 2017 report "The Changing Landscape of Online Education", the most prominent strategic priorities driving online programs is enrollment growth and student completion gains (Legon & Garrett, 2017, p. 24). The report also highlighted a higher degree of growth in distance graduate enrollment from 2012-2015 with 20% growth in graduate programs as compared to 6% growth in undergraduate programs.

In response to the overall trend of growth in online learning, it is also important to closely examine how these new learning programs emerge and evolve in order to better understand those practices that are most effective. In a report of a project focused on innovation in higher education, Magda and Buban (2018) explained that "promoting student success appears to be the chief focus of innovation initiatives, with 68 percent of respondents ranking student success as a top- three goal for innovation" (p.7), despite the fact that there is little consistency in how institutions define the term innovation. The report also noted that structural issues and cultural factors were identified as top barriers to innovation and that "these barriers can be overcome by strong leadership-shaping processes to better promote collaboration, as well as

rewards and incentives to encourage shifts in culture" (p. 7). Although innovation can emerge in many aspects of higher education, the success of online learning programs is certainly impacted by the support for, and barriers to, innovation at an institution.

Within this broad and growing field of online learning, it is imperative that models of teaching and learning are closely examined through research in an effort to improve student learning outcomes. Each time an online course is offered, a certain degree of variability is common. Variations in class sizes, student and instructor personalities, teaching and learning styles, prior experience and expertise, as well as innumerable other differences can impact the overall student experience in higher education courses. In an academic program that is offered in a blended or fully online format, where students are separated from the instructor and other students by physical space and time, these variations can have a direct influence on the degree to which students are engaged in learning activities and connected with their peers and those teaching the courses. Of specific interest in this study was whether variations in the number of instructors assigned to teach courses, as well as the degree to which some instructors chose to teach collaboratively, impacted student perceptions regarding the learning experience in a blended graduate teacher preparation program.

Research Context and Focus

The current study was conducted within the context of a blended graduate teacher preparation program in special education at a mid-sized public institution in the midwestern United States. Courses within this program meet the definition of *Blended Learning* as described by the Online Learning Consortium (2015) as each course features online coursework throughout an eight to 12-week academic term as well as 90-minute face-to-face class sessions at the beginning, middle, and near the end of a term. Within this academic program, some of the courses that were offered during the study period were taught by a single instructor, whereas other courses were taught by a team of two instructors. Additionally, some instructors taught two sections of the same course in the same academic term. As a result, there were three primary variations in the instructional model observed: (1) single instructor teaching a single course section, (2) single instructor teaching two combined course sections, and (3) two instructors co-teaching two course sections.

The number of course sections that are available during a typical academic term, however, varied during the study period based on the enrollment. The typical class size is 25 students for a course section for which a single instructor is assigned. When enrollment was expected to be above 35 students, two sections of a course were offered. The assignment of instructors to teach these courses was done according to multiple factors as defined by departmental governance policy, including (1) subject matter expertise, (2) experience and/or qualifications related to teaching online courses, and (3) instructor availability and teaching load. As a result of the effort to balance these factors, the manner in which instructors were assigned to courses was variable. When two sections of a course were offered in the same term, a single instructor was sometimes assigned to teach both sections; whereas in other cases, two different instructors were assigned to each of the course sections. When two different instructors were assigned to teach the same course, the manner in which the instructors chose to collaborate (or not) in teaching the different sections led to additional variability in the teaching model with some instructors choosing to teach collaboratively and some instructors teaching independently. Additionally, some sections of courses were combined, while others remained separate.

The overall outcome of these practices has been that, within this graduate degree program, several different models of independent and collaborate teaching have been

implemented. All courses in the program were offered at least twice during the study period, thus an opportunity existed to compare these different instructional models: such as team teaching vs. independent teaching of the same course as well as a single section of a course vs. combined sections of the same course. A student-reported measure was used to compare these different instructional models that included ratings of social, teaching, and cognitive presence as defined by the Community of Inquiry model (Arbaugh et al, 2008; Garrison, Anderson, & Archer, 2000; Garrison & Arbaugh, 2007) and measured by latest version (draft 14b) of the Community of Inquiry survey instrument (Arbaugh et al., 2008; Swan et al., 2008).

The focus of this study was to investigate the degree to which the four variations in the instructional model impacted the student perceptions of the quality of learning experience. As previously noted, these different teaching models were the result of both administrative policy for teaching assignments as well as instructor preference for collaboration. Because all courses were offered multiple time during the two-year study period, however, it was possible to examine the impact of these variations across the entire program as well as in specific courses where the subject matter and online course design were consistent.

Review of Relevant Literature

Collaborative Teaching Models

Collaborative approaches to providing instruction to students in one course, or co-teaching (Cook & Friend, 1995), is a relatively common practice in K-12 settings. Co-teaching is defined by having at least two individuals working to design and plan instruction for a specific course or content area. This collaborative model may be established in a variety of ways. Friend, Cook, Hurley-Chamberlain, and Shamberger (2010) define six approaches to co-teaching: (1) one teaches, one assists, (2) teaming, (3) station teaching, (4) parallel teaching, (5) alternate teaching, and (6) one teaches, one observes. While the one teaches, one observes or assists models of co-teaching are most prevalent within the research literature, findings indicate this approach to collaborative teaching has little effect on student achievement (Solis, Vaughn, Swanson, & Mcculley, 2012). Teaming, station teaching, parallel teaching, and alternative teaching allow for a more team-based approach to lesson design and implementation. Teaming is when both teachers are present in the course at the same time and equally share responsibility. Station teaching also requires both teachers to be in the same space but they are individually instructing small groups of students at one time. Parallel teaching is when two teachers are designing and planning in tandem but delivering the content to separate groups or classes. Alternative teaching is when both teachers provide instruction in the same space or course, but one teacher takes the lead while the other supports a small subset of students requiring more individualized attention.

Co-Teaching Models in Higher Education

While co-teaching is a common practice across K-12 settings, particularly when supporting students with disabilities in inclusive settings, co-teaching is not as prevalent in higher education. Co-teaching that is established at the college or university level, is typically developed for an alternative purpose such as interdisciplinary teaching, mentoring and modeling, and providing a more diverse set of pedagogical approaches (Krometis, Clark, Gonzalez, & Leslie, 2011; Morelock et al., 2017; Walters & Misra, 2013). Despite the large research base for

co-teaching at the K-12 level, limited research exists on the effects of co-teaching college and university level courses exists.

Jones and Harris (2012) investigated student and instructor perspectives of a multiple instructor, or co-teaching, model used within undergraduate science courses. Students reported that they preferred the increase in expertise from multiple instructors and the varied teaching and learning experience. On the contrary, the students reported that a disadvantage of this model was the need to frequently adjust to the differences between instructors' teaching style. Instructors of the co-taught classes identified diverse expertise and perspectives as the greatest advantage of co-teaching. More recent research examining co-teaching models in higher education by Morelock et al. (2017) paralleled these findings suggesting that the additional expertise within co-teaching pairs allows for increased diversity of thought. While both of these studies highlight advantages and disadvantages related to the overall learning experience for students, they do not investigate the effects of the models on student achievement.

While there is little research on the effectiveness of co-teaching within higher education, some benefits of increased expertise, diverse perspectives, and improved student experience have been noted in the literature. In addition, some key components and possible pitfalls of collaborative teaching in online learning have also been outlined within the small research base. Strohschen and Heaney (2000) studied the impact of a team approach in a graduate online course and identified collaborative preplanning for the course as an essential component of team teaching online. Additionally, the authors also provide a description and reflection of a power-struggle revealed in complex interactions between the team of instructors as the course progressed. The importance of collaborative planning is echoed by Fuller and Bail (2011) who conducted an action research study over five semesters of an online graduate course on Disaster Medicine and Management Applied Research and Statistics. The results of this study highlighted the importance of the instructors' collaboration in course planning and the benefits of team teaching for continued instructor learning of pedagogical approaches as well as subject matter expertise. The effect of team teaching on student attitudes towards the course was that students in the course perceived "greater instructor presence in this online course than in other similar online experiences" (p. 79). Although the benefits and challenges of collaborative teaching approaches are similar in both online and traditional course environments, additional research examining these teaching approaches within online education would strengthen the limited research base while also providing insight relevant to the academic policies related to online courses at the postsecondary level.

The Community of Inquiry Framework

The Community of Inquiry framework was first described by Garrison, Anderson, and Archer (2000) as a conceptual model for the essential characteristics for effective online learning experiences that includes three primary, though interconnected, forms of presence. Derived from research examining text-based interaction among students and instructors in online courses, the Community of Inquiry framework includes three essential elements of asynchronous learning networks: teaching presence, social presence, and cognitive presence. In further description of these elements, Garrison et al. (2000), describe multiple categories of events as well as indicators of these three elements that were used in coding and analyzing narrative data to further investigate and refine the Community of Inquiry framework.

Perhaps the most succinct and clear definitions of each of the three elements of the Community of Inquiry framework is from Garrison and Arbaugh (2007). The following excerpts

are the descriptions provided in that article describing the current state, as of 2007, and future directions of the framework:

- Social presence is "the ability of learners to project themselves socially and emotionally, thereby being perceived as 'real people' in mediated communication" (p. 159).
- Teaching presence is "the design, facilitation, and direction of cognitive and social processes for the purpose of realizing personally meaningful and educationally worthwhile learning outcomes" (p.163).
- Cognitive presence is "the extent to which learners are able to construct and confirm meaning through sustained reflection and discourse" (p.161).

Although the literature concerning the Community of Inquiry framework describes these elements of presence as unique, they are also interconnected. In the context of an online course, the design and facilitation of course activities (part of teaching presence) will foster higher levels of social presence if an instructor or course designer has chosen to include course activities in which students interact with their peers. Further, the degree to which these activities are connected to specific learning outcomes will influence the level of cognitive presence that a student might experience.

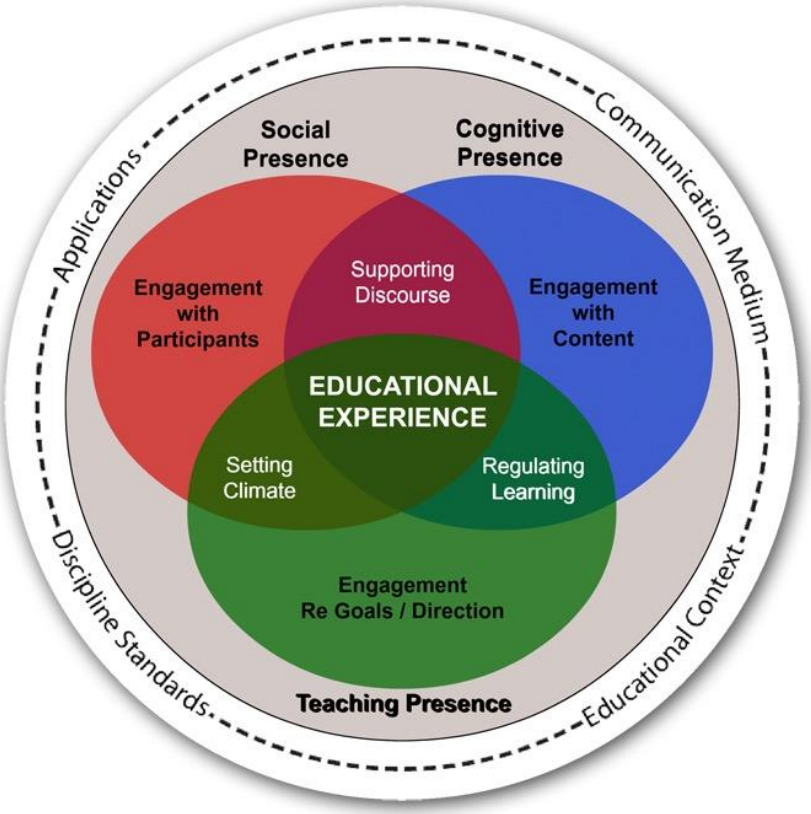

Figure 1. A visual depiction of the Community of Inquiry Model from http://www.thecommunityofinquiry.org/coi and reproduced under Creative Commons license (CC-BY-SA)

Garrison, Cleveland-Innes, and Fung (2010) suggest through their research exploring student perceptions related to Community of Inquiry that teaching presence is key to establishing a Community of Inquiry within online courses. Although Fuller and Bail (2011) identified increased perceptions of instructor presence as a positive outcome of their study of team teaching, there is insufficient research regarding the impact of collaborative teaching models on student perceptions on other facets of online learning as described by the Community of Inquiry framework.

Research Questions

The following research questions were investigated in order to provide a basis for investigating the impact of differences in the teaching models on measures from the Community of Inquiry Survey:

1. To what degree are differences found in student measures of teaching presence (TP), cognitive presence (CP), and social presence (SP) among all courses in the academic program in which various teaching models were used?
2. To what extent are differences in student measures of TP, CP, SP found when comparing individual courses in which different teaching models were used?

Methodology

Data for this study were collected over a two year period in the context of an evaluation of the student experience in a blended graduate program in special education. This academic program provides a total of twelve 3-credit graduate level courses leading to licensure as an Intervention Specialist and/or a Master of Education degree. Enrollment for each section of a course within the program during the two year period ranged between five and 51 students with a mean enrollment of 20.8 students per section.

The program that was being evaluated is cohort-based. Although there is no specific requirement for students to remain within a specific cohort group, most students progress through the program on a similar timeline due to course availability and scheduling. Students are admitted in either fall or spring semesters and most students enroll in two courses per term with other students who were admitted at the same. A small number of students choose to complete courses over a more extended time period and may enroll in only one course in some academic terms. Students who are pursuing only licensure will typically complete the program within a single academic year, whereas students seeking licensure and the M.Ed. degree can complete the program in 18-24 months.

Instrument and Procedures

The Community of Inquiry survey was first described in research literature in 2008 by a research team (Arbaugh et al., 2008; Swan et al., 2008) with the goal of creating a "structurally valid and psychometrically sound survey instrument with the potential to expand the study of online and blended learning" (Arbaugh et al., 2008, p. 134). In this initial study of the instrument, the authors administered the survey to graduate students in education and business programs at four institutions, from which 287 students participated in the study. The survey included 34 items that were derived from the multiple elements, categories, and indicators identified in prior research concerning the Community of Inquiry framework. A Principal Components Analysis of the responses revealed a three-factor solution consistent with the three elements of the Community of Inquiry framework (TP, SP, and CP). The survey has been the

focus of additional research and further refinement since the initial work published in 2008. Subsequent research has supported the validity of the Community of Inquiry survey as a measure of teaching, social, and cognitive presence (Bangert, 2009; Boston et al., 2009; Díaz, Swan, Ice, & Kupczynski, 2010; Garrison, Cleveland-Innes, & Fung, 2010; Rubin & Fernandes, 2013). The survey has also been used to evaluate online learning issues such as student retention (Boston et al., 2009), student engagement (Nagel & Kotzé, 2010; Shea & Bidjerano, 2009; Vaughan, 2010), as well as elements of course and program design (Akyol, Vaughan, & Garrison, 2011; Kumar, Dawson, Black, Cavanaugh, & Sessums, 2011; Kumar & Ritzhaupt, 2014; Richardson et al., 2012; Russell & Curtis, 2013; Vaughan, 2010) in addition to many other research studies in online and blended learning.

In the context of this study, the Community of Inquiry survey instrument (Arbaugh et al., 2008) was used to measure SP, CP, and TP. The survey was administered at the approximate midpoint of each academic term and was open for participation for approximately seven days. The survey was made available using the Qualtrics Survey system and an email invitation was sent via Qualtrics to all enrolled students to solicit participation in the study. Participation in the study was not required, nor was an incentive offered for participation. Follow-up invitation reminders were sent every two days to participants who had not yet participated. The survey procedure was reviewed and approved by the appropriate Institutional Review Board. As shown in Table 1, between 53.6% and 76.1% of students responded to the survey during each academic term during the study period.

Table 1
Response rate by term for Community of Inquiry survey

Academic Term	Invitations	Responses	Response Rate
Fall 15-16	91	66	72.5%
Winter 15-16	94	67	71.3%
Spring 15-16	84	50	59.5%
Summer 15-16	96	66	68.8%
Fall 16-17	88	67	76.1%
Winter 16-17	75	45	60.0%
Spring 16-17	85	52	61.2%
Summer 16-17	95	51	53.6%

The Community of Inquiry survey items focus on a respondents attitudes in reference to a specific course, thus the data collected also included a variable that identified the course in which the responses were collected. After the data collection was complete, the course identifier variable was used to link additional variables to the dataset regarding the teaching model that was used in that course. This variable indicated whether the course was taught by a single instructor or a teaching team. For those courses taught by a single instructor, a variable was utilized to indicate whether the instructor was teaching one section of the course or teaching two sections of the course that were combined. Table 2 provides an overview of the number of

responses from students in courses where different independent or co-teaching models were present.

Table 2
Frequency of survey responses from courses using each instructional model

Type	n	%
Single Instructor, 1 course section	203	28.8%
Single Instructor, 2 course sections combined	286	40.6%
Co-Teaching	133	18.9%

Rasch Person-Measures

One unique aspect of this study is the Rasch analysis that was used to calculate the students' measures of TP, SP, and CP from the survey responses. Previous studies using the Community of Inquiry survey have typically calculated these measures by first coding the responses to the Likert-scale items as 1=strongly disagree, 2=disagree, 3=neutral, 4=agree, and 5=strongly agree. Three measures (TP, SP, and CP) are then calculated using these coded numeric values by calculating a mean of these values. For example, a mean value of the numerically coded responses for items from the CP scale would be used to represent a respondent's CP, thus the possible measurements ranged from one (indicating a respondent selected strongly disagreed to all items in the subscale) to five (indicating a respondent selected strongly agree to all items in the subscale). While preparing for data analysis for this study, however, a review of the literature regarding the development, validation, and early use of this instrument (Arbaugh et al., 2008; Bangert, 2009; Díaz et al., 2010; Swan et al., 2008) revealed that this particular calculation is not explicitly discussed.

Although much of the prior research on the Community of Inquiry instrument used common methods for evaluating subscale measurements (e.g. factor analysis), there are multiple concerns with such procedure for calculating subscale scores that this current study sought to avoid. Among the most prominent concerns is that the use of numerically coded responses to represent data that is categorical in nature (Likert-scale responses) and are non-linear. However, it is often the case with the Community of Inquiry instrument that these numeric values (1=strongly disagree, 2=agree, etc.) are analyzed in a manner that is more appropriate for more precise linear scale measurements. These data have ordinal properties that indicate the relative position of responses but lack the interval properties or more precise measurement scales, thus the common procedure of using a mean value of all items on a scale to represent a measure is inappropriate for this type of data. Researchers often make the incorrect assumption that the difference between, for example, a response of "1" (strongly disagree) and a "2" (disagree) is the same difference as between a "4" (agree) and a "5" (strongly agree). All that can be accurately stated is the for any specific item is that a response of strongly agree means more agreement than agree, a agree means more agreement than neutral, and so on. An additional limitation of the procedure of calculating a measure (e.g. CP) as a mean of the coded responses is that all of the items in a subscale are weighted equally in this calculation. This means a response of agree to a

survey item that is very easy to agree with is treated as having the same meaning as an answer of agree to an item that is very difficult to agree to.

Because of these limitations with the computation of a mean rating using raw scores (e.g. Participant A has an average response of 3.5 for the CP scale), the authors of this study chose to utilize a Rasch analysis to calculate what is named a *person measure* for each of the subscales (TP, SP, and CP). The Rasch analysis and calculation of the person measure are specially designed for non-linear rating scale data. Rasch analysis uses the Likert-scale responses from the survey with coded using values one to five, with a value of one corresponding to a response of strongly disagree and five corresponding to a response of strongly agree. What is distinctly different, however, is that the Rasch model is utilized to address issues such as the non-linearity of rating scales and also take into consideration the fact that all survey items are not equally agreeable. Details of the Rasch model are presented in the seminal works *Best Test Design* (Wright & Stone, 1979) and *Rating Scale Analysis* (Wright & Masters, 1982). It should be noted, however, that the selection of a different process for calculating subscale measures is not necessarily intended to suggest that prior research is flawed, but rather the Rasch analysis process is more suited to the type of data that was collected and is sensitive to the nuances of the ordinal rating scale.

The Rasch analysis of this data set was conducted using the widely used WinSteps (Linacre, 2018) Rasch software. The student responses to items for each subscale were analyzed and Rasch person measure was calculated for each respondent. The Rasch person measures are often expressed using the "logit" (log odds) scale. For this study, once the person measures in logits were computed for each respondent for the CP, TP and SP scale, the logits were converted to a scale from 0-1000. This is a similar step taken for studies such as "Programme for International Student Assessment" (OECD, 2012) and "Trends in International Mathematics and Science Study" (Yamamoto & Kulick, 1999), thus it is these person measures which were used for the subsequent statistical analyses presented in this paper.

The Rasch person measures express where each respondent is located on the variables of TP, SP, and CP. A respondent with a CP score of 600, for example, represents an individual for whom the perceived TP in a course was higher than that of someone with a TP score of 550. Higher person-measure values represent a higher degree of the trait and vice-versa. It is also important to note that, although the three subscales of TP, SP, and CP were scaled in a similar manner (0-1000), these resulting person-measures are only comparable with other measures on the same scale. It is acceptable to compare, for example, two respondents person-measures on the CP scale, but a person measure on CP variable should not be directly compared to a measure on the SP, or TP scale. Although a more in-depth description of the Rasch analysis is beyond the scope of this article, readers are encouraged to consult Boone, Townsend, and Staver (2011), Boone, Staver, and Yale (2014), or Bond and Fox (2013) for further information regarding this procedure.

Results

Program-Level Comparison of Teaching Models

The first analysis that was conducted focused on comparing teaching models that differed across the academic program. To that end, the measures of TP, CP, and SP were compared among courses in which various teaching models were used. In order to select the appropriate inferential analysis for a comparison, the person-measure for each of the subscales were first evaluated for normality using a Shapiro-Wilks test. As shown in Table 3, when examined by group, each of the variables was found to be non-normal, thus a non-parametric Kruskal-Wallis test was selected as the appropriate analysis to compare these values across each category of course type.

Table 3
Mean Values and Shapiro-Wilks Tests of Normality for Subscale Person-Measures

				Test of Normality		
Subscale	Course Type	Mean	SD	W	df	Sig.
CP	Single Instructor, 1 course section	643.8	213.2	.919	203	.000
	Single Instructor, 2 course sections combined	610.8	189.8	.937	286	.000
	Co-teaching	709.4	212.6	.931	133	.000
TP	Single Instructor, 1 course section	681.1	215.8	.917	203	.000
	Single Instructor, 2 course sections combined	620.8	215.7	.951	286	.000
	Co-teaching	761.3	215.7	.895	133	.000
SP	Single Instructor, 1 course section	608.9	172.8	.906	203	.000
	Single Instructor, 2 course sections combined	589.7	156.7	.945	286	.000
	Co-teaching	678.9	202.6	.909	133	.000

Note. CP = cognitive presence, TP = teaching presence, and SP = social presence.

The results of the Kruskal-Wallis ANOVA analysis shown in Table 4 suggest that a significant difference was observed. These results indicate that some of the pairwise comparisons of the subscale person-measures were found to be significantly different between the instructional models on each of the subscale measures.

Table 4
Kruskal-Wallis ANOVA results for comparison of each subscale by instructional model used.

Subscale	Test Statistic	Sig.
CP	21.80	<.001
TP	36.80	<.001
SP	16.55	<.001

Note. n = 622, df = 2. CP = cognitive presence, TP = teaching presence, and SP = social presence.

Because these differences were observed, the Pairwise Comparisons table in SPSS was reviewed to evaluate the significance of each comparison between the different instructional models. Table 5 includes the results of the pairwise comparison on the TP measure, all pairwise comparisons were found to be significant with the largest difference found between courses with one instructor teaching two combined sections and those courses in which a co-teaching model was present.

Table 5
Pairwise Comparisons of instructional models on Teaching Presence Person-Measures

Category 1	Category 2	Test Statistic	*p* (adj.)*
1 Instructor, 1 course section	Co-teaching	-65.05	.003
1 Instructor, 2 course sections combined	Co-teaching	-112.97	<.001
1 Instructor, 1 course section	1 Instructor, 2 course sections combined	47.92	.011

** Significance values have been adjusted using the Bonferroni correction for multiple tests*

Regarding the SP variable, two of the three pairwise differences were found to be significant. The largest difference was again found when comparing courses with one instructor teaching two combined sections to those courses in which a co-teaching model was present (H=-75.77, p_{adj}<.001). The comparison between courses in which one instructor was teaching a single course section and those in which a co-teaching model was present was also significant (H=-59.98, p_{adj}=.008). The comparison between courses with one instructor (two combined sections versus one single section) was not statistically significant on the measure of SP.

Table 6
Pairwise Comparisons of instructional models on Social Presence Person-Measures

Category 1	Category 2	Test Statistic	*p* (adj.)*
1 Instructor, 1 course section	Co-teaching	-59.98	.008
1 Instructor, 2 course sections combined	Co-teaching	-75.77	<.001
1 Instructor, 1 course section	1 Instructor, 2 course sections combined	15.79	1.0

** Significance values have been adjusted using the Bonferroni correction for multiple tests*

A pattern similar to that found with the SP measurement was also observed in an analysis of the CP scale. The largest difference was found on this measure when comparing courses with one instructor, two sections combined to those courses in which a co-teaching model was present (H=-87.18, p_{adj}<.001). The comparison between courses in which a single instructor was teaching a single course section and those in which a co-teaching model was present was also significant (H=-65.12, p_{adj}=.003). As with the SP scale, the comparison between courses with one instructor (two combined sections versus one single section) was not statistically significant on the measure of CP.

Table 7
Pairwise Comparisons of instructional models on Cognitive Presence Person-Measures

Category 1	Category 2	Test Statistic	*p* (adj.)*
1 Instructor, 1 course section	Co-teaching	-65.12	.003
1 Instructor, 2 course sections combined	Co-teaching	-87.18	<.001
1 Instructor, 1 course section	1 Instructor, 2 course sections combined	22.06	.534

** Significance values have been adjusted using the Bonferroni correction for multiple tests*

Course-Level Comparison of Teaching Models

In addition to the program-level comparison of teaching models, this study also explored the Community of Inquiry measures in which additional comparisons could be made to investigate evidence of the impact of specific co-teaching models. As previously stated, the specific teaching model for courses was impacted by administrative decisions regarding teaching assignments and by the choices of assigned instructors to co-teach or not. As such, the courses selected for further analysis were those in which a reasonable comparison could be made between different sections of the same course, but where there were differences in the instructional model utilized.

For this additional targeted analysis, three courses were identified where a single instructor taught two combined sections of a course during a single academic term. Then, in a later academic term, a team of two instructors used a co-teaching model for two combined sections of the same course. In all three courses, the use of a shared online course model resulted in online course material and activities that were similar between each section of the course, regardless of the instructional model. In two courses, identified in Table 8 as Course A and Course B, the instructor who taught the course alone was also a member of the co-teaching team. Also, in Course A and Course B, the instructors used a parallel co-teaching model. In Course C, a similar change occurred from a single instructor to a co-teaching teach. In this course, however, the instructors in the co-teaching team did not include the previous instructor. Also, in Course C, an alternate co-teaching model was used in the second academic term.

Each test was conducted separately and the data were evaluated for normality. In all cases, the data were found to have a non-normal distribution. Additionally, the size of the sample was small in several groups. Given these characteristics of the dataset, a non-parametric Mann-Whitney test was selected to evaluate the difference between the subscale student measures for each type of course. Table 8 includes the mean and standard deviation of the Rasch person-measures for each of the subscales, as well as the mean difference between two instructional models. Although the mean difference value is not directly evaluated by the Mann-Whitney U

test, it is included here to provide perspective on the magnitude of differences between the measurements.

Table 8
Mann-Whitney U Test Results of Comparison of Person-Measures from Community of Inquiry survey subscale in Co-Teaching vs Single Instructor Courses

	Single instructor course Mean (SD)	Co-teaching course Mean (SD)	Mean Difference	Mann-Whitney U Test	*p*
Course A	N=30	N=13			
Teaching Presence	603.9 (216.6)	680.8 (265.6)	-76.9	227.5	.387
Social Presence	535.5 (131.2)	647.5 (191.7)	-112.0	268.5	.051
Cognitive Presence	597.3 (151.2)	640.6 (295.4)	-43.3	199.0	.915
Course B	N=34	N=25			
Teaching Presence	704.9 (205.3)	707.7 (243.6)	-2.8	432.50	.908
Social Presence	597.2 (164.8)	664.7 (192.6)	-67.5	516.00	.160
Cognitive Presence	652.3 (198.0)	669.6 (216.8)	-17.3	460.50	.584
Course C	N=24	N=21			
Teaching Presence	508.8 (154.5)	835.4 (189.8)	-326.7	453.5	<.001
Social Presence	548.8 (148.2)	744.4 (201.9)	-196.6	396.50	.001
Cognitive Presence	581.0 (167.7)	781.0 (213.3)	-200.0	386.0	.002

Note. Rasch person measures expressed utilizing a scale from 0 to 1000.

The analysis revealed that, although the mean ratings of all variables were higher in courses where a co-teaching model was present, this difference was only statistically significant in Course C. In Courses A and B, the largest difference between the two models was found in measures of SP. In Course C, the largest difference was found between measures of TP. One unique aspect of these results is that, although there were no significant differences in TP, SP, or CP in Courses A and B, all three of these variables were found to be significantly different when comparing the single-instructor model to the co-teaching model in Course C.

Discussion and Conclusions

The data analysis suggests that the practice of co-teaching resulted in higher ratings of teaching, cognitive, and SP when examined across all courses in the program. When examined at the level of specific courses, however, the results are inconsistent and suggest co-teaching resulted in improved student perceptions in some cases, while there were no significant differences in other cases.

Limitations

It is important to note that this study was conducted on data that were collected as a part of a program evaluation process. Among the factors that may have influenced the results, the most prominent were those related to the structure of the program as it occurred as well as the degree to which faculty collaborated in the co-teaching models. The instructional models were selected by the instructors and influenced by the administrative process for assigning instructors to courses, thus the instructional models were not experimentally controlled, nor were participants randomly assigned to groups that were used for comparison. With respect to the statistical analysis, the number of respondents that were available to participate was limited and found to be suitable only for a non-parametric statistical analysis. These factors should be taken into account when considering the degree to which these results are generalizable or applicable to other circumstances.

Program-Level Outcomes

The results of the program-level comparison of the three primary instructional models used consistently indicate that student measures of CP, TP, and SP were higher in courses in which a co-teaching model was used than in courses with a single instructor teaching either one section or two sections combined. Across all three measurements, these differences were statistically significant and suggest that, in general, co-teaching resulted in a positive impact on ratings of teaching, social, and cognitive presence as compared to either of the single-instructor models.

The differences between the two models where an instructor taught either one versus two combined sections of a course differed significantly only on the TP variable. This suggests that these two models do not differ significantly with respect to SP, which Garrison and Arbaugh (2007) described as "the ability of learners to project themselves socially and emotionally, thereby being perceived as 'real people' in mediated communication" (p. 159). Additionally, these models were not found to be significantly different on measures of CP, or "the extent to which learners are able to construct and confirm meaning through sustained reflection and discourse" (p.161). The impact of an instructor teaching two combined course sections is observed in only ratings related to TP, or the "design, facilitation, and direction of cognitive and social processes for the purpose of realizing personally meaningful and educationally worthwhile

learning outcomes" (p.163). This result should be carefully considered when assigning instructors to teach courses. For example, if two sections of a course factor into an instructor's overall teaching load as two different course assignments, it would be reasonable to assume that an instructor's workload would be distributed accordingly as well and that ratings of TP would not be adversely impacted. In this study, however, the results suggest that an assignment to teach two sections of the same course will result in diminished perceptions of TP. Although the results to do specifically point to an exact cause of this result, it is possible that the practice of combining two sections of a course leads to instructor behaviors that could be similar to a single higher enrollment course.

When considering these program-level results as a whole, it is tempting to assert that co-teaching offers clear advantages over the other two models. It is important, however, to consider some of the nuances of these comparisons that reveal an alternate interpretation. If we consider the student-teacher ratio, then both the co-teaching model and single instructor teaching one course section are similar in terms of the number of students per instructor (approximately 25 students per instructor). The model in which a single instructor is teaching two sections combined, however, has a student-teacher ratio of approximately 50:1. Although the co-teaching model courses had higher ratings across all three variables, in terms of improving the student experience in a course, simply maintaining a lower student-teacher ratio will likely result in improvement in the TP component of the Community of Inquiry model.

Course-Level Outcomes

When we consider the results of course-level comparisons, we begin to see how individual differences in a course, such as teaching style, subject matter, teaching experience, and willingness to collaborate may mediate the differences that were noted at the program level. If co-teaching is truly a better teaching model than that of a single instructor, then we would expect to see significant differences across the board when comparing these two models within the same course. This type of across-the-board difference was only found within one course (Course C) where an alternate co-teaching approach was used. Although there was a common course syllabus, online material, and assignments used in both terms, the team of co-teaching instructors did not include the previous independent instructor. As such, any interpretation of these results as clear evidence in favor of co-teaching should be tempered by the fact that other factors could have also impacted this difference and that additional research is necessary to establish a clearer indication of the factors leading to this difference.

An additional aspect of the course-level comparisons is that Courses A and B used a parallel teaching approach in which the instructors collaborated on planning and design, but each instructor interacted with students only within their assigned section of the course with respect to grading and feedback. For both Course A and Course B, a single instructor taught two combined sections of the course in the first term, then that same instructor was part of a co-teaching team in the second term. The online course design, material, and assignments remained largely consistent over time, thus the primary difference between the first term (single instructor) and the second term (co-teaching) for both Course A and Course B was the student-teacher ratio (approximately 50:1 vs. approximately 25:1). Given the results of the program-level analysis, we would expect that the different student-teacher ratio would correspond to differences in the student measurements of CP, TP, or SP. Although mean differences were observed in favor of the co-teaching model, these differences were not found to be statistically significant. Given the fact that these two courses used a parallel co-teaching model, as opposed to a more collaborative approach like alternate teaching, it is possible that the parallel approach has less of an impact on

the student measures of TP, SP, and CP than other approaches. Prior research by Fuller and Bail (2011) emphasized that one goal of co-teaching in online learning is to increase sharing of expertise and gaining the benefit of continued instructor learning regarding pedagogical and subject matter knowledge. It is possible that the benefit of the parallel co-teaching model, when considered within the context of an online learning course, is limited to collaborative planning and that other benefits of co-teaching that may have a more direct influence on student perceptions of the learning experience are less prominent in a parallel teaching model.

Implications for Practice and Future Research

When viewed within the context of a program evaluation, the results of this study suggest that co-teaching is a promising practice, in some cases, and that practices and policies regarding assigning instructors to online courses should consider both the ability of an instructor to teach collaboratively and the potential impact of higher student-teacher ratio. Within this study, lower ratings of CP, SP, and TP were found in courses where a single instructor was teaching multiple sections of the same course in one academic term. However, further research is necessary in order to determine what aspects of this teaching model are critical in this relationship as well as what facets of an online learning environment may reverse this pattern. Although assigning a single instructor to two sections of a single course may fit well in terms of teaching loads, these results suggest that there is a cost to this practice. In practice, however, administrators and instructors should carefully consider the issue of student-teacher ratio and examine a broader view of instructors' overall workload when making teaching assignments so that an instructor can successfully attend to the demands of an online course.

With respect to co-teaching models, the results of this study suggest that co-teaching is effective under certain conditions and alternate teaching was more promising that parallel teaching. Administrators and stakeholders are encouraged to explore these teaching models more deeply and create policies and practices to enhance the student experience in online courses by supporting effective collaborative teaching models. The next steps in future research into co-teaching in online courses would benefit from a focus on the nature of the collaboration among instructors in order to identify the key characteristics of successful collaborative teaching models.

Conclusion

A program evaluation of this nature involves many variables, some of which cannot be controlled when analyzing student perceptions of various teaching models. The overall results from this examination of diverse teaching formats within online courses suggest however that co-teaching and course sizes have an impact on the TP, SP, and CP. Previous research on co-teaching in K-12 and face-to-face college settings support these findings and suggest that students experience a greater level of satisfaction in the learning environment when there are multiple instructors. It should not be a surprise that students receiving instruction using a co-teaching model in an online environment also find the course more meaningful and well designed.

The role of teacher educators is to prepare their students for a lifelong career in education. As we seek to implement a rigorous curriculum that supports our students' development of their own teaching philosophy, it is necessary that we are providing this instruction in an environment that promotes multiple perspectives of thought on hot topics within our field. Morelock et al. (2017) found that students in courses where co-teaching was present appreciated not only the range of expertise brought with multiple instructors but also the diverse views and approaches. The benefits observed through previous research on co-teaching and the

findings within this current study warrant more rigorous research on the effective implementation and comparisons of various teaching models used within online teacher education.

References

Akyol, Z., Vaughan, N., & Garrison, D. R. (2011). The impact of course duration on the development of a community of inquiry. *Interactive Learning Environments*, *19*(3), 231–246.

Arbaugh, J. B., Cleveland-Innes, M., Diaz, S. R., Garrison, D. R., Ice, P., Richardson, J. C., & Swan, K. (2008). 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.

Bangert, A. W. (2009). Building a validity argument for the community of inquiry survey instrument. *The Internet and Higher Education*, *12*(2), 104–111.

Bond, T. G., & Fox, C. M. (2013). *Applying the Rasch model: Fundamental measurement in the human sciences*. New York, NY: Routledge

Boone, W. J., Staver, J. R., & Yale, M. S. (2014). Rasch analysis in the human sciences. New York, NY: Springer.

Boone, W. J., Townsend, J. S., & Staver, J. (2011). Using Rasch theory to guide the practice of survey development and survey data analysis in science education and to inform science reform efforts: An exemplar utilizing STEBI self-efficacy data. *Science Education 95*, 258–280. Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/sce.20413/full

Boston, W., Ice, P., Diaz, S., Richardson, J., Gibson, A., & Swan, K. (2009). An exploration of the relationship between indicators of the community of inquiry framework and retention in online programs. *Journal of Asynchronous Learning Networks*, *13*(3), 67–83.

Cook, L., & Friend, M. (1995). Co-Teaching: Guidelines for creating effective practices. *Focus on Exceptional Children*, *28*(3), 1–16.

Díaz, S. R., Swan, K., Ice, P., & Kupczynski, L. (2010). 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), 22–30.

Friend, M., Cook, L., Hurley-Chamberlain, D., & Shamberger, C. (2010). Co-Teaching: An illustration of the complexity of collaboration in special education. *Journal of Educational and Psychological Consultation: The Official Journal of the Association for Educational and Psychological Consultants*, *20*(1), 9–27.

Fuller, R. G., & Bail, J. (2011). Team teaching in the online graduate environment: Collaborative instruction. *International Journal of Information & Communication Technology Education*, *7*(4), 72–83.

Garrison, D. R., Anderson, T., & Archer, W. (2000). Critical inquiry in a text-based environment: Computer conferencing in higher education. *The Internet and Higher Education*, *2*(2), 87–105.

Garrison, D. R., & Arbaugh, J. B. (2007). Researching the community of inquiry framework: Review, issues, and future directions. *The Internet and Higher Education*, *10*(3), 157–172.

Garrison, D. R., Cleveland-Innes, M., & Fung, T. S. (2010). Exploring causal relationships among teaching, cognitive and social presence: Student perceptions of the community of inquiry framework. *The Internet and Higher Education*, *13*(1), 31–36.

Jones, F., & Harris, S. (2012). Benefits and drawbacks of using multiple instructors to teach single courses. *College Teaching*, *60*(4), 132–139.

Krometis, L.-A. H., Clark, E. P., Gonzalez, V., & Leslie, M. E. (2011). The "death" of disciplines: Development of a team-taught course to provide an interdisciplinary perspective for first-year students. *College Teaching*, *59*(2), 73–78.

Kumar, S., Dawson, K., Black, E. W., Cavanaugh, C., & Sessums, C. D. (2011). Applying the community of inquiry framework to an online professional practice doctoral program. *The International Review of Research in Open and Distributed Learning*, *12*(6), 126–142.

Kumar, S., & Ritzhaupt, A. D. (2014). Adapting the community of inquiry survey for an online graduate program: Implications for online programs. *E-Learning and Digital Media*, *11*(1), 59–71.

Legon, R., & Garrett, R. (2017). *The changing landscape of online education*. Quality Matters &

Eduventures. Retrieved from https://www.qualitymatters.org/sites/default/files/research-docs-pdfs/CHLOE-First-Survey-Report.pdf

Linacre, J. M. (2018). WinSteps (Version 4.1.0) [Windows]. Retrieved from http://www.winsteps.com/winsteps.htm

Magda, A., & Buban, J. (2018). *The state of innovation in higher education: A survey of academic administrators*. Louisville, KY: Online Learning Consortium and The Learning House, Inc. Retrieved from https://olc-wordpress-assets.s3.amazonaws.com/uploads/2018/04/The-State-of-Innovation-in-Higher-Education-A-Survey-of-Academic-Administrators.pdf

Morelock, J. R., Lester, M. M., Klopfer, M. D., Jardon, A. M., Mullins, R. D., Nicholas, E. L., & Alfaydi, A. S. (2017). Power, perceptions, and relationships: A model of co-teaching in higher education. *College Teaching*, *65*(4), 182–191.

Nagel, L., & Kotzé, T. G. (2010). Supersizing e-learning: What a CoI survey reveals about teaching presence in a large online class. *The Internet and Higher Education*, *13*(1), 45–51.

OECD. (2012). Scaling PISA Cognitive Data. In *Programme for International Student Assessment 2009 Technical Report*. OECD Publishing, Paris.

Online Learning Consortium. (2015, July 7). E-Learning Definitions. Retrieved March 24, 2018, from https://onlinelearningconsortium.org/updated-e-learning-definitions-2/

Richardson, J., Arbaugh, J. B., Cleveland-Innes, M., Ice, P., Swan, K., & Garrison, D. R. (2012). Using the community of inquiry framework to inform effective instructional design. In *The Next Generation of Distance Education* (pp. 97–125). New York, NY: Springer.

Rubin, B., & Fernandes, R. (2013). Measuring the community of online classes. *Journal of Asynchronous Learning Networks*, *17*(3), 1–21.

Russell, V., & Curtis, W. (2013). Comparing a large- and small-scale online language course: An examination of teacher and learner perceptions. *Internet and Higher Education*, *16*, 1–13.

Seaman, J. E., Allen, I. E., & Seaman, J. (2018). *Grade increase: Tracking distance education in the United States*. Babson Survey Research Group. Retrieved from https://onlinelearningsurvey.com/reports/gradeincrease.pdf

Shea, P., & Bidjerano, T. (2009). Community of inquiry as a theoretical framework to foster "epistemic engagement" and "cognitive presence" in online education. *Computers & Education*, *52*(3), 543–553.

Solis, M., Vaughn, S., Swanson, E., & Mcculley, L. (2012). Collaborative models of instruction: The empirical foundations of inclusion and co-teaching. *Psychology in the Schools*, *49*(5), 498–510.

Strohschen, G., & Heaney, T. (2000). This isn't Kansas anymore, toto: Team teaching online. *New Directions for Adult & Continuing Education*, (87), 33.

Swan, K., Richardson, J., Ice, P., Garrison, D. R., Cleveland-Innes, M., & Arbaugh, J. B. (2008). Validating a measurement tool of presence in online communities of inquiry. *E-Mentor*, *2*(24). Retrieved from www.e-mentor.edu.pl/eng

Vaughan, N. D. (2010). A blended community of inquiry approach: Linking student engagement and course redesign. *The Internet and Higher Education*, *13*(1), 60–65.

Walters, K., & Misra, J. (2013). Bringing collaborative teaching into doctoral programs: Faculty and graduate student co-teaching as experiential training. *The American Sociologist*, *44*(3), 292–301.

Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis*. Chicago, IL: MESA Press.

Wright, B. D., & Stone, M. H. (1979). *Best test design*. Chicago, IL: MESA Press. Retrieved from https://research.acer.edu.au/measurement/1/

Yamamoto, K., & Kulick, E. (1999). Scaling methodology and procedures for the TIMSS mathematics and science scales. In M. O. Martin, K. D. Gregory, & S. E. Stemler (Eds.), *TIMSS 1999 Technical Report* (pp. 237–263).