The purpose of the Online Learning Consortium (OLC) is to help learning organizations continually improve the quality, scale, and breadth of online programs according to their own distinctive missions, so that education will become a part of everyday life, accessible and affordable for anyone, anywhere, at any time, in a wide variety of disciplines.

Please direct all inquiries to publications@onlinelearning-c.org.

To learn more about submission guidelines for authors and subscription information, please visit: https://olj.onlinelearningconsortium.org/index.php/olj/about.

Copyeditors: Crossfield Professional Writers, Dr. Leslie Pourreau, and Dr. Mary Rice

Printed in the United States of America 0 9 8 7 6 5 4 3 2 1

ISSN: 2472-5730 (online)

The Online Learning Consortium, Inc. is a consortium of higher-education providers sharing the common bonds of understanding, supporting and delivering education via asynchronous learning networks (ALNs). With the mission of providing learning to anyone anywhere, OLC seeks to provide new levels of learning capability to people seeking higher and continuing education. For more information about OLC please visit: http://onlinelearningconsortium.org.

Online Learning

Volume 21 Issue 1 – March 2017

Table of Contents

Introduction to Issue 21:1

Peter J. Shea
University at Albany, State University of New York

This issue of OLJ explores several important themes related to online teaching and learning including faculty, student, and institutional concerns. The issue begins with four articles related to faculty, starting with a review of the literature by Nancy Pope Wingo, Nataliya V. Ivankova, Jacqueline A. Moss of The University of Alabama at Birmingham titled *Faculty Perceptions about Teaching Online: Exploring the Literature Using the Technology Acceptance Model as an Organizing Framework.* This study investigates an area of particular importance to faculty and institutional leaders – what are the factors that facilitate adoption of online teaching? Results to date in this area tend to be bleak: repeated studies indicate low level of acceptance of online learning by faculty and troubling faculty attitudes regarding learner outcomes in online coursework. This study provides a much needed, theoretically-framed approach to the issue. This paper is crucial reading for faculty developers, institutional leaders, and especially scholars seeking an updated and comprehensive resource for considering faculty attitudes toward online education.

The next article looks at a more specific dimension of faculty adoption of online education, willingness to attend training in the use of the institution's course management system. In *Course Management System's Compatibility with Teaching Style Influences Willingness to Complete Training,* Audrey Pereira, of Fitchburg State University, and Monika Wahi of Laboure College ground their study in Diffusion of Innovation theory to understand conditions under which faculty are more likely to participate in training necessary to use the CMS effectively. Using survey methods, the authors conclude that compatibility, defined as the degree to which instructors perceive the CMS as being consistent with their existing values, past experiences, and current or future teaching needs, was the only factors statistically significantly associated with willingness to complete online and in-person training. The study provides a useful set of recommendations for faculty development practices, theory and implications for future research.

Following the theme of faculty professional development and learning, Pamela Beach of Queen's University and Dale Willows of the Ontario Institute for Studies in Education, at the University of Toronto provide a related study titled *Understanding Teachers' Cognitive Processes during Online Professional Learning: A Methodological Comparison*. This paper used three think-aloud protocols as a lens to understand faculty learning that may lead to adoption of online pedagogy, in this case among educators in pre-college settings. How do we know what people learn when they are involved in professional development activities? The most common approaches include survey questions after the experience. However, we may benefit more from understanding cognition through verbalization and think-aloud protocols can offer such a view. In this study the authors examined different approaches to studying verbalization and the underlying cognitive processes exhibited when educators are learning about online instruction. The paper is valuable for scholars seeking insight into the benefits and limitations to employing each type of think-aloud method in the context of online professional development.

Online educator learning is again the focus in *Moving Beyond Smile Sheets: A Case Study on the Evaluation and Iterative Improvement of an Online Faculty Development Program* by Ken-Zen Chen of National Chiao-Tung University and Patrick R. Lowenthal, Christine Bauer, Allan Heaps, and Crystal Nielsen of Boise State University. In this study the authors are also interested in what faculty derive from professional development and demonstrate their seriousness by employing a mixed methods approach to data collection. The study investigates not only faculty perceptions but also their participation, skills, dispositions, and concerns related to involvement in a sustained faculty development program. The paper is a vital resource for others seeking to understand the impact of professional training for online educators.

The next paper shifts away from faculty issues and takes on issues of access. If we are to provide equitable access to online education, instructional content needs to be available to all. For the deaf and hard of hearing (as well as other audiences) the use of a free service that automatically captions online video content would seem to be a nearly miraculous solution. In *Reading Between the Lines: Accessing Information via YouTube's Automatic Captioning System* Chad Smith and Tamby Allman of Texas Woman's University and Samantha Crocker of Weatherford Regional Day School Program for the Deaf analyze such a service. Automated captioning is far from perfect. In this paper, the authors identify 11 categories of different errors in the captioning of videos targeted to middle-school audiences and then assigned college students to interact with videos containing different error types. The authors conclude that, when automatic-captions contain significant numbers of errors, and when no audio content is available, even hearing, college-educated adult readers are unable to comprehend the messages being delivered. Clearly more work needs to be done before we use free captioning options to serve all students.

The next section, on online academic integrity, includes *Examining the Effect of Proctoring on Online Test Scores* by Helaine M. Alessio, Nancy Malay, Karsten Maurer, A. John Bailer, and Beth Rubin of Miami University. Numerous reports indicate that the majority of undergraduate students admit to some form of academic dishonesty in both classroom and online settings. In this paper the authors go beyond student self-reports to look at online students taking tests with and without a proctoring solution. They find both temporal differences and that students in proctored conditions scored lower on tests than students in un-proctored conditions. The finding strongly suggests that cheating occurs in the absence of monitoring. This paper is important in both its design and rigor and results indicate that either we need to develop instruction and assessment that avoids high-stakes tests or invest in proctoring to ensure students don't engage in academic dishonesty when confronted with high-stake tests.

In the next section of this issue a series of articles address issues around students, community, and online learning.

Large format online courses create challenges in enacting pedagogies, such as dialogic forms of teaching and learning that might deter academic dishonesty found in the previous study. However, enacting productive approaches that focus on quality interaction with hundreds of students is difficulty to achieve. In *Creating a Community of Inquiry in Large- Enrollment Online Courses: An Exploratory Study on the Effect of Protocols within Online Discussions* Baiyun Chen and Aimee deNoyelles of the University of Central Florida with Kerry Patton and Janet Zydney of the University of Cincinnati explore how to use online discussion protocols to promote substantial learning in higher enrollment online courses. By using and iteratively redesigning these structured approaches to guiding online discussion the authors document improvements in both student perceptions of forms of presence and quality improvements in the nature of student discussion posts.

A second article in this section, *Exploring Small Group Analysis of Instructional Design Cases in Online Learning Environments* by Jesus Trespalacios of Boise State University, also examines the use of student interaction with significant guidance to enhancing learning. In this study the author analyzes case-study teaching approaches and seek to determine the effectiveness of small group analysis of cases in instructional design when compared with experts' analysis and to understand students use of VoiceThread for engaging for analyzing these cases. Results indicate that creating a small group discussion and requiring students to develop a VoiceThread presentation following scaffolding guidelines to analyze ID case studies assisted learners to identify relevant issues about the cases.

The next paper in this section is *Utilization of an Educational Web-Based Mobile App for Acquisition and Transfer of Critical Anatomical Knowledge by* Kevin Hannon of the Department of Basic Medical Sciences in College of Veterinary Medicine at Purdue University. In this article the author addresses students' need for greater preparation in lab sections in anatomy courses. The paper analyzes two uses of a web application to prepare students in contrast to more traditional modes of content delivery. The paper concludes that in contrast to a traditional reading tasks, use of the app significantly enhanced initial learning of anatomy and the transfer of content learned to a related, but new area. The author proposes that students using the app were better prepared for lecture and lab than students reading a textbook. The app may increase opportunity for time on task or engage cognitive processes central to the assessment task more effectively than reading static text, but more research into underlying processes and theoretical framing are needed.

The next two studies examine or use Wenger's Community of Practice Framework. The first of these is *A Critical Review of the Use of Wenger's Community of Practice (CoP) Theoretical Framework in Online and Blended Learning Research, 2000-2014* by Sedef Uzuner Smith of the University of Houston Downtown and Suzanne Hayes and Peter Shea of the State University of New York at Albany. In this paper my colleagues and I provide an integrative research review to address three questions on this most influential of theories. We examine which studies make central use of the CoP framework, which of these establish strong linkages between the framework and their findings, and among this latter group, identify studies that provide productive opportunities for future CoP research in online and blended teaching and learning. We conclude that online/blended learning research employing the CoP theory should enter a new phase of development. There is a need for studies that not only employ different aspects of Wenger's CoP theory but also extend the traditional practice of theory verification to provide more complex and nuanced understandings of online/blended learning environments.

A second paper in this issue, *Institutional Factors for Supporting Electronic Learning Communities* by Jayme N. Linton of Lenoir-Rhyne University attempts to provide such nuanced understanding through Wenger's CoP framework. This study explored how the electronic learning community process at a state virtual high school supported online high school teachers through dimensions of communities of practice (CoP) framework. To answer the study's research question, the author collected data related to five strategies identified by Wenger as effective methods for an organization to support and enhance the effectiveness of the work of CoPs. These include valuing the work, creating time and space, encouraging participation, removing barriers, and connecting to the organizational strategy. Results indicate that these strategies supported and increased the effectiveness of the electronic learning community (eLC) that was studied but also showed that the institutionally-driven nature of the eLC process could block alignment with the CoP framework.

This issue of OLJ closes with an article by Amy Roberts of University of Nebraska and Jennifer LoCasale-Crouch, Bridget K. Hamre and Jordan M. Buckrop of the University of Virginia titled *Adapting for Scalability: Automating the Video Assessment of Instructional Learning.* The questions addressed in this study relate to enhancing efficiency of assessment processes in large scale online educational formats. The authors examine whether the assessment of teaching skills collected through videotaped observations (Video Assessment of Instructional Learning or VAIL) could be automated rather than manually scored. Results indicate relatively high correlations between the manually scored and auto-scored assessments and that the strength of the associations between automated and hand-scored systems varied by what was assessed. The paper holds promise in assisting efforts to deliver online education at greater scale, but additional research on other contexts is needed.

We invite you to read, share, and cite the high-quality papers in this issue and help us to continue to enhance the research and practice of online learning.

References

Alessio, H. M., Malay, N., Maurer, K., Bailer, A. J., & Rubin, B. (2017). Examining the effect of proctoring on online test scores, *Online Learning, 21* (1), 146-161. doi: 10.24059/olj.v21i1.885

Beach, P, & Willows, D. (2017). Understanding teachers' cognitive processes during online professional learning: A methodological comparison, *Online Learning 21*(1), 60-84. doi: 10.24059/olj.v21i1.949

Chen, B., deNoyelles, A., Patton, K., & Zydney, J. (2017). Creating a community of inquiry in large-enrollment online courses: An exploratory study on the effect of protocols within online discussions, *Online Learning 21*(1), 165-188. doi: 10.24059/ olj.v21i1.816

Chen, K., Lowenthal, P.R., Bauer, C., Heaps, A., & Nielsen, C. (2017). Moving beyond smile sheets: A case study on the evaluation and iterative improvement of an online faculty development program, *Online Learning 21*(1), 85-111. doi: 10.24059/olj.v21i1.810

Hannon, K. (2017). Utilization of an educational web-based mobile app for acquisition and transfer of critical anatomical knowledge, thereby increasing classroom and laboratory preparedness in veterinary students, *Online Learning 21*(1), 201-208. doi: 10.24059/olj.v21i1.882

Lee-Post, A. & Hapke, H (2017). Online learning integrity approaches: Current practices and future solutions, *Online Learning 21*(1),135-145. doi: 10.24059/olj.v21i1.843

Linton, J. N. (2017). Institutional factors for supporting electronic learning communities, *Online Learning 21*(1), 238-256. doi: 10.24059/olj.v21i1.953

Pereira, A. S. & Wahi, M.M. (2017). Course management system's compatibility with teaching style influences willingness to complete training. *Online Learning 21*(1), 36-59. doi: 10.24059/olj.v21i1.763

Roberts, A. M., LoCasale-Crouch, J., Hamre, B. K., & Buckrop, J. M. (2017). Adapting for scalability: Automating the video assessment of instructional learning, *Online Learning 21*(1), 257-272. doi: 10.24059/olj.v21i1.961

Smith, C., Crocker, S. & Allman, T. (2017) Reading between the lines: Accessing information via *YouTube*'s automatic captioning, *Online Learning 21*(1), 115-131. doi: 10.24059/olj.v21i1.823

Smith, S. U., Hayes, S., & Shea, P (2017). A critical review of the use of Wenger's Community of Practice (CoP) theoretical framework in online and blended learning research, 2000-2014, *Online Learning 21*(1), 209-237. doi: 10.24059/olj.v21i1.963

Trespalacios, J. (2017). Exploring small group analysis of instructional design cases in online learning environments, *Online Learning 21*(1) 189-200. doi: 10.24059/ olj.v21i1.928

Wingo,N. P., Ivankova, N. V., & Moss, J. A. (2017) Faculty perceptions about teaching online: exploring the literature using the technology acceptance model as an organizing framework, *Online Learning 21*(1), 15-35. doi: 10.10.24059/olj.v21i1.761

SECTION I:

Faculty, Professional Development, and Online Teaching

Faculty Perceptions about Teaching Online: Exploring the Literature Using the Technology Acceptance Model as an Organizing Framework

Nancy Pope Wingo, Nataliya V. Ivankova, Jacqueline A. Moss
The University of Alabama at Birmingham

Abstract

Academic leaders can better implement institutional strategic plans to promote online programs if they understand faculty perceptions about teaching online. An extended version of a model for technology acceptance, or TAM2 (Venkatesh & Davis, 2000), provided a framework for surveying and organizing the research literature about factors that have influenced faculty's adoption of online delivery methods for courses and their willingness to continue to teach online. This paper presents the results of a synthesis of 67 empirical studies about faculty teaching online published between 1995 and 2015, using TAM2 constructs as an organizing framework. This validated model provided a lens for understanding research about faculty perceptions of the user-friendliness and ease of use of technology for online course delivery, as well as the overall experience of teaching online. Studies in this review revealed concerns among faculty regarding their perceived barriers to student success in online classes, uncertainty about their image as online instructors, technical support needs, and their desire for reasonable workload and manageable class enrollments in online classes.

Keywords: online faculty, online teaching, Technology Acceptance Model, literature review, faculty adoption

Wingo,N. P., Ivankova, N. V., & Moss, J. A. (2017) Faculty perceptions about teaching online: exploring the literature using the technology acceptance model as an organizing framework, *Online Learning 21*(1), 15-35. doi: 10.10.24059/olj.v21i1.761

Introduction

Higher education faculty in the United States are increasingly being asked to teach online (Allen & Seaman, 2015). Yet faculty may be reluctant to embrace different forms of online teaching, due to fear of change, concerns about the reliability of technology, skepticism about student outcomes in online learning environments, workload issues, and other factors (Bacow et al. 2012; Betts & Heaston, 2014; Bolliger & Wasilik, 2009; McQuiggan, 2012). Fostering faculty's acceptance of online delivery methods is critical for institutions that consider online learning to be a key part of their strategic plan; to accomplish this, administrators need to understand how faculty perceive teaching online and what factors shape those perceptions.

The Technology Acceptance Model (TAM) (Davis, 1989) has been used for decades to explain how users accept new technologies. An extended version of the original model, "TAM2" (Venkatesh & Davis, 2000), confirmed the effects of various factors on key constructs of the original model. Because it illustrates influences on technology acceptance in a clearly structured format, this validated model can provide a framework to better understand faculty's perceptions about teaching online. Despite its utility for understanding acceptance of new technology by the users, the TAM2 has not been applied to systematically study faculty acceptance of technology to deliver online courses. This paper addresses this gap and presents the results of the synthesis of the research literature regarding faculty perceptions about teaching online, using constructs in the TAM2 (Venkatesh & Davis, 2000) as an organizing framework.

Theoretical Framework

We used a model of technology acceptance by users in organizations, validated by Venkatesh and Davis (2000) and based on the Technology Acceptance Model (TAM) (Davis, 1989) to guide the review of research literature that explored factors influencing faculty's perceptions of online teaching. We chose this model because it includes factors regarding users' technical experiences and their perceptions about how using technology might affect their status in an organization, providing a broad scope for surveying research about faculty's experiences as online instructors. A meta-analysis by King and He (2006) of 88 studies in different fields determined that the TAM was a "powerful and robust predictive model" (p. 751) to understand technology acceptance of users in various contexts. The original TAM is an empirically validated framework initially developed by Davis (1989) to explain end users' willingness to use new technologies in organizations. Its two key constructs are *perceived usefulness* (PU), or the degree to which a person believes a technology will improve his or her job performance, and *perceived ease of use* (PEU), or the amount of effort a person believes he or she will need to expend to master that technology.

The TAM was developed further when researchers sought to understand determinants of PEU (Venkatesh & Davis, 1996) and PU (Venkatesh & Davis, 2000). Venkatesh and Davis (1996) found that users' computer self-efficacy significantly affected PEU both before and after exposure to a technological system. They later explored the determinants of PU over four longitudinal studies at various sites and found that PU was significantly affected by "social influence processes" (subjective norm, voluntariness, and image) and "cognitive instrumental processes" (job relevance, output quality, result demonstrability, and perceived ease of use) (Venkatesh & Davis, 2000, p. 187). The resulting model, or "TAM2," (Venkatesh & Davis, 2000), showed a more detailed relationship among various factors that influenced technology acceptance. The updated model continued to be used by researchers in different fields; in fact, Marangunić and Granić (2015) concluded, after a review of 85 publications using the TAM model, that the "TAM has evolved to become the key model in understanding the predictors of human behavior toward potential acceptance or rejection of the technology" (p. 92).

Figure 1 illustrates the TAM2 model with its two key constructs (PU and PEU) and shows the various factors found to influence PU (Venkatesh & Davis, 2000). These factors included the *subjective norm*, or users' perceptions of whether others in an organization believed they should use a technological system. The subjective norm was moderated by whether users had prior *experience* using the technology, and whether using it was mandatory or *voluntary* in an organization.The subjective norm also influenced a user's perceptions of how his or her *image* might be affected as a result of using technology. Other factors influencing PU included *job relevance*, or user's perceptions of how a technological system could help them accomplish significant goals; *output quality*, or the quality of technology needed to accomplish specific tasks; and *result demonstrability*, or the perceived tangible results and benefits of using a technological system. Each of these factors will be discussed in more detail further in this article.

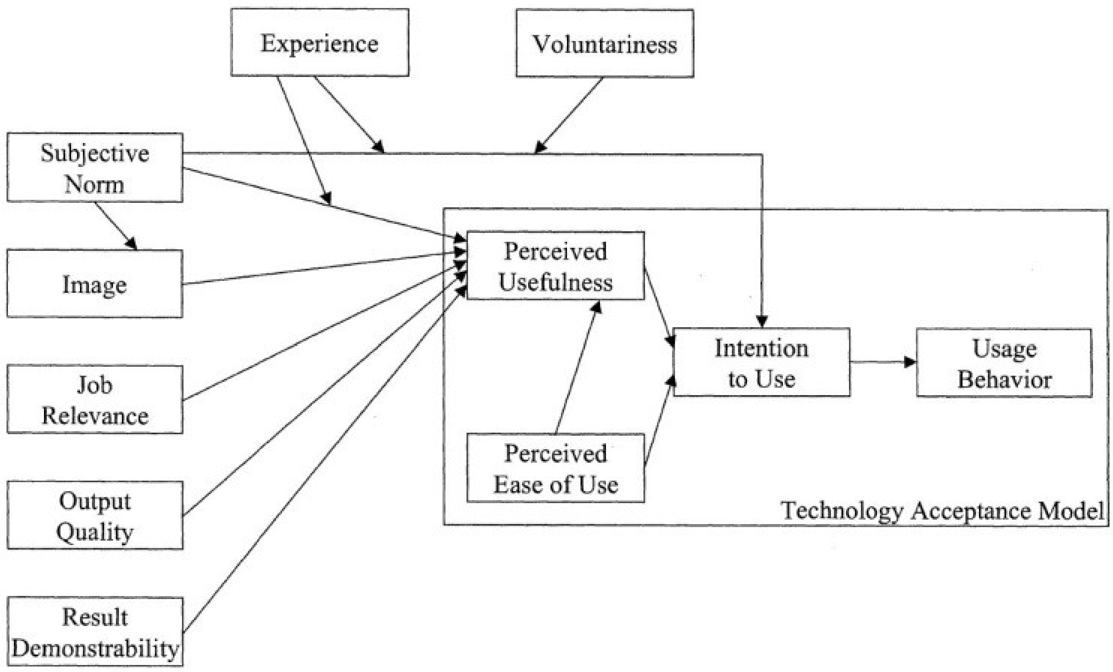


Figure 1. TAM2. Reprinted by permission, (Viswanath Venkatesh, Fred D. Davis), A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies, *Management Science, 46, 2.* Copyright (2000). The Institute for Operations Research and the Management Sciences, 5521 Research Park Drive, Suite 200, Catonsville, Maryland 21228 USA.

Methods

To explore empirical literature regarding faculty teaching online in the context of technology acceptance models, a search for studies that used the TAM or TAM2 in various disciplines was first conducted to gain a broader understanding of the applications of these models. Then, the search was narrowed to studies that used either of the models as theoretical frameworks for the experience of teaching online. After finding few studies meeting these criteria, the search was adjusted further to explore research that addressed specific elements of technology acceptance and adoption outlined in the TAM2 model in terms of teaching online, even if the authors had not used constructs of the model as a framework. Reviewing the literature in this manner allowed for a fuller picture of numerous factors related to faculty experiences in teaching online that have been explored by researchers, although they may not have explicitly stated the use of the TAM or TAM2 models. The procedures and results of those searches are discussed in more detail in the following sections.

Searching for Research Applying the TAM to Various Disciplines

Many researchers have applied some version of the TAM to various disciplines over the years. To understand the extent of its application, we searched empirical literature, using the terms "Technology Acceptance Model" or "TAM" in three databases: Academic Search Premier, ERIC, and Education Full Text (H. W. Wilson). The search yielded over 14,000 results. The same search terms in Google Scholar returned over 44,000 possible articles. Obviously, this model is popular for providing a framework for technology research, yet it also has a broad range of possible applications. To determine the application of the TAM by discipline, we narrowed our search to major journals in various fields that were likely to be concerned with technology acceptance, again using the search terms "Technology Acceptance Model" or "TAM." Results showed significantly more research rooted in the TAM in business than in other disciplines: Business Search Premier = 124 (59% of total); Education Full Text = 34 (16% of total); Library/Information Science/Technology Abstracts = 32 (15% of total); and CINAHL Plus with Full Text = 20 (9.5% of total).

These numbers were not surprising, since the models' developers were business professors. Yet in the twenty-first century, with technology use permeating so many fields, the TAM might be increasingly used to better understand technology acceptance in any discipline. In fact, a statistical meta-analysis of TAM constructs performed by King and He (2006) across 88 empirical articles in the social sciences led the authors to conclude that the TAM was "a valid and robust model that has been widely used, but which potentially has wider applicability" (p. 740).

Searching for Research Applying the TAM to Online Teaching

To understand more about the application of the TAM in higher education, specifically concerning technology acceptance among faculty teaching online, we combined the search terms "Technology Acceptance Model" and "TAM" in various combinations with "online," "distance education," "faculty," and "instructors." Using these search terms in Academic Search Premier, CINAHL Plus with Full Text, ERIC, and Education Full Text (H. W. Wilson), we found only three articles from peer-reviewed journals that specifically applied the TAM to higher education faculty who were teaching in an online environment (Alsofyani, Aris, Eynon, & Majid, 2012; Gibson, Harris, & Colaric, 2008; Huang, Deggs, Jabor, & Machtmes, 2011). A search on Google Scholar using "Technology Acceptance Model," "faculty," and "online" yielded 598 results, but a scan of these abstracts again revealed that most of the studies were using data about students or were focused on marketing or online employee training programs. The Google Scholar search did reveal two additional empirical articles that used the TAM to better understand faculty's intentions to accept online education (Stewart, Bachman, & Johnson, 2010; Wang & Wang, 2009).

To complete the search, we first examined 102 different articles about faculty who teach online that we had collected from various search engines since 2007, searching for themes and keywords that related to constructs in the TAM2, even though they had not specifically mentioned the TAM2. We then conducted searches for other articles specifically related to each construct, pairing terms such as “computer self-efficacy” and “job relevance” with “online,” “distance education,” “faculty,” and “instructors.”

Data Analysis

Our combined searches yielded 67 empirical studies about faculty teaching online published between 1995 and 2015. To analyze these articles, we first made a list of construct components and their descriptions from the TAM2 model. We then carefully read each article, making notes about the reported findings that reflected various TAM2 construct components. This process revealed that these articles addressed issues that were described by at least one construct in the TAM2 model, even though they did not use the TAM2 model explicitly. We then organized all studies in a table grouped by the TAM2 construct components and developed short summaries of the major results that reflected these components. These 67 articles revealed a fuller picture of faculty’s inclinations about technology adoption for online learning. Table 1 (below) shows all articles (N=67) included in this literature review with summarized findings that aligned with TAM2 constructs.

In our analysis of these articles, we did not attempt to make significant distinctions between “faculty acceptance” and “faculty satisfaction.” In making this decision, we drew from a conceptualization of faculty satisfaction by Hagedorn (2000), illustrating disengagement at one end of a continuum, acceptance or tolerance in the middle, and job appreciation or engagement at the other end. This continuum suggested that faculty satisfaction would occur only after faculty had accepted some aspect of teaching online.

Results

We organized this review according to each construct in the model, beginning with studies that addressed faculty’s PEU of technology for online course delivery and then exploring the various determinants of PU.

Table 1

Articles Reviewed in this Study, Aligned with TAM2 Constructs and Major Findings

TAM2 Construct	Studies Referenced	Major Findings
Perceived ease of use (PEU)	Bolliger & Wasilik (2009) Christianson, Tiene, & Luft (2002)	Faculty were less satisfied with teaching online when they had technical problems.
	Compeau & Higgins (1995) Conceição (2006) DeGagne & Walters (2010) Green, Alejandro, & Brown (2009)	Faculty who were more confident about their technical skills were more willing to teach online.
	Osika, Johnson, & Buteau (2009)	Faculty who were more skilled with technology were more satisfied with teaching online.
	Panda & Mishra (2007)	Faculty valued continuing education, even when

	Ryan, Hodson-Carlton, & Ali (2005) Shea (2007) Shea, Pickett, & Li (2005) Tabata & Johnsrud (2008) Zhen, Garthwait, & Pratt (2008)	they were skilled online instructors.
Subjective norm	Allen & Seaman (2012b) Allen & Seaman (2015) Betts & Heaston (2014) Chapman (2011) Dooley & Murphrey (2000) Huang, Deggs, Jabor, & Machtmes (2011) Lee (2001) Maguire (2009) Orr, Williams, & Pennington (2009) Wang & Wang (2009) Wickersham & McElhany (2010)	Faculty did not view online education with the same optimism that administrators did. Faculty desired clearer statements of institutional goals and policies regarding online education. Faculty teaching online needed strong institutional support in various forms.
Voluntariness	Betts (2009) Dooley & Murphrey (2000) Hixon, Barczyk, Buckenmeyer, & Feldman (2011) Jacobsen (2000) Lackey (2011) McQuiggan (2012) Shea (2007)	Motivating faculty to teach online required different strategies, depending on whether faculty were required to teach online or chose to do so. Training faculty to teach online could promote faculty satisfaction, despite whether teaching online was mandatory or voluntary.
Experience	Allen & Seaman (2012a) Orr et al. (2009) Ryan et al. (2005) Seaman (2009) Shea et al. (2005) Ulmer, Watson, & Derby (2007)	Faculty who had taught online were more positive about the effectiveness of online teaching. Faculty who had taught online were more willing to continue to teach online.
Image	Alexander, Polyakova-Norwood, Johnston, Christensen, & Loquist (2003) Allen & Seaman(2012a) Allen & Seaman (2013) Allen & Seaman (2015) Bacow, Bowen, Guthrie, Lack, & Long (2012)	Faculty had concerns about how teaching online would affect their image. Faculty worried that teaching online would negatively affect their promotion and tenure process. Faculty were unsure about how their teaching in online courses would be assessed. Faculty were more negative about the legitimacy

	Christianson et al. (2002) Dooley & Murphrey (2000) Gaytan (2009) Green et al. (2009) Mason et al. (2010) McQuiggan (2012) Orr et al. (2009) Shea (2007) Stewart et al. (2010) Ulmer et al. (2007)	of online education than administrators were.
Job relevance	Allen & Seaman (2012a) Bacow et al. (2012) Bolliger & Wasilik (2009) Chao, Saj, & Hamilton (2010) DeGagne & Walters (2010) Gibson, Harris, & Colaric (2008) Haber & Mills (2008) Jaschik & Lederman (2014) Johnson (2008) McQuiggan (2012) Orr et al. (2009) Osborne, Kriese, Tobey, & Johnson (2009) Panda & Mishra (2007) Ryan et al. (2005) Seaman (2009) Shea et al. (2005) Shovein, Huston, Fox, & Damazo (2005) Stewart et al. (2010)	Faculty were concerned about interacting with students in online courses. Faculty valued collaboration to design online courses that were student-centric. Faculty were more satisfied teaching online when they believed students were achieving learning outcomes
Output quality	Adkins, Kenkel, & Lim (2005) Arend (2009) Arif (2001) Bacow et al. (2012) Bolliger & Wasilik (2009) Chapman, Davis, Toy, & Wright (2004) Green et al. (2009) Grijalva, Nowell, & Kerkvliet (2006) Haber & Mills (2008) Harmon, Lambrinos, & Buffolino (2010) King, Guyette, & Piotrowski (2009) Lackey (2011)	Faculty were concerned about the effectiveness of various forms of technology used in online courses. Faculty were concerned about students' technical skills, their access to equipment, and their abilities to use technology effectively in online courses. Faculty were concerned about the potential for students to cheat in online courses.

	Lee, Srinivasan, Trail, Lewis, & Lopez (2011) Luck & McQuiggan (2006) Mason et al. (2010) Mazzolini & Maddison (2007) McGee (2013) Mupinga, Nora, & Yaw (2006) Osika et al. (2009) Ryan et al. (2005) Sahin & Shelley (2008) Semple, Hatala, Franks, & Rossi (2010) Stuber-McEwen, Wiseley, & Hoggatt (2009) Trenholm (2007) Ward, Peters, & Shelley (2010) Watson & Sottile (2010)	
Result demonstrability	Alsofyani et al. (2012) Bacow et al. (2012) Bolliger & Wasilik (2009) Chao et al. (2010) Chapman (2011) Christianson et al. (2002) Conceição (2006) DeGagne & Walters (2010) Gautreau (2011) Gaytan (2009) Green et al. (2009) Haber & Mills (2008) Huang et al. (2011) Mason et al. (2010) McQuiggan (2012) Orr et al. (2009) Pandra & Mishra (2007) Seaman (2009) Shea (2007) Shea et al. (2005) Wang & Wang (2009)	Faculty were concerned about their workload in online courses. Extra time to teach online was a barrier for some faculty. Stipends could be an incentive to teach online. (No consistent compensation models for faculty teaching online were identified.) Flexibility was a strong incentive for faculty to teach online. Faculty valued professional development opportunities associated with teaching online. Faculty valued training, support, and mentoring to help them succeed in teaching online. Faculty were gratified when their online teaching was recognized publicly by their institution.

Perceived Ease of Use (PEU)

Research has shown that PEU of educational technology affects faculty satisfaction with teaching online. Shea, Pickett, and Li (2005) surveyed 913 faculty and determined through factor and multiple regression analysis that technological barriers were strongly correlated to levels of online faculty satisfaction. Bolliger and Wasilik (2009) later surveyed 102 instructors and found that the issue that most impacted faculty satisfaction with teaching online was struggles with technology. A number of researchers (Bolliger & Wasilik, 2009; Christianson, Tiene, & Luft, 2002; Conceição, 2006; DeGagne & Walters, 2010; Green, Alejandro, & Brown, 2009) also found that faculty were dissatisfied if they thought that using a system would take more time or increase their workload, factors that could be considered "ease" of use, especially if users were struggling to learn how to operate a system.

Another factor that should be considered in discussions of PEU is *computer self-efficacy*, or a person's beliefs about his or her competence using computers (Compeau & Higgins, 1995). Zhen, Garthwait, and Pratt (2008) determined that self-efficacy in using online course management applications effectively was the single most important factor affecting instructors' decision to adopt an application for online teaching. Other studies also showed correlations between faculty's computer self-efficacy and their intent to teach online or willingness to continue to teach online. Shea (2007) surveyed 386 faculty at 36 institutions and found that instructors who were more skilled in technology reported that they were also more willing to move new subject areas online. Similarly, Tabata and Johnsrud (2008) found in a study involving 2,048 participants that faculty's beliefs that they were skilled in using technology were significantly correlated with their intention to participate in online education. A smaller study (Osika, Johnson, & Buteau, 2009) surveying 36 participants at an urban university in the Midwest found that the number one factor influencing faculty's decision to use an LMS to move courses online was users' previous success with other technologies. Taken together, these studies indicated that faculty's confidence about their own computer skills played a critical role in their willingness to teach online.

Perceived Usefulness (PU)

The TAM2 constructs included seven factors (subjective norm, voluntariness, experience, image, job relevance, output quality, and result demonstrability) that have been shown to affect a user's PU of a system. We found that each of these factors has been addressed in studies of faculty teaching online.

Subjective norm. The TAM2 showed that users' understanding of the value of using a system is driven in part by their perceptions about whether others in an organization feel that they *should* use that system. In higher education, administrators often determine who will be teaching online and what kinds of technology they might use to do so. Research suggested that administrators who communicated reasons for why faculty should teach online could create a stronger subjective norm that might encourage faculty participation in online initiatives (Betts & Heaston, 2014; Huang et al., 2011; Wang & Wang, 2009; Wickersham & McElhany, 2010). Other studies noted faculty's desire for clearer institutional goals and policies concerning online education (Dooley & Murphrey, 2000; Orr, Williams, & Pennington, 2009) and their interest in playing a role in the development of these goals and policies (Maguire, 2009). Faculty also expressed a need for more institutional support in various forms, including enrollment caps, instructional design support, development of online faculty communities, and security or proctoring software (Chapman, 2011; Lee, 2001; Wickersham & McElhany, 2010). These studies suggested that faculty satisfaction with online teaching could improve if leaders who contribute to creating the subjective norm met communication and support needs.

Voluntariness. Venkatesh and Davis (2000) found that the subjective norm had a direct effect on intention to adopt a system when using that system was mandatory, but not when it was voluntary. Some researchers found that institutions that made training for teaching online mandatory saw gains in online faculty satisfaction, even if faculty were not initially enthusiastic about using LMS (Betts, 2009; Lackey, 2011; McQuiggan, 2012). However, other researchers found differences in motivating factors to teach online, depending on whether participation was mandatory or voluntary. Some of the studies that addressed motivating instructors to teach online looked at early adopters to learn more about how to encourage other faculty to use online teaching tools (Dooley & Murphrey, 2000; Hixon, Barczyk, Buckenmeyer, & Feldman, 2011; Jacobsen, 2000; Shea, 2007). These studies confirmed that early adopters - most of whom volunteered to teach online - had different motivations than other faculty. For example, Shea's (2007) research involving 386 faculty at 36 institutions revealed that volunteers were more motivated by intrinsic factors (renewed passion for teaching, opportunities to experiment with new pedagogical methods, etc.), while faculty who were required to teach online were more motivated by extrinsic factors such as compensation and job security. Jacobsen (2000) also found that early adopters were driven to use technology in innovative ways; however, the majority of instructors in that study were not early adopters, and they were hesitant to use new technologies until they understood what benefits they would gain from doing so.

Experience. The TAM2 confirmed that direct experience with technology affected users' subsequent intentions to use that technology. Other research on online education reinforced this idea. Studies showed that faculty adapted well to the online environment and were more satisfied as they gained more experience. Shea et al. (2005) reported that 90% of over 900 faculty surveyed immediately after teaching an online course were satisfied with developing and delivering online courses, and almost 98% of those faculty said they would like to teach online again. Another study (Ulmer, Watson, & Derby, 2007) surveyed 137 faculty and found significant differences in attitudes toward online education based on instructors' experience level. In this study, faculty who had more experience teaching online had significantly more positive perceptions of the overall effectiveness of instructor-student interaction and the ability to increase student performance in online courses.

Other research also revealed positive attitudes toward teaching online by faculty who had previously experienced online teaching. A survey of almost 11,000 faculty (Seaman, 2009) showed that 86.4% of faculty who were teaching an online course at the time of the survey had recommended an online course to a student. Another report by Allen and Seaman (2012a) demonstrated that faculty at institutions that offered more online courses and programs were more optimistic about online learning in general. Their survey showed that faculty who had taught online held the most positive views about it, with two-thirds of them reporting that they felt more excited than fearful about online education; in contrast, less than one-third (32.4%) of faculty who had *not* taught online or blended courses viewed online education with more excitement than fear (Allen & Seaman, 2012a). These reports suggest that, once faculty experience teaching online, they are more likely to be willing to continue to teach online courses.

Image. Research showed that many faculty had some anxiety about how teaching online would affect their status or their prestige at a university (Allen & Seaman, 2013; Green et al., 2009; Mason et al., 2010; Ulmer et al., 2007). Some of their concern was rooted in skepticism about the image of online education in general. For instance, in one study, almost 70% of faculty surveyed answered "No" to the question "Do you think an online degree is as prestigious as a traditional degree?" (Stewart et al., 2010).

Various studies have addressed image in terms of faculty's beliefs about whether learning outcomes in online courses were inferior to those in face-to-face classes (Allen & Seaman, 2012a; Allen & Seaman, 2015; Bacow et al., 2012; McQuiggan, 2012; Stewart et al., 2010). Many administrators are aware of this kind of skepticism on the part of faculty; in fact, a recent report by Allen and Seaman (2015) showed that administrators believed that online teaching has had a negative image among faculty for over a decade. Furthermore, in annual reports by Allen and Seaman since 2002, there has never been a majority of administrators who believed that their faculty accepted the "value and legitimacy of online education" (Allen & Seaman, 2015, p. 21).

Other studies showed that faculty were concerned about whether their role as an online educator might have some bearing on their promotion or tenure (Alexander, Polyakova-Norwood, Johnston, Christensen, & Loquist, 2003; Gaytan, 2009; Green et al., 2009; Mason et al., 2010; Orr et al., 2009; Shea, 2007). Some faculty worried that teaching online could make them more vulnerable and result in poor evaluations, thus threatening their job security (Dooley & Murphrey, 2000; Gaytan, 2009). Junior faculty members, in particular, were apprehensive about how their courses would be assessed for quality by the institution (Shea, 2007). Clearly, faculty had concerns about their image and the effects on their career as a result of teaching online.

Job relevance. The TAM2 showed that PU was affected by *job relevance*, or users' perceptions of the degree to which a system might be important in their jobs by allowing them to accomplish significant goals. In online education, faculty's perceptions about using technology to engage students and accomplish learning objectives have been critical issues related to job relevance, because of the importance of student progression as a measure of success. The idea that students might learn less in online courses is an issue of image, as we have already discussed, but it is also an issue of job relevance. Studies have shown that instructors were deeply concerned about students' ability to learn in online courses (Allen & Seaman, 2012a; Gibson et al., 2008; Osborne, Kriese, Tobey, & Johnson, 2009; Seaman, 2009; Shovein, Huston, Fox, & Damazo, 2005; Stewart et al., 2010). One survey of 10,700 faculty teaching online in the United States showed that 70% believed that learning outcomes for students in online courses were inferior or somewhat inferior to those experienced by students in face-to-face classes (Seaman, 2009). Another survey of 2,799 faculty and 288 campus administrators across the United States found that only 26% of faculty agreed or strongly agreed with a statement that student learning outcomes in online courses were at least equivalent to those in face-to-face ones; in contrast, 67% of campus administrators agreed or strongly agreed with that idea (Jaschik & Lederman, 2014).

Faculty have also claimed that they valued collaboration with instructional designers who could help them design their online courses to make them more student-centric (Chao, Saj, &

Hamilton, 2010; DeGagne & Walters, 2010; Johnson, 2008; McQuiggan, 2012; Ryan et al., 2005). This student-centric approach was also valued by the 10 faculty participating in a qualitative study by Orr and colleagues (2009), as all participants claimed that their major motivation in teaching online was meeting the needs of students. Clearly, in terms of job relevance, faculty were most concerned about their ability to help students thrive in an online learning environment.

Output quality. *Output quality* in the TAM2 concerned how well technology performed functions needed to accomplish specific tasks (Venkatesh & Davis, 2000). Research has shown that many faculty teaching online have been concerned about various technical aspects of learning management systems (LMS) and other educational technology (Bolliger & Wasilik, 2009; Green et al., 2009; Luck & McQuiggan, 2006; Ryan et al., 2005; Ward, Peters, & Shelley, 2010). Various researchers found that faculty who complained about feeling disengaged from their students did not find many forms of online communication (discussion boards, web conferencing, etc.) satisfactory for the level of interaction they desired (Arend, 2009; Haber & Mills, 2008; Mazzolini & Maddison, 2007; Ward et al., 2010). Some studies showed that simply accessing communication tools could be problematic, as faculty and/or students experienced issues with Internet connectivity, log-in problems, or manipulating the LMS (Lackey, 2011; Ward et al., 2010).

Another serious concern about technology among faculty teaching online has been the potential for students to cheat (Bacow et al., 2012; Chapman, Davis, Toy, & Wright, 2004; Haber & Mills, 2008; Trenholm, 2007). Some researchers suggested that these fears were unfounded, either because they found no significant difference between cheating in online and face-to-face classes (Grijalva, Nowell, & Kerkvliet, 2006) or they actually found that students cheated more in face-to-face classes (Stuber-McEwen, Wiseley, & Hoggatt, 2009; Watson & Sottile, 2010). Even so, faculty often perceived that students had more opportunities to cheat in online courses, and some research has supported this idea (Adkins, Kenkel, & Lim, 2005; Harmon, Lambrinos, & Buffolino, 2010; King, Guyette, & Piotrowski, 2009; Mason et al., 2010).

Faculty sometimes feared that students with strong technical skills could manipulate technology to their advantage, by finding technological loopholes to avoid taking tests or submitting assignments (McGee, 2013; Stuber-McEwen et al., 2009). In contrast, some instructors have worried that students might not have the technical skills to allow them to learn effectively in online environments. Various studies have shown that, even as online education has become more prevalent, students' essential technical skills for online learning have varied widely (Arif, 2001; Lee, Srinivasan, Trail, Lewis, & Lopez, 2011; Sahin & Shelley, 2008). Indeed, instructors' concerns about students' technical skills are output quality issues, since students' abilities to use technologies are essential for those technologies to be effective.

Result demonstrability. *Result demonstrability*, or perceived tangible results and benefits of using a technological system, also affected PU in the TAM2. A number of studies have noted benefits that faculty have received or would like to receive as a result of teaching online. The most obvious tangible results for employees are related to money and time. Issues of compensation, time, and workload have recurred throughout the literature on online teaching. In fact, Bolliger and Wasilik (2009) found that some of the most significant institutional issues affecting faculty satisfaction were a higher workload and increased time commitment for online instructors. Many other studies have acknowledged issues of substantial time commitments and workload for online faculty (Bacow et al., 2012; Christianson et al., 2002; Conceição, 2006; DeGagne & Walters, 2010; Green et al., 2009; Haber & Mills, 2008; Mason et al., 2010). Most research has shown that time commitment and workload were barriers or demotivators to faculty, though some studies claimed that the extra time commitment did not affect faculty satisfaction or preference for teaching online (Christianson et al., 2002; Orr et al., 2009; Shea et al., 2005). Some researchers also determined that faculty would not mind the extra time it took to deliver online courses if they were adequately compensated (Haber & Mills, 2008; Huang et al., 2011; Shea, 2007).

Many studies have cited one main factor that played a role in faculty's satisfaction with online education: the flexibility of teaching online courses. Recurring throughout the reviewed studies was the idea that faculty appreciated the fact that online education was not bound by time or space. For example, Shea's (2007) survey of 386 faculty teaching online in 36 colleges found that the top motivator was a flexible work schedule. Similarly, Green et al. (2009) determined that 82.22% of 135 faculty they surveyed claimed they enjoyed the flexibility of online instruction. Additional research by Chapman (2011) included full-time and part-time instructors; the author surveyed 294 tenured/tenure-track and adjunct instructors and found that the strongest motivation for both groups to teach online was a flexible schedule.

A variety of tangible rewards for instructors teaching online was found to be an effective means to attract faculty to teach online. For example, some researchers discovered that faculty viewed new technologies positively because learning about them was a professional development opportunity or a way to grow intellectually (Chapman, 2011; Green et al., 2009; McQuiggan, 2012; Pandra & Mishra, 2007; Seaman, 2009). Faculty also appreciated access to high-quality training and support programs and other forms of mentoring to help them be successful teaching online (Alsofyani et al., 2012; Chao et al., 2010; Chapman, 2011; Green et al., 2009; McQuiggan, 2012; Shea et al., 2005; Wang & Wang, 2009). Other studies showed that instructors were motivated when their achievements in teaching online were highlighted or recognized with an award by their institution (Bacow et al., 2012; Gautreau, 2011; Mason et al., 2010). In general, faculty in these studies expressed that teaching online afforded them opportunities for professional growth and allowed them more control over their own schedules.

Conclusions

Even though we initially found few studies applying the TAM or the TAM2 to faculty acceptance of technology for online teaching, this literature review revealed that researchers have addressed various elements of the TAM2 in terms of faculty teaching online. Table 1 shows a myriad of issues, attitudes, and concerns aligning with TAM2 constructs that our reviewed studies addressed in different ways. Exploring these empirical studies in this context provided a lens to better understand faculty perceptions, not only of the user-friendliness and usefulness of technological tools, but of the overall experience of teaching online. This synthesis is important, since recognizing faculty's needs and desires in their roles as instructors is critical for institutions offering online courses and programs. In addition, understanding more about how faculty accept and implement technology used for online learning could help higher education administrators promote positive attitudes and support faculty efforts to foster student success in online courses.

Some studies in this review indicated a gap between views of administrators and faculty concerning the usefulness of online education. Instructors were particularly concerned about issues affecting student success, such as effective communication, technical proficiency, and legitimate achievement of learning outcomes without cheating. Faculty also were concerned about their own status as online instructors in the larger institutional culture. Some instructors worried about how teaching online would affect their image. Instructors were also unsure about how their online teaching would be evaluated, particularly in promotion and tenure processes. Issues of time commitment and workload were viewed as barriers to teaching online as well.

At the same time, many of the studies showed that instructors adapted well to the online environment as they gained more experience. Faculty who were teaching online were gratified when institutions provided mentoring, training, support, and recognition of their success. Instructors also valued the personal and professional rewards that resulted from their online teaching, such as flexible schedules and professional development opportunities. The reviewed studies also found that, whether or not they were already teaching online, faculty's perceptions about the user-friendliness of technology and their own skills in mastering LMS and other tools played a role in their satisfaction with online teaching and learning.

This review also highlighted gaps in research concerning faculty's experiences in teaching online. The TAM2 provided a well-defined framework for understanding faculty perceptions, and more research that directly applies the TAM2 to faculty teaching online is warranted. Further studies could also shed more light on faculty's perceptions of how teaching online would help them achieve their goals (*job relevance*) and accomplish essential tasks (*output quality*). More research could also help administrators understand how to communicate the value of online education to faculty in terms that resonate with them. Exploring more about potential tangible benefits for faculty teaching online (*result demonstrability*) could also help institutions address issues of compensation and workload.

The knowledge gleaned from this literature review has significant implications for institutions that seek to build and maintain strong online programs. Knowing more about faculty's views about the user-friendliness of technology used to teach online (including LMS, web conferencing tools, or other technology used for communicating or delivering content via the Internet) could help universities determine how to train faculty to teach in their online courses and programs. It is also important to understand more about faculty's computer self-efficacy to assist them in using technologies to their fullest potential. Administrators who determine teaching assignments and cap class enrollments could address barriers such as time commitment and workload for online teaching. Academic leaders responsible for strategic plans could also involve faculty in planning processes and clearly communicate institutional mission and goals for their online programs.

Using the TAM2 as a framework to explore research on faculty satisfaction with teaching online also highlighted important factors for institutions to consider if they want their faculty to thrive in online teaching and learning environments. The various facets of the model provide guideposts for universities to focus on in recruiting and retaining online instructors. Understanding faculty's computer self-efficacy could help institutions plan what kind of training programs would be needed to encourage more instructors to teach online. Recognizing the need for continuous training and support, even for experienced instructors, could convince administrators to leverage resources to provide these kinds of programs. Indeed, academic leaders who are aware of the effects of social influence processes could use their power to change institutional culture. By consistently supporting faculty and demonstrating to them that their efforts are just as valuable as those of instructors teaching face-to-face, administrators could create a strong positive image of online education at their institution.

References

Adkins, J., Kenkel, C., & Lim, C. (2005). Deterrents to online academic dishonesty. *The Journal of Learning in Higher Education*, *1*(1), 17-22. Retrieved from http://jwpress.com/JLHE/Issues/v1i1/Deterrents%20to%20Online%20Academic%20Dishonesty.pdf

Alexander, J., Polyakova-Norwood, V., Johnston, L., Christensen, P., & Loquist, R. (2003). Collaborative development and evaluation of an online nursing course. *Distance Education*, *24*(1), 41-56. doi:10.1080/01587910303046

Allen, I. E. & Seaman, J. (2012a). Conflicted: Faculty and online education. Babson Survey Research Group and Quahog Research Group, LLC. Retrieved from http://www.insidehighered.com/sites/default/server_files/files/IHE-BSRG-Conflict.pdf

Allen, I. E. & Seaman, J. (2012b). Digital faculty: Professors, teaching and technology, 2012. Babson Survey Research Group and Quahog Research Group, LLC. Retrieved from https://www.insidehighered.com/sites/default/server_files/DigitalFaculty.htm

Allen, I. E. & Seaman, J. (2013). Changing course: Ten years of tracking online education in the United States. Babson Survey Research Group and Quahog Research Group, LLC. Retrieved from http://www.onlinelearningsurvey.com/reports/changingcourse.pdf

Allen, I. E., & Seaman, J. (2015). Grade level: Tracking online education in the United States. Babson Survey Research Group and Quahog Research Group, LLC. Retrieved from http://www.onlinelearningsurvey.com/reports/gradelevel.pdf

Alsofyani, M., Aris, B., Eynon, R., & Majid, N. (2012). A preliminary evaluation of short blended online training workshop for TPACK development using technology acceptance model. *The Turkish Online Journal of Educational Technology, 11*(3), 20-32. Retrieved from http://www.tojet.net/articles/v11i3/1133.pdf

Arend, B. (2009). Encouraging critical thinking in online threaded discussions. *The Journal of Educators Online, 6*(1). Retrieved from http://www.thejeo.com/Archives/Volume6Number1/Arendpaper.pdf

Arif, A. (2001). Learning from the web: Are students ready or not? *Educational Technology & Society, 4*(4), 32-38. Retrieved from http://www.ifets.info/others/journals/4_4/arif.html

Bacow, L., Bowen, W., Guthrie, K., Lack, K., & Long, M. (2012). Barriers to adoption of online learning systems in U.S. higher education. Ithaka S+R Consulting. Retrieved from http://www.sr.ithaka.org/research-publications/barriers-adoption-online-learning-systems-us-higher-education

Betts, K. (2009). Online Human Touch (OHT) training & support: A conceptual framework to increase faculty engagement, connectivity, and retention in online education, part 2. *Journal of Online Learning and Teaching, 5*(1). Retrieved from http://jolt.merlot.org/vol5no1/betts_0309.htm

Betts, K., & Heaston, A. (2014). Build it but will they teach? Strategies for increasing faculty participation and retention in online and blended education. *Online Journal of Distance Learning Administration, 17,* (2). Retrieved from http://www.westga.edu/~distance/ojdla/summer172/betts_heaston172.html

Bolliger, D., & Wasilik, O. (2009). Factors influencing faculty satisfaction with online teaching and learning in higher education. *Distance Education, 30*(1), 103-116. doi:10.1080/01587910902845949

Chao, I., Saj, T., & Hamilton, D. (2010). Using collaborative course development to achieve online course quality standards. *International Review of Research in Open and Distributed Learning, 11*(3), 106-126. Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/912

Chapman, D. (2011). Contingent and tenured/tenure-track faculty: Motivations and incentives to teach distance education courses. *Online Journal of Distance Learning Administration, 14*(3), 1-15. Retrieved from http://www.westga.edu/~distance/ojdla/fall143/chapman143.html

Chapman, K., Davis, R., Toy, D., & Wright, L. (2004). Academic integrity in the business school environment: I'll get by with a little help from my friends. *Journal of Marketing Education, 26*(3), 236-249. doi:10.1177/0273475304268779

Christianson, L., Tiene, D., & Luft, P. (2002). Web-based teaching in undergraduate nursing programs. *Nurse Educator*, *27*(6), 276-282. doi:10.1097/00006223-200211000-00009

Compeau, D., & Higgins, C. (1995). Computer self-efficacy: Development of a measure and initial test. *MIS Quarterly*, *19*(2), 189-211. doi:10.2307/249688

Conceição, S. (2006). Faculty lived experiences in the online environment. *Adult Education Quarterly, 57*(1), 26-45. doi:10.1177/1059601106292247

Davis, F. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319-340. doi:10.2307/249008

Davis, F. (1993). User acceptance of information technology: System characteristics, user perceptions and behavioral impacts. *International Journal of Man-Machine Studies, 38*, 475-487. doi:10.1006/imms.1993.1022

DeGagne, J., & Walters, K. (2010). The lived experience of online educators: Hermeneutic phenomenology. *Journal of Online Learning and Teaching, 6*(2). Retrieved from http://jolt.merlot.org/vol6no2/degagne_0610.htm

Dooley, K., & Murphrey, T. P. (2000). How the perspectives of administrators, faculty and support units impact the rate of distance education adoption. *Online Journal of Distance Learning Administration, 3*(4). Retrieved from http://www.westga.edu/~distance/ojdla/winter34/dooley34.html

Gautreau, C. (2011). Motivational factors affecting the integration of a learning management system by faculty. *The Journal of Educators Online*, *8*(1). Retrieved from http://files.eric.ed.gov/fulltext/EJ917870.pdf

Gaytan, J. (2009). Analyzing online education through the lens of institutional theory and practice: The need for research-based and -validated frameworks for planning, designing, delivering, and assessing online instruction. *The Journal of Research in Business Education, 51*(2), 62-75. Retrieved from http://www.dpe.org/publications.htm

Gibson, S., Harris, M., & Colaric, S. (2008). Technology acceptance in an academic context: Faculty acceptance of online education. *Journal of Education for Business, 83*(6), 355-359. doi:10.3200/JOEB.83.6.355-359

Green, T., Alejandro, J., & Brown, A. (2009). The retention of experienced faculty in online distance education programs: Understanding factors that impact their involvement. *The International Review of Research in Open and Distance Learning, 10*(3), 1-15. Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/683/1279

Grijalva, T., Nowell, C., & Kerkvliet, J. (2006). Academic honesty and online courses. *College Student Journal*, *40*(1), 180-185. Retrieved from http://www.projectinnovation.com/college-student-journal.html

Haber, J., & Mills, M. (2008). Perceptions of barriers concerning effective online teaching and policies: Florida community college faculty. *Community College Journal of Research and Practice, 32*, 266-283. doi:10.1080/10668920701884505

Hagedorn, L.S. (2000). Conceptualizing faculty job satisfaction: Components, theories, and outcomes. *New Directions for Institutional Research, 2000*(105), 5-20. doi: 10.1002/ir.10501

Harmon, O., Lambrinos, J., & Buffolino, J. (2010). Assessment design and cheating risk in online instruction. *Online Journal of Distance Learning Administration, 13*(3). Retrieved from http://www.westga.edu/~distance/ojdla/Fall133/harmon_lambrinos_buffolino133.html

Hixon, E., Barczyk, C., Buckenmeyer, J., & Feldman, L. (2011). Mentoring university faculty to become high quality online educators: A program evaluation. *Online Journal of Distance Learning Administration, 14*(5). Retrieved from http://distance.westga.edu/~distance/ojdla/winter144/hixon_Barczyk_Buckenmeyer_feldman144.html

Huang, R.T., Deggs, D. M., Jabor, M. K., & Machtmes, K. (2011). Faculty online technology adoption: The role of management support and organizational climate. *Online Journal of Distance Learning Administration*, *14*(2). Retrieved from http://www.westga.edu/~distance/ojdla/summer142/huang_142.html

Jacobsen, D. M. (2000). Examining technology adoption patterns by faculty in higher education. Proceedings of ACEC2000: Learning Technologies, Teaching and the Future of Schools, July 6 to 9, Melbourne, Australia. Retrieved from http://people.ucalgary.ca/~dmjacobs/acec/

Jaschik, S., & Lederman, D. (2014). The 2014 Inside Higher Ed survey of faculty attitudes on technology. Gallup and Inside Higher Ed. Retrieved from https://www.insidehighered.com/node/71103/download/3543a7f4860331d55489c20b4f6fa059

Johnson, A. (2008). A nursing faculty's transition to teaching online. *Nursing Education Perspectives 29*(1), 17-22. Retrieved from http://www.nln.org/nlnjournal/

King, C., Guyette, R., & Piotrowski, C. (2009). Online exams and cheating: An empirical analysis of business students' views. *The Journal of Educators Online, 6*(1). Retrieved from http://www.thejeo.com/Archives/Volume6Number1/Kingetalpaper.pdf

King, W., & He, J. (2006). A meta-analysis of the technology acceptance model. *Information & Management*, *43*(6), 740-755. doi:10.1016/j.im.2006.05.003

Lackey, K. (2011). Faculty development: An analysis of current and effective training strategies for preparing faculty to teach online. *Online Journal of Distance Learning Administration, 14*(5). Retrieved from http://www.westga.edu/~distance/ojdla/winter144/lackey144.html

Lee, J. (2001). Instructional support for distance education and faculty motivation, commitment, satisfaction. *British Journal of Educational Technology, 32*(2), 153-160. doi:10.1111/1467-8535.00186

Lee, S., Srinivasan, S., Trail, T., Lewis, D., & Lopez, S. (2011). Examining the relationship among student perception of support, course satisfaction, and learning outcomes in online learning. *Internet & Higher Education, 14*(3), 158-163. doi:10.1016/j.iheduc.2011.04.001

Luck, A., & McQuiggan, C. (2006). Discovering what faculty really need to know about teaching online. Proceedings of the 22nd Annual Conference on Distance Teaching & Learning, August 2-6. Madison, WI. Retrieved from http://www.uwex.edu/disted/conference/Resource_library/proceedings/06_4160.pdf

Maguire, L. (2009). The faculty perspective regarding their role in distance education policy making. *Online Journal of Distance Learning Administration, 12*(1). Retrieved from http://www.westga.edu/~distance/ojdla/spring121/maguire121.html

Marangunić, N., & Granić, A. (2015). Technology acceptance model: a literature review from 1986 to 2013. *Universal Access in the Information Society, 14*: 81-95. doi: 10.1007/s10209-014-0348-1

Mason, J., Hickman, C., Dyer, A., Koproske, C., Fry, G., & Taha, M. (2010). Engaging faculty in online education: Rightsizing incentives and optimizing support. Washington, DC: University Leadership Council of The Advisory Board Company. Retrieved from http://www20.csueastbay.edu/oaa/files/student_success/EngFacOnlineEd.pdf

Mazzolini, M., & Maddison, S. (2007). When to jump in: The role of the instructor in online discussion forums. *Computers & Education, 49*(2), 193-213. Retrieved from doi:10.1016/j.compedu.2005.06.011

McGee, P. (2013). Supporting academic honesty in online courses. *Journal of Educators Online, 10*(1). Retrieved from http://www.thejeo.com/Archives/Volume10Number1/McGee.pdf

McQuiggan, C. (2012). Faculty development for online teaching as a catalyst for change. *Journal of Asynchronous Learning Networks, 16*(2), 27-61. Retrieved from http://files.eric.ed.gov/fulltext/EJ971044.pdf

Mupinga, D., Nora, R., & Yaw, D. (2006). The learning styles, expectations, and needs of online students. *College Teaching, 54*(1), 185-189. doi:10.3200/CTCH.54.1.185-189

Orr, R., Williams, M., & Pennington, K. (2009). Institutional efforts to support faculty in online teaching. *Innovative Higher Education, 34*(4), 257-268. doi:10.1007/s10755-009-9111-6

Osborne, R., Kriese, P., Tobey, H., & Johnson, E. (2009). And never the two shall meet? Student vs. faculty perceptions of online courses. *Journal of Educational Computing Research, 40*(2), 171-182. doi:10.2190/EC.40.2.b

Osika, E., Johnson, R., & Buteau, R. (2009). Factors influencing faculty use of technology in online instruction: A case study. *Online Journal of Distance Learning Administration, 12*(1). Retrieved from http://www.westga.edu/~distance/ojdla/spring121/osika121.html

Panda, S., & Mishra, S. (2007). E-learning in a mega open university: Faculty attitude, barriers and motivators. *Educational Media International, 44*(4), 323-338. doi:10.1080/09523980701680854

Ryan, M., Hodson-Carlton, K., & Ali, N. (2005). A model for faculty teaching online: Confirmation of a dimensional matrix. *Journal of Nursing Education, 44*(8), 357-365. Retrieved from http://www.nottingham.ac.uk/~ntzcl1/literature/teaching/ryan-2005.pdf

Sahin, I., & Shelley, M. (2008). Considering students' perceptions: The distance education student satisfaction model. *Educational Technology & Society, 11*(3), 216-223. Retrieved from http://ifets.info/journals/11_3/15.pdf

Seaman, J. (2009). "Volume II: The paradox of faculty voices: Views and experiences with online learning." In "Online learning as a strategic asset: Results of a national faculty survey," part of the online education benchmarking study conducted by the APLU Sloan National Commission on Online Learning. Retrieved from http://olc.onlinelearningconsortium.org/publications/survey/APLU_Reports

Semple, M., Hatala, J., Franks, P., & Rossi, M. (2010). Is your avatar ethical? On-line course tools that are methods for student identity and verification. *Journal of Educational Technology Systems, 39*(2), 181-191. doi:10.2190/ET.39.2.h

Shea, P. (2007). Bridges and barriers to teaching online college courses: A study of experienced online faculty at 36 colleges. *Journal of Asynchronous Learning Networks, 11*(2), 73-128. Retrieved from http://www.sunyresearch.net/hplo/?page_id=13

Shea, P., Pickett, A., & Li, C. (2005). Increasing access to higher education: A study of the diffusion of online teaching among 913 college faculty. *The International Review of Research in Open and Distance Learning, 6*(2). Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/238

Shovein, J., Huston, C., Fox, S., & Damazo, B. (2005). Challenging traditional teaching and learning paradigms: Online learning and emancipatory teaching. *Nursing Education Perspectives, 26*(6), 340-343. Retrieved from http://www.nln.org/nlnjournal/

Stewart, C., Bachman, C., & Johnson, R. (2010). Predictors of faculty acceptance of online education. *MERLOT Journal of Online Learning and Teaching, 6*(3), 597-616. Retrieved from http://jolt.merlot.org/vol6no3/stewartc_0910.pdf

Stuber-McEwen, D., Wiseley, P. & Hoggatt, S. (2009). Point, click, and cheat: Frequency and type of academic dishonesty in the virtual classroom. *Online Journal of Distance*

Learning Administration, 12(3). Retrieved from http://www.westga.edu/~distance/ojdla/fall123/stuber123.html

Tabata, L., & Johnsrud, L. (2008). The impact of faculty attitudes toward technology, distance education, and innovation. *Research in Higher Education, 49*(7), 625-646. doi:10.1007/s11162-008-9094-7

Trenholm, S. (2007). A review of cheating in fully asynchronous online courses: A math or fact-based course perspective. *Journal of Educational Technology Systems, 35*(3), 281-300. doi:10.2190/Y78L-H21X-241N-7Q02

Ulmer, L., Watson, L., & Derby, D. (2007). Perceptions of higher education faculty members on the value of distance education. *The Quarterly Review of Distance Education, 8*(1), 59-70. Retrieved from http://faculty.weber.edu/eamsel/Research%20Groups/On-line%20Learning/Umer%20et%20al.%20%282007%29.pdf

Venkatesh, V., & Davis, F. (1996). A model of the antecedents of perceived ease of use: Development and test. *Decision Sciences, 27*(3), 451- 481. doi:10.1111/j.1540-5915.1996.tb00860.x

Venkatesh, V., & Davis, F. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science, 46*(2), 186-204. doi:10.1287/mnsc.46.2.186.11926

Wang, W., & Wang, C. (2009). An empirical study of instructor adoption of web-based learning systems. *Computers & Education, 53*(3), 761-774. doi:10.1016/j.compedu.2009.02.021

Ward, M., Peters, G., & Shelley, K. (2010). Student and faculty perceptions of the quality of online learning experiences. *International Review of Research in Open & Distance Learning, 11*(3), 57-77. Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/867

Watson, G., & Sottile, J. (2010). Cheating in the digital age: Do students cheat more in online courses? *Online Journal of Distance Learning Administration, 13*(1). Retrieved from http://www.westga.edu/~distance/ojdla/spring131/watson131.html

Wickersham, L., & McElhany, J. (2010). Bridging the divide: Reconciling administrator and faculty concerns regarding online education. *Quarterly Review of Distance Education, 11*(1), 1-12. Retrieved from http://www.infoagepub.com/quarterly-review-of-distance-education.html

Zhen, Y., Garthwait, A., & Pratt, P. (2008). Factors affecting faculty members' decision to teach or not to teach online in higher education. *Online Journal of Distance Learning Administration, 11*(3). Retrieved from http://www.westga.edu/~distance/ojdla/fall113/zhen113.html

Course Management System's Compatibility with Teaching Style Influences Willingness to Complete Training

Audrey Smith Pereira, Ph.D.
Fitchburg State University, Fitchburg, Massachusetts

Monika Maya Wahi
Laboure College, Milton, Massachusetts
Vasanta Health Science LLC, Cambridge, Massachusetts

Abstract

Although course management systems (CMSs) provide technology platforms that help faculty members adopt better techniques for teaching and learning, and training contributes to faculty information technology (IT) use, many higher education faculty members do not complete CMS training programs, resulting in underuse of CMSs. Therefore, the overall purpose of this research was to address how instructor perceptions influence willingness to complete IT training on CMSs, and to discern techniques university administrators can implement to improve training completion rates and, ultimately, CMS adoption rates. The basic design of the study was a cross-sectional survey. Data were obtained from 102 public university faculty members who responded to an anonymous, web-based survey about their perceptions of the relative advantage, compatibility, complexity, trialability, and observability of their institution's CMS. The data were analyzed using multiple linear regression models. Compatibility, defined as the degree to which instructors perceive the CMS as being consistent with their existing values, past experiences, and current or future teaching needs, was statistically significantly associated with willingness to complete online and in-person CMS training after controlling for other factors. Major findings suggest that faculty training on the CMS is not "one size fits all." If greater use of CMSs by faculty is to be achieved, university administrators should consider compatibility of teaching style with CMS adoption when developing and promoting CMS training.

Keywords: higher education faculty members, course management system, technology adoption, educational technology, faculty training and development, diffusion of innovation theory

Pereira, A. S. & Wahi, M.M. (2017). Course management system's compatibility with teaching style influences willingness to complete training. *Online Learning 21*(1), 36-59. doi: 10.24059/olj.v21i1.763

Introduction

Many higher education administrators offer course management systems (CMSs) to their faculty members to use in their courses (Green, 2010), and these CMSs help instructors improve teaching and learning (Tsai & Talley, 2013; Yidana, Sarfo, Edwards, Boison, & Wilson, 2013). However, CMS adoption rates by faculty are low (Green, 2010; Unwin et al., 2010). The lack of faculty training on information technology (IT) is one factor that contributes to low faculty IT adoption rates (deNoyelles, Cobb, & Lowe, 2012; Goktas, Yildirim, & Yildirim, 2009; Masalela, 2009; Smolin & Lawless, 2011). Yet, researchers have found that many faculty members are unwilling to complete IT training (Hassan, 2011; Hurtado, Eagan, Pryor, Pereira, 2015; Whang, & Tran, 2012). Faculty members who do not complete IT training on the CMS will be less likely to adopt the CMS, resulting in lost opportunities to increase the quality of teaching and learning at their institutions.

Therefore, this study examined faculty members' perceptions of their organization's CMS that may influence their willingness to complete IT training on the CMS. The research was grounded in components of Rogers' (2003) diffusion of innovations (DOI) theory. According to Rogers (2003), five perceived attributes of an innovation partially explain technology adoption: relative advantage, compatibility, complexity, trialability, and observability. Rogers asserted that perceived relative advantage, compatibility, trialability, and observability of an innovation relates positively to its adoption rate, whereas an innovation's perceived complexity has a negative influence on its adoption.

Research has shown that the quality of teaching and learning increases if faculty members more broadly adopt their organization's CMS (Tsai & Talley, 2013; Yidana et al., 2013). Research has also shown that faculty training on their CMS improves faculty adoption of these systems (deNoyelles et al., 2012; Hixon, Buckenmeyer, Barczyk, Feldman, & Zamojski, 2012; McBride & Thompson, 2011), but unfortunately, the rate of faculty training on CMSs is low (Hassan, 2011; Hurtado et al., 2012; Pereira, 2015). Therefore, increasing faculty willingness to complete CMS training on their organization's CMS, the topic of this research, should ultimately lead to higher CMS adoption rates by faculty members, and consequently, improved quality of teaching and learning in higher education.

This paper will first provide a review of the literature associated with faculty adoption and willingness to complete training on educational technology, including CMSs. Next, the research questions and methodology for the study will be described. Results will then be presented, followed by a discussion, which will include recommendations for administrative approaches to improving CMS training completion rates among faculty members, as well as recommendations for future research.

Review of Related Literature

The literature indicates that the use of IT positively contributes to teaching and learning in the higher education classroom (Archambault, Wetzel, Foulger, & Williams, 2010; Newhouse, Buckley, Grant, & Idzik, 2013). Consequently, CMSs, including Blackboard, have been developed to improve the teaching and learning process; to facilitate this goal, they also offer online course management tools (Blackboard, Inc., 2017). This suggests that use of CMSs, such as Blackboard, in the higher education classroom has a considerable potential to increase the quality of teaching and learning. This contention is supported by research conducted by Yidana et al. (2013) who found that learning was improved through the provision of a CMS that permitted students to control their learning process and learn independently, as well as research conducted by Tsai and Talley (2013) who reported that foreign language students' reading comprehension improved when they used a CMS. Also, Unal and Unal (2011) described a study in which students rated different teaching and learning functions within two CMSs (Blackboard and Moodle) but that, regardless of the CMS, the students rated these teaching and learning functions highly, indicating that students were ready to adopt CMSs. Additionally, from their research on course design and delivery elements that affect student satisfaction, Simon, Jackson, and Maxwell (2013) suggested that CMSs are valuable scholastic tools, in that they can represent "a rigorous alternative or supplement to traditional instruction" (p. 112). However, they also concluded that professors should not be replaced by CMSs in the learning process.

Although researchers have found that IT has the potential to improve the quality of teaching and learning in higher education (Archambault et al., 2010; Newhouse et al., 2013), and CMSs are now commonly present in higher education (Green, 2010) many faculty members are slow to integrate IT (and CMSs) into their classrooms (Abrahams, 2010; Bothma & Cant, 2011; Unwin et al., 2010; Yohon & Zimmerman, 2006) and resist using IT for teaching and learning (Hicks, 2011). Additionally, faculty members are more proficient in basic rather than high-level technologies (Allen & Seaman, 2012; Chitiyo & Harmon, 2009; Kinuthia, 2005; Rocca, 2010), and Allen and Seaman (2012) as well as Ertmer and Ottenbreit-Lefwich (2010) suggested that faculty are more likely to use IT to facilitate traditional rather than new instructional techniques. Also, although the literature suggests that faculty IT training is one factor that influences adoption of IT in the classroom (deNoyelles et al., 2012; Goktas et al., 2009; Kidd, 2010; Masalela, 2009; McBride & Thompson, 2011; Porter, 2011; Potter & Rockinson-Szapkiw, 2012; Samarawickrema & Stacey, 2007; Smolin & Lawless, 2011), it also suggests that many faculty members are unwilling to complete formal IT training (Hassan, 2011; Hurtado et al., 2012; Pereira, 2015; Yohon & Zimmerman, 2006).

Because of the low faculty adoption rates (Green, 2010; Unwin et al., 2010) and costs associated with implementing a CMS at higher education organizations, many researchers have focused on studying barriers to CMS adoption, as well as studying factors that may improve faculty adoption rates (Bennett & Bennett, 2003; Green, 2010; Keesee & Shepard, 2011; Mallinson & Krull, 2013; Samarawickrema & Stacey, 2007; West, Waddoups, & Graham, 2007). For example, Bennett and Bennett (2003) studied 20 higher education faculty members and concluded that workshop-based training improves faculty attitudes toward the CMS, and West et al. (2007) asserted that this indicates that faculty training increases the probability of CMS adoption.

Based on this body of research related to barriers to IT adoption in general and CMS adoption specifically, and factors that may influence adoption, some writers have recommended improvements to faculty IT training, as a way to improve instructional IT adoption. This is because offering higher-quality training may result in increased faculty willingness to complete the training. These recommendations include the following: offering pedagogical as well as technological training (Calderon et al., 2012; Iorio, Kee, & Decker, 2012; Kidd, 2010; Mark, Thadani, Santandreu Calonge, Pun, & Chiu, 2011; Samarawickrema & Stacey, 2007); developing research-based technology training programs (Onyia & Onyia, 2011); ensuring IT training is relevant to faculty needs (Kidd, 2010) and is accessible (Keengwe, Kidd, & Kyei-Blankson, 2009); requiring training as part of employment obligations (Onyia & Onyia, 2011); aligning IT training with institutional policies and procedures (Korr, Derwin, Greene, & Sokoloff, 2012); and offering in-person as well as online training (Kidd, 2010).

Scholars have also researched the influence of CMS training specifically (Allen & Seaman, 2012; Bennett & Bennett, 2003; Samarawickrema & Stacey, 2007). Samarawickrema and Stacey (2007) asserted that CMS training is more valuable to faculty members if it is applicable, appropriate, timely, and relevant. Further, Allen and Seaman (2012) found that while administrators rated their CMS training offerings as high quality, faculty attitudes about the CMS training were less positive.

However, fewer studies focused on the factors that influence instructors' willingness to attend, and presumably complete, IT training (on the institution's CMS or otherwise). This body of literature indicates that the following factors influence faculty willingness to attend or complete training: professional growth (Kinuthia, 2005); time away from duties (Kinuthia, 2005; Sandford, Dainty, Belcher, & Frisbee, 2011); free hardware/software (Kinuthia, 2005); timing of training programs (Roman, Kelsey, & Lin, 2010; Sandford et al., 2011), skill level (Chen et al., 2000); travel distance (Sandford et al., 2011); teaching experience (Sandford et al., 2011); and specific pedagogical competencies (Carril, Sanmamed, & Sellés, 2013). These studies also indicated that incentives, including monetary rewards, release time, and positive impacts on tenure and promotion encourage faculty to attend IT training (Kinuthia, 2005; Sandford et al., 2011). These results should be considered alongside studies that found that incentives are an important factor in improving faculty IT adoption rates (Allen & Seaman, 2012; Al-Senaidi, Lin, & Poirot, 2009; Aremu, Fakolujo, & Oluleye, 2013; Keengwe et al., 2009; Masalela, 2009; McKissic, 2012; Yidana et al., 2013).

The current state of the literature on this subject suggests that research on factors that improve higher education faculty's willingness to complete CMS training, both online and in-person, are lacking. Thus, this study aims to fill this important gap in the literature. Although higher education administrators invest considerable portions of their institutions' budgets in providing high-quality CMS services (Green, 2010) and CMS training (Meyer, 2014; Pereira, 2015), unfortunately, many faculty members are unwilling to complete the CMS training offered (Hassan, 2011; Hurtado et al., 2012; Pereira, 2015). This contributes to low faculty CMS adoption rates, resulting in lower quality teaching and learning than would be possible with CMS adoption, and lost opportunities to improve student learning experiences in higher education.

This study explored how higher education faculty perceptions of the relative advantage, compatibility, complexity, trialability, and observability (as defined by Rogers' (2003) DOI theory) of their institution's CMS influence their willingness to complete online and in-person IT training on use of the CMS. Specific research questions were the following: (a) What is the relative contribution of faculty perceptions of the *relative advantage* of using their institution's

CMS in teaching and learning to their willingness to complete online and in-person IT training on the CMS? (b) What is the relative contribution of faculty perceptions of the *compatibility* of using their institution's CMS to their willingness to complete online and in-person IT training on the CMS? (c) What is the relative contribution of faculty perceptions of the *complexity* of the CMS to their willingness to online and in-person IT training on the CMS? (d) What is the relative contribution of faculty perceptions of the *trialability* of their CMS to their willingness to complete online and in-person IT training on the CMS? (e) and What is the relative contribution of faculty perceptions of the *observability* of their CMS to their willingness to complete online and in-person IT training on the CMS?

Methods

Study Population/Sampling

All 392 full-time and part-time faculty members who taught undergraduate and graduate courses at Fitchburg State University (FSU) in Fitchburg, Massachusetts, a public university in the northeast United States, were invited to participate in an anonymous, web-based survey in late 2014. The survey included questions about their demographics, perceptions of their institution's CMS, and willingness to complete CMS training. The response rate was 29%.After exclusions for ineligible responses, 102 surveys were used for data analysis, yielding a revised response rate of 26%.

Forty-seven percent of the respondents were male and 46% female, while eight respondents did not identify their gender. At FSU, the following ranks are available: instructor, assistant professor, associate professor, and full professor (called "professor"). The instructor level can be occupied by full-time faculty, but is typically the level assigned to adjunct and part-time faculty. In the sample, 26% were instructors, 24% assistant professors, 23% associate professors, and 27% professors. Respondents taught in the following departments: 34% Science, Technology, Engineering, and Mathematics; 16% Social Science, Economics, History, and Political Science; 15% Education, Communication, and Game Design; and 36% taught in other departments, including Business Administration, English Studies, Industrial Technology, Interdisciplinary Studies, and Nursing.

Data Collection

After obtaining approval from the appropriate Institutional Review Boards for the Protection of Human Subjects in Research (IRB), data were collected anonymously using a public link through SurveyMonkey (SurveyMonkey, 2015). The survey questions that measured perceptions of the CMS were considered independent variables in this study, and were based on subscales developed by Keesee (2010). Keesee (2010) named her instrument the CMS Diffusion of Innovations Survey (CMS-DOIS). Perceptions of the relative advantage, compatibility, complexity, trialability, and observability were measured using statements asking respondents to rate them on a 5-point Likert scale where 1 = strongly disagree, 2 = disagree, 3 = undecided/neutral, 4 = agree, and 5 = strongly agree. To score each subscale, the mean of the Likert scale questions were taken (for number of questions per subscale, see Table 1). The survey questions that measured willingness to complete in-person and online CMS training were considered dependent variables, and were developed specifically for this study (Pereira, 2015). These were measured with two statements (one for online training and one for in-person training) using a 5-point Likert scale, where 1= not at all willing, 2 = somewhat unwilling, 3 = neither willing nor unwilling, 4 = somewhat willing, and 5 = very willing (see Table 1).

Table 1

Survey Question Origins, Subscale Definitions, Number of Items, Cronbach's Alpha, and Sample Questions

Instrument Name or Source	Subscale Name	Subscale Definition*	Number of Items	Cronbach's Alpha	Example Survey Question
CMS-DOIS	Relative advantage	The degree to which faculty members perceive that incorporating the use of their institution's CMS in teaching and learning is better than their current method.	15	0.939	Based on my experiences with the Blackboard CMS, I think using the Blackboard CMS enables (would enable) me to significantly improve the overall quality of my teaching.
CMS-DOIS	Compatibility	The degree to which faculty members perceive the CMS as being consistent with their existing values, past experiences, and current or future teaching needs.	10	0.821	Based on my experiences with the Blackboard CMS, I think using the Blackboard CMS fits (would fit) well with my teaching style.
CMS-DOIS	Complexity	The degree to which faculty members perceive the CMS as relatively difficult to understand and use.	10	0.916	Based on my experiences with the Blackboard CMS, I think learning to use the Blackboard CMS is (would be) easy for me.
CMS-DOIS	Trialability	The degree to which faculty members perceive that they may experiment with the CMS before they decide to incorporate it into their instruction.	7	0.767	Based on what I know right now, I think I was (am) permitted to use the Blackboard CMS on a trial basis long enough to see what it could/can do.
CMS-DOIS	Observability	The degree to which faculty members perceive the results of use of the CMS to be visible to others.	6	0.762	Based on what I know right now, I think I have observed how other teachers are using the Blackboard CMS in their teaching.
Pereira 2015	Willingness to complete online Blackboard training	At time of survey, over the next 12-month period, how willing faculty members were to complete any Blackboard CMS online training offered by FSU.	1	NA	Over the next 12-month period, how willing are you to complete any Blackboard CMS online training module(s) offered by FSU?
Pereira 2015	Willingness to complete in-person Blackboard training	At time of survey, over the next 12-month period, how willing faculty members were to complete any Blackboard CMS in-person training offered by FSU.	1	NA	Over the next 12-month period, how willing are you to complete any Blackboard CMS face-to-face training offered by FSU?

Note: * Based on Rogers' (2003) classifications of the five perceived attributes of an innovation.

Demographic information was also collected, as this information has been shown to have the potential to mediate the relationship between the dependent variables and independent variables (Allen & Seaman, 2012; Al-Senaidi et al., 2009; Ertmer & Ottenbreit-Leftwich, 2010, Hurtado et al, 2012; Keengwe et al., 2009; Keesee, 2010; Onyia & Onyia, 2011; Yidana et al., 2013). These mediating variables were gender, age, department, tenure status, rank, length of CMS use, and level of CMS expertise (see Table 2 for mediating variable definitions).

Table 2
Mediating Variable Definitions

Mediating Var	Measurement	Levels	Definition	Descriptive Analysis Classification
Gender	Categories	Male Female Other/refused	Gender at time of survey	Same as levels
Age	Categories	20 - 29 30 - 39 40 - 49 50 - 59 60 - 69 70 - 79 80+	Age at time of survey	Collapsed into the following groups due to low sample size: 20 - 39 40 - 49 50 - 59 60 - 69 70+
Department	Categories	Science, Technology, Engineering, and Math Social Science Education Economics, History, and Political Science Communications, Game Design All other departments	Primary department where faculty worked at time of survey	Collapsed into the following groups due to low sample size: Science Technology, Engineering, and Math Social Science, Economics, History, and Political Science Education, Communication, and Game Design Other
Tenure Status	Categories	Full-time tenured Full-time tenure-track Full-time nontenure-track Part-time	Faculty tenure status at time of survey	Collapsed into the following groups due to low sample size: Full-time tenured Full-time tenure-track Full-time and part-time nontenure-track
Rank	Categories	Instructor Assistant Professor Associate Professor Professor Other (please specify)	Faculty rank at time of survey	Collapsed into the following groups after analyzing "other" responses: Instructor Assistant Professor Associate Professor Professor
Length of CMS use	Years	0 - 30 (0 for less than 1 year Or if faculty did not use the CMS)	Number of years faculty had used the CMS at time of survey	Same as levels

Mediating Var	Measurement	Levels	Definition	Descriptive Analysis Classification
Level of CMS expertise	Likert Scale	1 = none 2 = little 3 = adequate 4 = more than adequate 5 = expert	Faculty level of expertise using CMS at time of survey	Same as levels

The survey was administered during a two-week period in October 2014. To ensure anonymity, the survey was accessed via a publicly available, universal link provided in e-mail to each faculty member. One week prior to survey administration, the university's chief information officer sent the faculty list an e-mail with details about the study. In addition, a reminder e-mail with the survey link was sent one week prior to the survey close date.

Data Analysis

After downloading the data from SurveyMonkey, subscales for the CMS-DOIS instrument were scored using SPSS (SPSS, n.d.). The subscales were found to be internally consistent through a Cronbach's alpha analysis (see Table 2). The values ranged from .762 to .939, which are considered reliable. The questions used to measure willingness to complete in-person and online training were found to have convergent validity with actual training participation (Pereira, 2015). Specifically, answers to questions on intention to complete online and in-person CMS training in the next 12 months were correlated with self-reports of training completion in the previous 12 months. The data indicated a trend that the more willing a person was to complete training, the more likely they were to have completed at least one training session over the previous 12 months.

Descriptive statistics were analyzed for the sample overall and separately for willingness to complete online versus in-person training. Means and distributions of continuous variables were considered, as were correlations.

To address the research questions, two separate analyses of variance (ANOVA) and linear regression models were developed, one to assess the association of the independent variables with the dependent variable "willingness to complete online training," and the other to assess the association of the independent variables with the dependent variable "willingness to complete in-person training." Mediating variables were entered as independent variables in the model to control for their potential influence on the dependent variable.

A best-subsets modeling procedure was followed (Hosmer, Borko, & Lemeshow, 1989; King, 2003). The best-subsets modeling approach is a method of selecting optimal predictor variables for a dependent variable, typically a binary one (Hosmer et al., 1989), but the procedure can be used in linear regression with a continuous dependent variable (King, 2003). The purpose of applying the best-subsets approach in this study was to use a data-driven rather than intuitive method of selecting an optimal set of predictor variables for the final model.

Results

Table 3 provides descriptive statistics for the faculty sample of n=102 with respect to categorical, demographic characteristics.

Table 3
Descriptive Statistics for Categorical Variables

Category	Levels	n (%)	*Online Willingness (M, SD)*	*In-person Willingness (M, SD)*
All	All	102, (100%)	3.52, (1.31)	3.46, (1.32)
Gender	Male	48, (47%)	3.27, (1.35)	3.42, (1.18)
	Female	46, (45%)	3.80, (1.22)	3.50, (1.46)
	Other	8, (8%)	3.38, (1.41)	3.50, (1.41)
Age Group	20-39 years	19, (19%)	3.58, (1.22)	3.16, (1.34)
	40-49 years	22, (22%)	3.41, (1.33)	3.41, (1.40)
	50-59 years	25, (25%)	3.64, (1.25)	3.52, (1.29)
	60 + years	21, (21%)	3.62, (1.40)	3.86, (1.2)
	Refused	15, (15%)	3.27, (1.49)	3.27, (1.39)
Tenure Status	Full-time Tenured	46, (45%)	3.22, (1.33)	3.39, (1.31)
	Full-time Tenure-track	24, (24%)	3.46, (1.32)	3.42, (1.38)
	Full time and Part-time Nontenure-track	32, (31%)	4.00, (1.16)	3.59, (1.32)
Rank	Instructor	27, (26%)	4.26, (0.94)	3.78, (1.37)
	Assistant Professor	24, (24%)	3.63, (1.35)	3.42, (1.38)
	Associate Professor	23, (23%)	3.00, (1.31)	3.13, (1.29)
	Professor	28, (27%)	3.14, (1.3)	3.46, (1.23)
Department	STEM	35, (34%)	3.31, (1.37)	3.40, (1.29)
	SEHP	16, (16%)	3.5, (0.97)	3.81, (1.05)
	ECG	15, (15%)	4.00, (1.31)	3.47, (1.41)
	Other	36, (35%)	3.53, (1.38)	3.36, (1.44)

Note: Online Willingness = willingness to complete online CMS training, In-person Willingness = willingness to complete in-person CMS training. STEM = Science, Technolgy, Engineering, and Mathmatics. SEHP = Social Science, Economics, History, and Political Science. ECG = Education, Communication, and Game Design. Other includes Business Administration, English Studies, Industrial Technology, Interdisciplinary Studies, and Nursing.

As indicated in Table 3, mean willingness to complete training fell mostly between 3 and 4, which is a small range. Females had a higher mean willingness to complete both online (3.80 vs. 3.27) and in-person training (3.50 vs. 3.42) than males. In most cases, older age levels were more willing to complete training, with the exception of the 40-49 level who were less likely to complete online training than the other levels (20-39 years = 3.58, 40-49 years = 3.41, 50-59 years = 3.64, and 60+ years = 3.62). Additionally, faculty members at earlier stages or not on the tenure track were more willing to complete training, especially in-person training (full-time tenured = 3.39, full-time tenure-track = 3.42, full-time and part-time non-tenure-track = 3.59). Likewise, in general, lower ranks expressed higher mean levels of willingness to complete training, excluding professors who were more willing to complete training than associate professors (online willingness: instructor = 4.26, assistant professor = 3.63, associate professor = 3.00, professor = 3.14; in-person willingness: instructor = 3.78, assistant professor = 3.42, associate professor = 3.13, professor = 3.46). Finally, members of the Education, Communication, and Game Design departments were much more willing to complete online training while members of the Social Science, Economics, History, and Political Science departments were much more willing to complete in-person training than the other departments (online willingness: Science, Technology, Math, and Science = 3.31, Social Science, Economics, History, and Political Science = 3.5, Education, Communications, and Game Design = 4.00, Other, including Business Administration, English Studies, Industrial Technology, Interdisciplinary Studies, and Nursing = 3.53; in-person willingness: Science, Technology, Mathematics, and Science = 3.40, Social Science, Economics, History, and Political Science = 3.81, Education, Communications, and Game Design = 3.47, Other, including Business Administration, English Studies, Industrial Technology, Interdisciplinary Studies, and Nursing = 3.36).

Table 4 provides summary statistics for the continuous variables.

Table 4
Summary Statistics for Continuous Variables

Variable	M*	SD
Online Willingness	3.52	1.31
In-person Willingness	3.46	1.32
Relative Advantage	3.58	0.77
Compatibility	3.66	0.73
Complexity	3.66	0.78
Trialability	3.36	0.70
Observability	3.48	0.72
Length use	6.16	4.22
Level expertise	3.26	1.04

Note: N=102. Online Willingness = willingness to complete online CMS training and In-person Willingness = willingness to complete in-person CMS training. *Length of use measured as discrete numerical variable ranging from 0 and 30 years, where 0 = less than 1 year or no use. Other variables measured on 5-point Likert scales.

As indicated in Table 4, the mean for the continuous variables that measured faculty perceptions of the CMS (independent variables) fell in a small range, between 3.36 and 3.66. Most of the independent variables had low to moderate positive correlations with each other, apart from relative advantage and compatibility, which were highly correlated with each other (see Table 5). The dependent variables (online and in-person willingness) were highly correlated with each other. Additionally, the dependent variables had low positive correlations with length of use and expertise level, excepting the correlation between in-person willingness and expertise level which was negative.

Table 5
Correlation Matrix

Variable	1	2	3	4	5	6	7	8	9
1. Online will	1.000	.709*	.443**	.432**	.241*	.173	.077	.041	.076
2. In-person will	.709**	1.000	.299**	.290**	.035	.088	.023	.088	-.058
4. Relative adv	.443**	.299**	1.000	.807**	.564**	.270**	.373**	.367**	.299**
5. Compatibility	.432**	.290**	.807**	1.000	.578**	.233**	.322**	.370**	.367**
6. Complexity	.241*	.035	.564**	.578**	1.000	.379**	.373**	.546**	.593**
7. Trialability	.173	.088	.270**	.233**	.379**	1.000	.527**	.169	.217*
8. Observability	.077	.023	.373**	.322**	.373**	.527**	1.000	.378**	.400**
9. Length	.041	.008	.367**	.370**	.546**	.169	.378**	1.000	.170**
10. Expert	.076	-.058	.299**	.367**	.593**	.217*	.400**	.170**	1.000

Note: N=102. Online willingness = willingness to complete online CMS training, In-person willingness = willingness to complete in-person CMS training. *$p < .05$. **$p < .01$.

Tables 6 and 7 present the ANOVA and linear regression results for the dependent variable "willingness to complete online training."

Table 6
Analysis of Variance for Predictors of Willingness to Complete Online Training

	Sum of Squares	df	Mean Square	F	Sig.
Regression	53.534	9	5.948	4.563	.000
Residual	119.927	92	1.304		
Total	173.461	101			

Note: Dependent variable measurement: Willingness to complete online CMS training.

Table 7
Predictors of Willingness to Complete Online Training

Predictor	*Beta (β)*	*t* statistic	*p*-value	VIF
Independent variables				
Compatibility	.490	5.451	.000	1.075
Mediating variables				
Department				
SEHP	.117	1.176	.243	1.319
CGE	.452	3.794	.000	1.889
Other	.221	2.079	.040	1.506
Age				
20-39 years	.066	.617	.539	1.527
40-49 years	-.010	-.087	.931	1.726
60+ years	-.033	-.307	.759	1.580
Refused	-.105	-.999	.321	1.461

Interaction variables				
CGE x 40-49 years	-.234	-2.104	.038	1.643

Note: Dependent variable is willingness to complete online training. SEHP = Social Science, Economics, History, and Political Science. ECG = Education, Communication, and Game Design. Other includes Business Administration, English Studies, Industrial Technology, Interdisciplinary Studies, and Nursing.

The ANOVA was statistically significant ($F = 4.563$ at 9 df., $p = 0.000$), so the model was interpreted. After modeling, only the perception of compatibility, defined as the degree to which faculty members perceive the CMS as consistent with their existing values, past experiences, and current or future teaching needs, was significantly positively associated with willingness to complete training online (standardized ß = 0.490, $p = 0.000$).

Tables 8 and 9 present the ANOVA and linear regression results for the dependent variable “willingness to complete in-person training.”

Table 8
Analysis of Variance for Predictors of Willingness to Complete In-person Training

	Sum of Squares	df	Mean Square	F	Sig.
Regression	49.284	15	3.286	2.242	.010
Residual	126.059	86	1.466		
Total	175.343	101			

Note: Dependent variable measurement: Willingness to complete in-person person training.

Table 9
Predictors of Willingness to Complete In-person Training

Predictor	*Beta (β)*	*t* statistic	*p*-value	VIF
Independent variables				
Compatibility	.242	2.469	.016	1.152
Mediating variables				
Tenure Status				
Full-time Tenure-track	.244	1.263	.210	4.458
Full-time and Part-time Nontenure-track	.125	.615	.540	4.937
Rank				
Instructor	.163	.925	.357	3.714
Assistant Professor	.264	1.363	.177	4.497
Associate Professor	-.030	-.251	.802	1.677
Gender				
Female	.367	2.414	.018	2.763
Other	.089	.702	.484	1.921
Age				
20-39 years	-.203	-1.613	.111	1.905

40-49 years	.117	.910	.366	1.983
60+ years	.149	1.269	.208	1.645
Refused	-.091	-.653	.516	2.315
Interaction variables				
Female x Full-time Tenure-track	-.425	-2.732	.008	2.893
Female x Full-time and Part-time Nontenure-Track	-.425	-2.455	.016	3.576
Assistant Professor x 40-49 age	-.384	-2.912	.005	2.082

Note: Dependent variable: Willingness to complete in-person training on the CMS.

For the in-person final model, the ANOVA was statistically significant ($F = 2.242$ at 15 df., $p = 0.010$), so the model was interpreted. This model demonstrates that of the independent variables, only compatibility was significantly positively associated with willingness to complete training in-person (standardized ß = 0.242, $p = 0.016$).

For this study, compatibility was defined as the degree to which faculty members perceive the CMS as consistent with their existing values, past experiences, and current or future teaching needs. Because in multivariate analysis, compatibility was the only independent variable statistically significantly associated with willingness to complete training on the CMS, both online and in-person, a bivariate analysis of the mean compatibility score for each mediating variable category was conducted. Table 10 provides a summary of this analysis.

Table 10
Descriptive Analysis of Compatibility Scores

Category	Levels	n (%)	*Compatibility Score (M, SD)*
All	All	102, (100%)	3.66 (.73)
Gender	Male	48, (47%)	3.71, (.66)
	Female	46, (45%)	3.67, (.75)
	Other	8, (8%)	3.38, (1.41)
Age Group	20-39 years	19, (19%)	3.68, (.72)
	40-49 years	22, (22%)	3.80, (.62)
	50-59 years	25, (25%)	3.74, (.70)
	60 + years	21, (21%)	3.56, (.78)
	Refused	15, (15%)	3.43, (.87)
Tenure Status	Full-time Tenured	46, (45%)	3.54, (.81)
	Full-time Tenure-track	24, (24%)	3.65, (.60)
	Full time and Part-time Non-tenure-track	32, (31%)	3.84, (.67)
Rank	Instructor	27, (26%)	3.95, (.65)
	Assistant Professor	24, (24%)	3.68, (.58)
	Associate Professor	23, (23%)	3.59, (.68)

	Professor	28, (27%)	3.42, (.87)
Department	STEM	35, (34%)	3.87, (.42)
	SEHP	16, (16%)	3.68, (.76)
	ECG	15, (15%)	3.41, (.67)
	Other	36, (35%)	3.56, (.73)

Note: STEM = Science, Technolgy, Engineering, and Mathmatics. SEHP = Social Science, Economics, History, and Political Science. ECG = Education, Communication, and Game Design. Other includes Business Administration, English Studies, Industrial Technology, Interdisciplinary Studies, and Nursing.

As shown in Table 10, most mediating variable levels were not considerably different with respect to mean compatibility scores. However, for tenure status there was a trend toward lower compatibility scores associated with higher tenure status (full-time tenured = 3.54, full-time tenure-track = 3.65, full-time and part-time non-tenure-track = 3.84). Similarly, for rank, mean compatibility scores decreased as ranks increased (instructor = 3.95, assistant professor = 3.68, associate professor = 3.59, professor = 3.42).

Table 9 also indicates that the mean compatibility score for males was slightly higher than for females (3.71 vs. 3.67), and the respondents who did not report their gender scored much lower than the two other groups (3.38). Generally, as age increased, mean compatibility scores decreased. Excepting those in the 20-39 year old range who scored less than the 40-49 and 50-59 age groups levels (20-39 years = 3.68, 40-49 years = 3.80, 50-59 years = 3.74, and 60+ years = 3.56). Like gender, participants who chose to not report their age scored the lowest in compatibility (3.43). Faculty members in the Science, Technology, Engineering, and Mathematics and Social Science, Economics, History, and Political Science departments reported the highest mean compatibility scores, followed by faculty members in the Other (included Business Administration, English Studies, Industrial Technology, Interdisciplinary Studies, and Nursing) group and the Education, Communication, and Game Design departments (Science, Technology, Engineering, and Mathematics = 3.87, Social Science, Economics, History, and Political Science = 3.68, Education, Communication, and Game Design = 3.41, Other = 3.56).

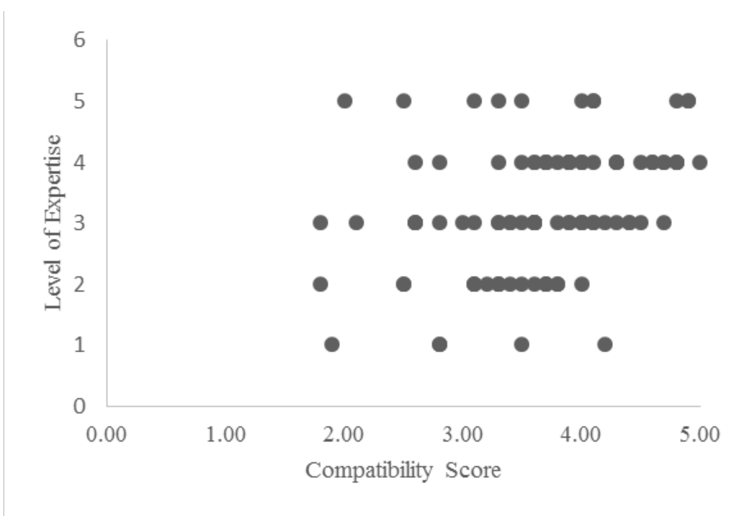


Figure 1. Scatterplot of compatibility score by level of expertise. Scale 1-5, where 1 = not at all willing, 2 = somewhat unwilling, 3 = neither willing nor unwilling, 4 = somewhat willing, and 5 = very willing.

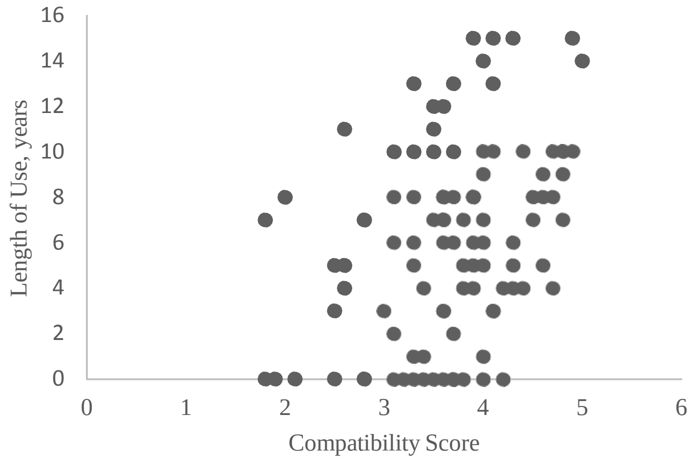


Figure 2. Scatterplot of compatibility score by length of use. 0 = less than 1 year or no use of CMS.

Self-rated level of expertise was also positively associated with compatibility score (r = .593, $p < 0.01$, see Figure 1). In addition, there was a moderate positive correlation between length of CMS use in years and compatibility perceptions (r = .546, $p < 0.01$, see Figure 2).

Discussion

Key Findings Summary

Of the independent variable measurements, only compatibility was significantly associated with willingness to complete training (both in-person and online). Consequently, the study results suggest that the other independent variables (relative advantage, complexity, trialability, and observability) did not significantly influence willingness to complete training on the CMS. Bivariate analyses indicated that higher tenure status was associated with lower perceptions of compatibility, and, similarly, that faculty perceptions of compatibility increased as their ranks decreased. Participants who rated themselves as having higher expertise levels also had the highest perceptions of compatibility. Lastly, in most cases, faculty members that used the CMS longer had higher perceptions of compatibility with the CMS than those that used the CMS for a shorter period of time.

Faculty members who perceived the CMS as compatible with their teaching styles were more willing to complete CMS training. For that reason, a strategy that can be used by universities is to identify and work with faculty members who view the CMS as incompatible with their teaching styles to help them discover methods to integrate the CMS into their classroom activities. This will likely not only increase willingness to engage in CMS training but will also promote CMS adoption.

Results of this study also revealed that faculty members who rated themselves as having a high level of expertise in using the CMS preferred online training. On the other hand, faculty members with lower expertise levels preferred in-person training. Faculty who reported lower expertise levels may benefit from assistance with integrating the CMS into their instruction as well as from technical help. Therefore, universities that design in-person CMS training for faculty members with low expertise levels and online training for faculty with higher expertise levels will likely increase faculty willingness to complete training on the CMS.

Connection to the Literature

In this study, compatibility was found to significantly positively influence faculty willingness to complete CMS training, both online and in-person. Although previous researchers did not specifically explore how perceptions of compatibility impact faculty willingness to complete training, they studied its influence on faculty willingness to adopt instructional technology. The findings of this study are generally consistent with the findings of other researchers. For instance, Tabata and Johnsrud (2008) found that if faculty members believe that distance education is compatible with their working styles then they are likelier to instruct distance education classes, and Sayadian et al.'s (2009) results suggested that if faculty members perceive that web-based instruction is consistent with their values and teaching methods then they are more disposed to integrating web-based instruction in their courses. Also, Tornatzky and Klein (1982), who studied general IT adoption, found that compatibility perceptions delivered one of the most constant, significant, positive associations within a large variety of innovation categories.

The findings of this research suggest that relative advantage did not significantly influence faculty willingness to complete CMS training, either positively or negatively. Although other studies did not specifically research how perceptions of relative advantage influence faculty willingness to complete training, researchers studied it in relation to faculty adoption and use of IT in the classroom. The findings of this study are inconsistent with prior research that found either positive or negative relationships. For example, Aremu et al. (2013) and Sayadian et al. (2009) found that relative advantage positively impacts faculty IT adoption, and Bennett and Bennet's (2003) research suggested that relative advantage positively influences faculty training program effectiveness. Conversely, Tabata and Johnsrud (2008) found an association between relative advantage and a decreased use of new technology practices. According to Tabata and Johnsrud, this may be because while faculty members perceive that distance education affords a relative advantage over current approaches, they do not believe that distance education instruction coincides with their responsibilities, needs, or values. Faculty members at FSU may have lacked a concept of courses without Blackboard. This is because most faculty members have adopted Blackboard for at least basic functions. Since Blackboard is widely adopted in at least some way, it may have been difficult for faculty members to gauge the "relative" advantage of not using it, given that it is rarely rejected completely at FSU in practice. This may explain why the results from this study differ from previous research.

This study found that perceptions of complexity do not significantly influence instructor willingness to complete CMS training. Previous research has not focused on how perceptions of the complexity of IT influences willingness to complete training, like this study. Rather, it has focused on how it affects faculty member willingness to use IT. The results of this research are consistent with studies conducted by Tabata and Johnsrud (2008) and Wang (2009). These studies found no significant relationship between faculty adoption of IT and complexity perceptions. Yet, these results contradict other study findings which suggested there is a significant inverse relationship between faculty IT adoption and complexity perceptions

((Bennett & Bennett, 2003; Keesee & Shepard, 2011; Motaghian, Hassanzadeh, & Moghadam, 2013; Prescott & Conger, 1995). Perhaps complexity has a strong impact only when faculty members perceive that the CMS is complex. Faculty members have used the Blackboard CMS at FSU for over 10 years, and Blackboard has been upgraded and improved over that time (Green, 2010). These advances may have lowered FSU faculty perceptions of its complexity to the extent that it was not much of an influence.

The findings of this study suggest that faculty perceptions of trialability do not significantly affect willingness to complete CMS training. Although previous studies did not focus on how perceptions of trialability influence faculty willingness to complete CMS training, they did focus on its influence of faculty willingness to use instructional technology. For example, Sayadian et al. (2009) suggested that faculty perceptions of trialability positively affects their incorporation of online instruction. Bennett and Bennett (2003) asserted that faculty members should be permitted to try IT in order to foster use, and Tabata and Johnsrud (2008) found that permitting faculty members to try instructional technology increased their adoption of it in distance education. It may be that trialability of the Blackboard CMS at FSU is not as critical to faculty because it has become much easier to edit courses in Blackboard. This is because, over time, Blackboard has become more functional (Blackboard, Inc., 2015).

This study found that faculty perceptions of observability do not significantly influence willingness to complete CMS training. This is in contradiction to previous findings. In particular, the results of three studies noted earlier (Bennett & Bennett, 2003; Sayadian et al., 2009; Tabata & Johnsrud, 2008) suggested that when faculty members thought their efforts would be observable, they would more likely to adopt IT. At FSU, although extensive adoption of all the functions of Blackboard is likely not occurring, at least some of its functions are being used in the majority of FSU classes. Therefore, the failure to use Blackboard altogether would become obvious to colleagues or students. Therefore, since observability is already consistently high, it may not influence willingness to complete CMS training. Although it may pressure faculty members to increase their Blackboard use, it does not directly result in increased willingness to complete training.

Findings Related to Theoretical Framework

This study was framed using components of the DOI theory. As conceptualized by Rogers (2003), the DOI theory suggests that perceived relative advantage, compatibility, trialability, and observability of an innovation positively influences its adoption rate, while perceived complexity negatively influences its adoption rate. Of the five attributes, only compatibility was related to faculty willingness to complete training on their institution's CMS, and this was a significantly positive relation for both online and in-person training. This coincides with Roger's (2003) theory.

Though perceptions of relative advantage, complexity, trialability, and observability may be influential in general for the adoption of technology as Rogers' (2003) postulated, there were no associations for this study's dependent variables (willingness of faculty to complete online and in-person training on the CMS) and for this technology (CMS). Perceptions of relative advantage may not have influenced faculty willingness to complete CMS training because Blackboard (in at least some capacity) is already used by most FSU faculty. Therefore, instructors may have had difficulty determining the "relative" situation of not using Blackboard. Regarding complexity, it is probable that FSU faculty did not perceive it to be relatively complex, given the high level of complexity of other current technology. Similarly, because CMSs, like Blackboard, permit faculty members to easily modify actions they take in the CMS, trialability may not be an important factor in their decisions to complete training. Similarly, in this study, observability did not influence willingness to complete CMS training, possibly because adoption of the CMS at FSU is already quite observable.

Study Limitations

This study is not without limitations. The results of this study are potentially generalizable to faculty members who teach at other state universities in the US. They are particularly generalizable to campuses that teach undergraduate and graduate students, have a faculty base similar to that of FSU, and have a CMS. However, the results may not be generalizable to other types of faculties and other settings. In addition, it is possible that the five perceived attributes associated with diffusion of innovation theory are not the most optimal attributes to explain willingness to complete training on a CMS in this population. Furthermore, the instrument used to measure the diffusion of innovation perceptions may not have been ideal. Furthermore, the best-subsets modeling approach may not have been the optimal choice for modeling, but sensitivity analysis showed that other approaches yielded similar results (Pereira, 2015), so the results are felt to be robust. Future research should consider other types of faculties in other settings, and measure other predictors felt to influence willingness to complete training on their CMS.

Conclusion and Recommendations

In conclusion, overall, the faculty in this study did not express high levels of willingness to complete CMS training, but CMS compatibility with teaching style was an influence. This study suggests that training for higher education faculty members on their institution's CMS should not be "one size fits all." Proper evaluation and categorization of teaching styles, as well as current utilization of the CMS are necessary before developing appropriate online and in-person training programs. This evaluation will help universities to better administer effective training that accommodates faculty members with different philosophies and pedagogical approaches to teaching as well as different perceived expertise levels. Further, it will foster enhanced and regular use of the institution's already-implemented CMS. More universal adoption by higher education faculty members of their institution's CMS will undoubtedly lead to an overall improvement in the quality of teaching and learning in higher education.

References

Abrahams, D. A. (2010). Technology adoption in higher education: A framework for identifying and prioritising issues and barriers to adoption of instructional technology. *Journal of Applied Research in Higher Education*, *2*(2), 34–49. http://doi.org/10.1108/17581184201000012

Allen, I. E., & Seaman, J. (2012). *Digital faculty: Professors, teaching and technology, 2012*. Babson Park, MA: Babson Survey Research Group. Retrieved from http://www.babson.edu/Academics/faculty/provost/Pages/babson-survey-research-group.aspx

Al-Senaidi, S., Lin, L., & Poirot, J. (2009). Barriers to adopting technology for teaching and learning in Oman. *Computers & Education*, *53*(3), 575–590. http://doi.org/10.1016/j.compedu.2009.03.015

Archambault, L., Wetzel, K., Foulger, T. S., & Williams, M. K. (2010). Professional development 2.0: Transforming teacher education pedagogy with 21st century tools. *Journal of Digital Learning in Teacher Education*, *27*(1), 4–11. http://doi.org/10.1080/21532974.2010.10784651

Aremu, A., Fakolujo, O., & Oluleye, A. (2013). Designing and Developing e-content in higher education: The University of Ibadan model. In *Unlocking the potential of ICT in higher education: Case studies of educational technology initiatives at African universities* (pp. 76–88). Johannesburg, South Africa: South African Institute for Distance Education. Retrieved from http://www.saide.org.za/

Bennett, J., & Bennett, L. (2003). A review of factors that influence the diffusion of innovation when structuring a faculty training program. *The Internet and Higher Education*, *6*(1), 53–63. http://doi.org/10.1016/S1096-7516(02)00161-6

Blackboard, Inc. (2017). Learn how Blackboard can help you grow in the evolving market of higher education. Retrieved from http://www.blackboard.com/higher-education/index.aspx

Bothma, C. H., & Cant, M. C. (2011). Adopting learning technologies: From belief to practice. *Educational Studies*, *37*(4), 375–389. http://doi.org/10.1080/03055698.2010.511697

Calderon, O., Ginsberg, A. P., Ciabocchi, L., Calderon, O., Ginsberg, A. P., & Ciabocchi, L. (2012). Multidimensional assessment of pilot blended learning programs: Maximizing program effectiveness based on student and faculty feedback. *Journal of Asynchronous Learning Networks*, *16*(4), 23–37. http://doi.org/http://sloanconsortium.org/publications/jaln_main

Carril, P. C. M., Sanmamed, M. G., & Sellés, N. H. (2013). Pedagogical roles and competencies of university teachers practicing in the e-learning environment. *The International Review of Research in Open and Distance Learning*, *14*(3), 462–487. http://doi.org/http://www.irrodl.org/index.php/irrodl

Chen, J. C., Ellis, M., Lockhart, J., Hamoush, S., Brawner, C. E., & Tront, J. G. (2000). Technology in engineering education: What do the faculty know and want? *Journal of Engineering Education*, *89*(3), 279–283. http://doi.org/10.1002/j.2168-9830.2000.tb00526.x

Chitiyo, R., & Harmon, S. W. (2009). An analysis of the integration of instructional technology in pre-service teacher education in Zimbabwe. *Educational Technology Research and Development*, *57*(6), 807–830. http://doi.org/10.1007/s11423-009-9136-7

deNoyelles, A., Cobb, C., & Lowe, D. (2012). Influence of reduced seat time on satisfaction and perception of course development goals: A case study in faculty development. *Journal of Asynchronous Learning Networks*, *16*(2), 85–98. http://doi.org/http://sloanconsortium.org/publications/jaln_main

Ertmer, P. A., & Ottenbreit-Leftwich, A. T. (2010). Teacher technology change: How knowledge, confidence, beliefs, and culture intersect. *Journal of Research on Technology in Education*, *42*(3), 255–284. http://doi.org/10.1080/15391523.2010.10782551

Goktas, Y., Yildirim, S., & Yildirim, Z. (2009). Main barriers and possible enablers of ICTs integration into pre-service teacher education programs. *Educational Technology & Society*, *12*(1), 193–204. http://doi.org/http://www.ifets.info/

Green, K. C. (2010). The 2010 campus computing survey. Retrieved from http://www.campuscomputing.net/survey

Hassan, S. S. (2011). The needs and perceptions of academics regarding their professional development in an era of educational transformation. *South African Journal of Higher Education*, *25*(3), 476–490. http://doi.org/http://jolt.merlot.org/

Hicks, S. D. (2011). Technology in today's classroom: Are you a tech-savvy teacher? *Clearing House: A Journal of Educational Strategies, Issues and Ideas*, *84*(5), 188–191. http://doi.org/http://journalseek.net/cgi-bin/journalseek/journalsearch.cgi?field=issn&query=0009-8655

Hixon, E., Buckenmeyer, J., Barczyk, C., Feldman, L., & Zamojski, H. (2012). Beyond the early adopters of online instruction: Motivating the reluctant majority. *The Internet and Higher Education*, *15*(2), 102–107. http://doi.org/10.1016/j.iheduc.2011.11.005

Hosmer, D. W., Borko, J., & Lemeshow, S. (1989). Best subsets logistic regression. *Biometrics*, *45*, 1265–1270. http://doi.org/10.2307/2531779

Hurtado, S., Eagan, K., Pryor, J. H., Whang, H., & Tran, S. (2012). *Undergraduate teaching faculty: The 2010-2011 HERI faculty survey*. University of California, Los Angeles: Higher Education Research Institute. Retrieved from http://www.heri.ucla.edu/

Iorio, J., Kee, K., & Decker, M. (2012). Instructional technology training: Developing functional and applied skill sets. In *Proceedings of the Society for Information Technology & Teacher Education International Conference* (pp. 1188–1193). Chesapeake, VA. Retrieved from http://www.editlib.org/p/39740/

Keengwe, J., Kidd, T., & Kyei-Blankson, L. (2009). Faculty and technology: Implications for faculty training and technology leadership. *Journal of Science Education and Technology*, *18*(1), 23–28. http://doi.org/10.1007/s10956-008-9126-2

Keesee, G. S. (2010). A perceived attributes and organizational support influencing course management system adopter status in historically black colleges and universities (Doctoral dissertation). Retrieved from ProQuest, UMI Dissertations Publishing. (Order No. 3426608).

Keesee, G. S., & Shepard, M. (2011). Perceived attributes predict course management system adopter status. *Online Journal of Distance Learning Administration*, *14*(1). Retrieved from http://www.westga.edu/~distance/ojdla/

Kidd, T. T. (2010). Butterfly under a pin: Exploring the voices and stories told of faculty who adopt ICT's for teaching and learning practices. *Education and Information Technologies*, *15*(3), 155–170. http://doi.org/10.1007/s10639-009-9102-3

King, J., E. (2003). Running a best-subsets logistic regression: An alternative to stepwise methods. *Educational and Psychological Measurement*, *63*(3), 392–403. http://doi.org/10.1177/0013164403063003003

Kinuthia, W. (2005). Planning faculty development for successful implementation of web-based instruction. *Campus-Wide Information Systems*, *22*(4), 189–200. http://doi.org/10.1108/10650740510617502

Korr, J., Derwin, E. B., Greene, K., & Sokoloff, W. (2012). Transitioning an adult-serving university to a blended learning model. *The Journal of Continuing Higher Education*, *60*(1), 2–11. http://doi.org/10.1080/07377363.2012.649123

Mallinson, B., & Krull, G. (2013). An investigation of the deployment of the Moodle virtual learning environment at eight African universities. In *Unlocking the potential of ICT in higher education: Case studies of educational technology initiatives at African universities* (pp. 30–45). Johannesburg, South Africa: South African Institute for Distance Education. Retrieved from http://www.saide.org.za/

Mark, K., Thadani, D. R., Santandreu Calonge, D., Pun, C. F. K., & Chiu, P. H. P. (2011). From neophyte to experienced facilitator: An interactive blended-learning course for graduate teaching assistants in Hong Kong. *Knowledge Management & E-Learning: An International Journal*, *3*(2), 153–169. http://doi.org/http://www.kmel-journal.org/ojs/index.php/online-publication

Masalela, R. K. (2009). Potential benefits and complexities of blended learning in higher education: The case of the University of Botswana. *Turkish Online Journal of Distance Education*, *10*(1), 66. http://doi.org/http://tojde.anadolu.edu.tr/

McBride, R., & Thompson, A. (2011). Using Moodle Academy to prepare technical college faculty for online teaching: An evaluation model. In *Proceedings of the Society for Information Technology & Teacher Education International Conference* (pp. 513–518). Chesapeake, VA. Retrieved from http://www.editlib.org/p/36320/

McKissic, S. C. (2012). Examining transformative faculty development factors to advance technology adoption and diffusion at a campus-based institution. (Doctoral dissertation). Retrieved from ProQuest, UMI Dissertations Publishing. (Order No. 3095210)

Meyer, K. A. (2014). An analysis of the cost and cost-effectiveness of faculty development for online teaching. *Journal of Asynchronous Learning Networks*, *18*(1). Retrieved from http://sloanconsortium.org/publications/jaln_main

Motaghian, H., Hassanzadeh, A., & Moghadam, D. K. (2013). Factors affecting university instructors' adoption of web-based learning systems: Case study of Iran. *Computers & Education*, *61*, 158–167. http://doi.org/10.1016/j.compedu.2012.09.016

Newhouse, R., Buckley, K. M., Grant, M., & Idzik, S. (2013). Reconceptualization of a doctoral EBP course from in-class to blended format: lessons learned from a successful transition. *Journal of Professional Nursing*, *29*(4), 225–232. http://doi.org/10.1016/j.profnurs.2012.05.019

Onyia, C. R., & Onyia, M. (2011). Faculty perception for technology integration in Nigeria university system: Implications for faculty quality curriculum design. *International Journal of Business and Social Science*, *2*(1), 81–92. http://doi.org/http://www.ijbssnet.com/

Pereira, A. S. (2015). Faculty willingness to complete information technology training on course management systems (Doctoral dissertation). Retrieved from Proquest, UMI Dissertation Publishing. (Order No. 3700990).

Porter, G. (2011). Specifics of course management system benefits for new university faculty. *Higher Education Studies*, *1*(2), 2–7. http://doi.org/10.5539/hes.v1n2p2

Potter, S. L., & Rockinson-Szapkiw, A. J. (2012). Technology integration for instructional improvement: The impact of professional development. *Performance Improvement*, *51*(2), 22–27. http://doi.org/10.1002/pfi.21246

Prescott, M. B., & Conger, S. A. (1995). Information technology innovations: A classification by IT locus of impact and research approach. *SIGMIS Database*, *26*(2-3), 20–41. http://doi.org/10.1145/217278.217284

Rocca, S. J. (2010). Determining the professional development needs of faculty in a college of agriculture. *College Student Journal*, *54*(1), 69–75. http://doi.org/http://www.projectinnovation.biz/csj.html

Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). New York, NY: Free Press.

Roman, T., Kelsey, K., & Lin, H. (2010). Enhancing online education through instructor skill development in higher education. *Online Journal of Distance Learning Administration*, *13*(4). Retrieved from http://www.westga.edu/~distance/ojdla/

Samarawickrema, G., & Stacey, E. (2007). Adopting web-based learning and teaching: A case study in higher education. *Distance Education*, *28*(3), 313–333. http://doi.org/10.1080/01587910701611344

Sandford, B. A., Dainty, J. D., Belcher, G. G., & Frisbee, R. L. (2011). Perceptions of the willingness of part-time instructors in community colleges in the US to engage in professional development opportunities and the best method(s) of delivering these experiences. *Journal of Career and Technical Education*, *26*(1), 48–61. http://doi.org/http://scholar.lib.vt.edu/ejournals/JCTE/

Sayadian, S., Mukundan, J., & Baki, R. (2009). Exploring the factors influencing UPM English language faculty members' adoption and integration of web-based instruction (WBI). *Journal of College Teaching & Learning*, *6*(6), 31–38. http://doi.org/http://journals.cluteonline.com/index.php/TLC

Simon, D., Jackson, K., & Maxwell, K. (2013). Traditional vs. online instruction: Faculty resources impact strategies for course delivery. *Business Education & Accreditation*, *5*(1), 107–116. http://doi.org/http://www.theibfr.com/bea.htm

Smolin, L., & Lawless, K. A. (2011). Evaluation across contexts: Evaluating the impact of technology integration professional development partnerships. *Journal of Digital Learning in Teacher Education*, *27*(3), 92–98. http://doi.org/10.1080/21532974.2011.10784663

Tabata, L. N., & Johnsrud, L. K. (2008). The impact of faculty attitudes toward technology, distance education, and innovation. *Research in Higher Education*, *49*(7), 625–646. http://doi.org/10.1007/s11162-008-9094-7

Tornatzky, L. G., & Klein, K. J. (1982). Innovation characteristics and innovation adoption-implementation: A meta-analysis of findings. *IEEE Transactions on Engineering Management*, *EM-29*(1), 28–45. http://doi.org/10.1109/TEM.1982.6447463

Tsai, Y., & Talley, P. C. (2013). The effect of a course management system (CMS)-supported strategy instruction on EFL reading comprehension and strategy use. *Computer Assisted Language Learning*, 1–17. http://doi.org/10.1080/09588221.2012.757754

Unal, Z., & Unal, A. (2011). Evaluating and comparing the usability of web-based course management systems. *Journal of Information Technology Education: Research*, *10*(1), 19–38. http://doi.org/http://www.informingscience.us/icarus/journals/jiteresearch

Unwin, T., Kleessen, B., Hollow, D., Williams, J. B., Oloo, L. M., Alwala, J., … Muianga, X. (2010). Digital learning management systems in Africa: Myths and realities. *Open Learning*, *25*(1), 5–23. http://doi.org/10.1080/02680510903482033

Wang, W.-T., & Wang, C.-C. (2009). An empirical study of instructor adoption of web-based learning systems. *Computers & Education*, *53*(3), 761–774. http://doi.org/10.1016/j.compedu.2009.02.021

West, R., Waddoups, G., & Graham, C. (2007). Understanding the experiences of instructors as they adopt a course management system. *Educational Technology Research & Development*, *55*(1), 1–26. http://doi.org/10.1007/s11423-006-9018-1

Yidana, I., Sarfo, F. K., Edwards, A. K., Boison, R., & Wilson, O. A. (2013). Using the Moodle learning management system for teaching and learning at the University of Education, Winneba. In *Unlocking the potential of ICT in higher education: Case studies of educational technology initiatives at African universities* (pp. 58–75). Johannesburg, South Africa: South African Institute for Distance Education. Retrieved from http://www.saide.org.za/

Yohon, T., & Zimmerman, D. (2006). An exploratory study of adoption of software and hardware by faculty in the liberal arts and sciences. *Journal of Technical Writing and Communication*, *36*(1), 9–27. http://doi.org/10.2190/8QGV-RYXQ-VY7Q-5RJT

Understanding Teachers' Cognitive Processes during Online Professional Learning: A Methodological Comparison

Pamela Beach
Queen's University

Dale Willows
Ontario Institute for Studies in Education, University of Toronto

Abstract

This study examined the effectiveness of three types of think aloud methods for understanding elementary teachers' cognitive processes as they used a professional development website. A methodology combining a retrospective think aloud procedure with screen capture technology (referred to as the virtual revisit) was compared with concurrent and retrospective think aloud procedures. Elementary teachers from a large metropolitan area were assigned to one of the three think aloud conditions (N = 45). Participants in the concurrent condition verbalized their thoughts while simultaneously navigating a professional development website for 20 minutes. Participants in the retrospective condition verbalized their thoughts following their 20-minute website navigation without any aids. Finally, participants in the virtual revisit condition verbalized their thoughts while viewing a screen recording of their website navigation. Think aloud protocols were analyzed to determine the frequency of cognitive processes verbalized by participants in each condition. The findings of this study indicated significant differences in the types of verbalizations produced by participants across the three think aloud conditions. In addition, findings reveal benefits and limitations of employing each type of think aloud method in the context of a professional development website.

Keywords: online learning, teacher cognition, think aloud methodology, teacher professional development

Beach, P, & Willows, D. (2017). Understanding teachers' cognitive processes during online professional learning: A methodological comparison, *Online Learning 21*(1), 60-84. doi: 10.24059/olj.v21i1.949

Introduction

Elementary teachers are a necessary foundation for building successful programs in the classroom (Gambrell & Anders Mazzoni, 1999; Pressley, Mohan, Raphael, & Fingeret, 2007). Successful programs begin with a repertoire of pedagogical knowledge, pedagogical content knowledge, and research-based instructional practices. This repertoire of information can be delivered to practicing elementary teachers through various professional development opportunities (Cervetti, Kulikowich, Drummond, & Billman, 2012; Desimone, 2009; Kao, Wu, & Tsai, 2011).

One facet of teacher professional development is online learning, which occurs when professional knowledge is constructed from multiple modes of digital information—photographs, videos, and interactive tools, to name a few (Mayer, 2002). Online learning is a favored approach to professional development because it creates accessible opportunities; online learning takes place within platforms that deliver information in a means that removes time, place, and situational barriers (Kanuka & Nocente, 2003). Online learning opportunities have also been shown to have positive effects on and even change teachers' pedagogical and content knowledge, classroom practice, and student outcomes (Weschke & Barclay, 2011). As elementary teachers increasingly turn to the Internet for their professional learning (Charalambousa & Ioannou, 2011; Kao et al., 2011), it is essential to examine how they use and learn from online resources and professional development websites.

Most studies that have examined online teacher learning have gathered data through surveys, questionnaires, and interviews (Duncan-Howell, 2010; Hur & Brush, 2009; Kao et al., 2011). These methods offer information about teachers' attitudes towards online professional learning; however, data generated from these methods is limited to participants' recollection of past events. A method that tracks teachers' cognitive processes as they make online choices is necessary to provide further insight into how teachers use and learn from online environments. The think aloud methodology is an approach that can track teachers' cognitive processes during decision-making activities. While think aloud research is extensive, studies that compare the effectiveness of different think aloud methodologies for understanding teachers' cognitive processes as they navigate online resources are limited (Kuusela & Paul, 2000; van Gog, Paas, van Marrienboer, & Witte, 2005). The purpose of this comparative study was to examine the effectiveness of three types of think aloud methods for understanding elementary teachers' cognitive processes as they used a professional development website. A methodology combining a retrospective think aloud procedure with screen capture technology (referred to as the virtual revisit) was compared with concurrent and retrospective think aloud procedures.

A detailed discussion of the think aloud methodology sets the foundation for this paper. The current study's methods are then presented, followed by the results and a discussion of the significant findings. The study's limitations and educational implications conclude this paper.

Literature Review

Thinking aloud has historical roots in introspection analysis, a form of data collection aimed at investigating psychological claims and theories of mind during the eighteenth century

(Boren & Ramey, 2000; Ericsson, 2002). The cognitive revolution of the 50s and 60s produced alternative types of verbal reports of thinking to gather information about cognitive structures and processes (Ericsson, 2003). Today, the think aloud method most widely employed is based on the techniques of protocol analysis by Ericsson and Simon (1984, 1993). As described by Ericsson and Simon (1984), thinking aloud captures cognitive processes in real time and verbal reports "provide the most informative data available on thinking during cognitive tasks" (Ericsson, 2003). Cognitive processes underlying decisions and behaviors are usually "hidden from direct observation" (Gaissmaier, Fifc, & Rieskany, 2010, p. 141). However, the think aloud method makes monitoring cognitive processes possible—the think aloud generates direct data about the ongoing cognitive processes that occur during task performance (Jaspers, Steen, van den Bos, & Geenen, 2004).

Ericsson and Simon (1984) describe three levels of verbalizations that can occur during the think aloud method. The first two levels require information processing in the participant's short term memory and the third level requires additional cognitive resources and retrieval of information from long term memory (Olmsted-Hawala, Murphy, & Hawala, 2010). While Ericsson and Simon (1984) state that Level 3 verbalizations or higher cognitive processes are less reliable because they involve access to long-term memory, usability researchers suggest that this type of data provides useful information about online learning, website user goals, and online behaviors (Boren & Ramey, 2000; Guan, Lee, Cuddihy, & Ramey, 2006; Olmsted-Hawala et al., 2010).

Usability researchers most often employ the concurrent and retrospective think aloud methods to gain insight into web seeking behaviors and to evaluate a website's content and ease of use (Aranyi, Schaik, & Barker, 2012; Barzilai & Zohar, 2012; Branch, 2006; Kuusela & Paul, 2000). During the concurrent procedure participants verbalize their thoughts aloud while they simultaneously complete a task. Verbal reports that result from the concurrent procedure generate data about the website user's navigational experience. For instance, Aranyi and his colleagues (2012) conducted an exploratory study of interaction experience with a news website. The concurrent think aloud yielded five categories of experience based on the participants' evaluative statements: impression, content, layout, information, architecture, and diversion (Aranyi et al., 2012). Similarly, Barzilai and Zohar (2012) utilized the concurrent procedure to examine epistemic thinking in action. Data was collected to shed light on the relationship between sixth grade students' knowledge construction and their online practices. Analysis revealed a positive relation between students' online strategies and their epistemic cognition (Barzilai & Zohar, 2012). Damico and Balidon (2007) also employed the concurrent procedure to examine how elementary students engage with an educational website. Findings from their study highlight how elementary students evaluated claims and evidence of online educational resources (Damico & Balidon, 2007).

The retrospective procedure is also referred to as post-task testing, retrospective report, and think after. Retrospective think alouds alone are used less often in the fields of online learning and website usability since they require participants to think aloud after a task has been completed. An international survey found that just 5% of think aloud studies (not limited to website usability) employed the retrospective technique, whereas 89% used the concurrent think aloud, and 6% used an alternative think aloud (McDonald, Edwards, & Zhao, 2012).

One of the main reasons why retrospective think alouds are used less often is due to the fact that the procedure relies on the ability to recall decisions after a task has been completed. As participants recall their decisions, information may be incomplete and include errors, omissions, and substitutions (Branch, 2006). For instance, a comparative study that examined retrospective and concurrent verbal protocol analysis in the context of a decision-making task found retrospective reports more prone to errors of omission whereas concurrent reports contained more relevant information about the decision making process (Kuusela & Paul, 2000).

While retrospective procedures are limited by the fact that they may be incomplete and include errors, omissions, and substitutions (Branch, 2006), they have the advantage of freeing cognitive resources by thinking aloud after the task has been completed—retrospective think alouds do not interfere with task performance (McDonald et al., 2012). Concurrent think alouds, on the other hand, can interfere with task performance since participants verbalize their thoughts while they simultaneously complete a given task—participants engage in two different processes at the same time. When two processes occur simultaneously there is an increase in cognitive load—"the level of mental energy required to process a given amount of information" (Ping Lim, 2004, p. 17). As a result of a higher cognitive load, task completion may be compromised during the concurrent procedure and resulting think aloud reports are often procedural in nature (McDonald et al., 2012; van Gog, Kester, Nievelstein, Giesbers, & Paas, 2009). Findings of such studies suggest that alternative think aloud methods warrant attention.

An alternative to the concurrent and retrospective procedures is the virtual revisit think aloud method. The virtual revisit is a variation of the retrospective think aloud method and allows participants to review and comment on a visual recording of how they interacted with a particular website. The goal of the virtual revisit is to aid recall of original events and thought processes by using a screen-capture recording of participants' navigational experiences. Similar to cued retrospective reporting where participants are given instructions to report retrospectively on the basis of a record of observations (van Gog, Paas, van Marrienboer, & Witte, 2005), the virtual revisit think aloud combines a retrospective think aloud with screen capture technology to aid recall of original events and thought processes.

Despite the limitations of the concurrent procedure, it has been widely used in usability research, mostly as a means to evaluate a given website—participants verbalize their thoughts about the ease of use and accessibility of information. While evaluative data contributes to the refinement of professional development websites, user experience is a complex process and usability research should go beyond evaluating websites to include a range of cognitive processes and learning strategies (Dillon, 2001); the virtual revisit think aloud has the potential to avoid the limitations of the concurrent and retrospective procedures. In addition, few studies have been undertaken to compare the relative utility of different think aloud procedures (Kuusela & Paul, 2000; van Gog, Paas, van Marrienboer, & Witte, 2005). The current study addresses these gaps in the literature by examining the utility of three think aloud methods during online professional learning.

The following research questions guided the study:

1. To what extent do participants' verbalizations differ across the three think aloud methods?

2. What are the benefits and limitations of employing each type of think aloud in the context of online learning?

Methods

Study Context

This research was undertaken within the context of the development and refinement of a professional development website. The website is a multimedia evidence-informed literacy professional development website that provides free professional learning resources for elementary teachers and educators. The website is highly complex and interactive, and includes virtual tours of exemplary classrooms (PreK-6), video clips of expert teachers explaining and demonstrating effective educational practices, detailed lesson plans, photographs of teaching materials, exemplars of student work, and links to related research articles.

Participants

Forty-five practicing elementary teachers from a large metropolitan area participated in this research over an eight-month period. All participants completed informed consent forms and volunteered to participate in this study.

Data Sources

Demographic questionnaire. A demographic questionnaire was administered to participants to obtain data on a range of relevant factors based on the literature related to teacher development and online learning (e.g., age, gender, years of teaching experience, education, extent of involvement with various web-based technologies). Participants were asked to complete the questionnaire online prior to a one-on-one meeting. The questionnaire was administered through an online survey program. After participants completed the questionnaire, they were placed in one of three think aloud conditions (concurrent, retrospective, and virtual revisit). The conditions are described below. Stratified random assignment was employed to ensure that certain demographic features were represented within each group as equally as possible. Stratification variables were selected based on the literature on web navigation and teacher professional development. Research has found that gender (Page, Robson, & Uncles, 2012; J. Pearson, A. Pearson, & Green, 2007), age (Laberge & Scialfa, 2005), subject matter knowledge and experience (Laberge & Scialfa, 2005), and computer self-efficacy (Page et al., 2012) influence the perception and navigation of websites; therefore, comparable aspects including gender, age, years of teaching experience, current teaching grade, comfort with technology, and frequency of Internet use for professional purposes, were selected as the key variables used for the group assignment.

Table 1 (next page) summarizes the demographic characteristics for the participants across the conditions. As can be seen from the table, participants in the three groups were quite similar with respect to gender, age, years of teaching, current grade, comfort with technology, and frequency of Internet use. In other words, the stratified random assignment was successful.

Table 1
Demographic Characteristics

Characteristic		Concurrent Condition (*n* = 15)	Retrospective Condition (*n* = 15)	Virtual Revisit Condition (*n* = 15)	Total (*N* = 45)
		n (%)	*n* (%)	*n* (%)	*n* (%)
Gender					
	Male	3 (20%)	3 (20%)	3 (20%)	9 (20%)
	Female	12 (80%)	12 (80%)	12 (80%)	36 (80%)
Age					
	25-29	3 (20%)	3 (20%)	3 (20%)	9 (20%)
	30-34	3 (20%)	5 (33%)	4 (27%)	12 (27%)
	35-39	4 (27%)	3 (20%)	4 (27%)	11 (24%)
	40-44	2 (13%)	2 (13%)	2 (13%)	6 (13%)
	45-49	1 (6%)	1 (6%)	1 (6%)	3 (7%)
	50-54	2 (13%)	0	0	2 (4%)
	55+	0	1 (6%)	1 (6%)	2 (4%)
Years Teaching					
	2-4 years	3 (20%)	3 (20%)	4 (27%)	10 (22%)
	5-9 years	6 (40%)	6 (40%)	5 (33%)	17 (38%)
	10-14 years	3 (20%)	5 (33%)	4 (27%)	12 (27%)
	15-19 years	1 (6%)	0	1 (6%)	2 (4%)
	20-24 years	1 (6%)	0	1 (6%)	2 (4%)
	25+ years	1 (6%)	1 (6%)	0	2 (4%)
Current Grade					
	Kindergarten	1 (6%)	1 (6%)	0	2 (4%)
	Primary	5 (33%)	5 (33%)	7 (47%)	17 (38%)
	Junior	7 (47%)	6 (40%)	6 (40%)	19 (42%)
	Primary/Junior	2 (13%)	3 (20%)	2 (13%)	7 (16%)
Comfort Using Internet					
	Very	15 (100%)	13 (87%)	13 (87%)	41 (91%)
	Somewhat	0	2 (13%)	2 (13%)	4 (9%)
	Not very	0	0	0	0
Frequency of Internet Use for Professional Purposes					
	> once/day	5 (33%)	3 (20%)	6 (40%)	14 (31%)

Once/day	6 (40%)	8 (53%)	7 (47%)	21 (47%)
Once/week	4 (27%)	4 (27%)	1 (6%)	9 (20%)
Once/month	0	0	1 (6%)	1 (2%)

Think aloud. Participants completed a think aloud during a one-on-one meeting.

Concurrent think aloud. Participants in the concurrent condition verbalized their thoughts for 20 minutes while simultaneously completing a website task.

Retrospective think aloud. Immediately following a 20-minute website task, participants in the retrospective condition recalled and verbalized their thought processes without any aids.

Virtual revisit think aloud. Immediately following a 20-minute website task, participants in the virtual revisit condition reviewed their online choices virtually and verbalized their thoughts while viewing the 20-minute screen recording of their explorations.

Screen capture technology. During participants' navigation of the website, each visual step was captured with Camtasia Studio, a screen-recording computer software program developed by TechSmith (Uppal, 2011).

Procedure

One-on-One Meeting. The one-on-one meetings followed a sequence of events and lasted approximately 45 minutes.

Website task and think aloud. The following website task instructions were presented to all participants:

> Your task is to use a professional development website as you normally would when seeking information online for your teaching practices.

While the website task instructions were consistent across the conditions, the think aloud instructions varied for each condition. Participants in the concurrent condition were given the think aloud instructions before they completed the website task, whereas participants in the retrospective and virtual revisit groups were given the think aloud instructions after they completed the website task. The purpose of informing participants in the retrospective and virtual revisit conditions of the think aloud instructions after their navigation was to reduce reactivity—"influences of the verbalizations on the decision process" (Ranyard & Svenson, 2010, p. 119)—as much as possible. The following passage outlines the think aloud instructions. The underlined portions state the different think aloud instructions given for each condition.

> In this study we are interested in what you think about when you explore a professional development website. In order to do this, I am going to ask you to think aloud (*concurrent condition*: as you explore the website; *retrospective condition:* about your exploration of the website; *virtual revisit condition:* while you view a recording of your exploration of the website). What I mean by think aloud is that I want you to tell me everything that you are/were thinking from the time you begin/began exploring the

website until the end of your exploration. I would like you to talk aloud constantly. I don't want you to try to plan out what you say or try to explain to me what you are saying. Just act as if you are alone in the room speaking to yourself. It is most important that you keep talking.

While think aloud studies most often employ a specific task, an open-ended task was used in this study to reflect as naturally as possible how teachers use and learn from professional development websites. To reduce disruption to the participants' cognitive processes, prompts, redirections, and interventions were kept to a minimum during the process of verbalizing (Jaspers, 2009). Participants who were silent for a period of 30 seconds were only told to "keep talking." This prompt was only given to one participant in the concurrent condition.

Data Analysis

Audio recordings were transcribed verbatim resulting in 45 think aloud transcripts. Word counts were calculated for each of the three conditions (concurrent, retrospective, and virtual revisit). As shown in Table 2 there are clear differences in the average number of words participants generated in the three think aloud conditions.

Table 2
Average Word Counts across the Three Conditions

	Concurrent Condition	Retrospective Condition	Virtual Revisit Condition
	M (*SD*)	*M* (*SD*)	*M* (*SD*)
Word Count	1676.13 (394.89)	658.80 (229.00)	2637.87 (359.94)

The coding scheme used to code all 45 transcripts was generated based on several studies relating to website usability (Aranyi et al., 2012; Cooke, 2010; Damico & Balidon, 2007; van Gog et al., 2005; Tan & Wei, 2006; Zhao & McDonald, 2010), teacher planning and decision making (Kansanen et al., 2000; Moos, 2014), and Ericsson and Simon's levels of verbalizations and suggested statements (1984). The final coding scheme used to code all 45 transcripts includes 11 categories. Table 3 summarizes the coding scheme and offers a description of each category and an example from the think aloud transcripts.

Table 3
Coding Scheme-Categories Used to Code the Transcripts

Verbalization Category	Description	Examples from the Transcripts
Planning	Referring to program planning, reorganizing information to form or develop new ideas; constructing and creating	"…that is a lesson I can just take and tweak for my class right away."

Connecting	Recalling information; activating prior knowledge in relation to information presented on the site; finding a past example or recalling a concept	"This reminds me of the mini lessons that I like to do at the beginning of the year."
Reasoning	Providing a rationale for making a navigational decision; explaining why	"I was considering looking at text structures and I ended up choosing reading comprehension strategies because I'm trying to make that one of the main focuses of our reading program."
Reflecting	Making meta-comments in reference to awareness of their own thinking and learning style	"I like to go over everything first and then go back and look at something more in-depth."
Evaluating Website Content	Making judgments or expressing opinions about an aspect of the website or information presented on the website	"This kind of photo tour really informs me in terms of good practice for classroom management and good classroom environments."
Evaluating User Experience	Expressing a positive or negative feeling towards the usability and accessibility of the website and its features	"It's nice that the link is already there for me, that I don't have to type in a separate search button or go onto Google, I can just click onto the link."
Diversion	Verbalizing difficulties, including utterances where participants indicate uncertainty and confusion	"Hmm, assessment, how do I go, so what do I do? Enter? Search?" "How do I go back? This is…am I doing something wrong?"
Understanding	Identifying and making sense of new information and web-based tools	"At this point I'm just looking at the main page and I'm understanding how the website is organized."
Describing Procedural Behavior	Describing what they are doing or going to do or just did; statements about participants' actions during their navigation	"I'm just looking at the videos right now."
Describing Website Features	Describing the spatial characteristics, website features and images; what participants notice	"It's showing various pictures and monthly virtual tours."

Reading	Reading words, phrases or sentences off the screen; reading out texts, headings, links and other on-screen text	"Motivation for literacy, oral language, knowledge building, concepts of print, writing conventions."

The categories listed in Table 3 are types of cognitive processes. Cognitive processes underlie the study of decision-making during professional learning experiences and can be categorized as higher order processes and lower order processes. Higher order processes, such as reasoning, involve access to "thematically related information in long-term memory" (Horz & Schnotz, 2010, p. 238). Information stored in memory "is interrelated and rearranged, and extended to achieve a specific purpose" (Lewis & Smith, 1993, p. 136). Lower order processes, such as procedural knowledge, are normally executed in an automated way and are "only marginally influenced by intentional processes" (Horz & Schnotz, 2010, p. 238). In the current study, higher order cognitive processes refer to planning, connecting, reasoning, reflecting, and evaluating, whereas lower order cognitive processes refer to diversion, understanding, describing, and reading. The higher order and more complex cognitive processes "involve the manipulation of information" (McLoughlin & Mynard, 2009, p. 148), whereas the lower order cognitive processes demand only "mechanical application of previously acquired information" (Lewis & Smith, 1993, p.133).

The organization of the categories into higher and lower order cognitive processes is consistent with Krathwohl's revised taxonomy of educational objectives (2002). Krathwohl (2002) organizes six major categories and 19 sub-categories of the cognitive domain hierarchically and discusses how they differ in complexity. Krathwohl (2002) distinguishes between higher and lower order cognitive processes: the more complex categories (e.g., create) are higher on the scale, whereas the less complex categories (e.g., recalling) are lower on the scale. This distinction was taken into consideration during the development of the coding scheme for the current study.

Prior to coding, the think aloud transcripts were first segmented or "unitized" into thought units—each utterance was deemed a separate segment or thought unit if it conveyed relevant information and was preceded and followed by a pause and a change of ideas (Lincoln & Guba, 1985). According to Ericsson and Simon (1993) this procedure, in which protocols are unitized into phrases or segments provides more reliable findings. A second researcher was trained on dividing the think aloud transcripts into segments or thought units to establish inter-rater agreement. The second researcher was not involved in the research project and had no specific interest in the outcomes (van Someren, Barnard & Sandberg, 1994). This was necessary to provide results that were as objective as possible (van Someren et al., 1994). Training involved a review of unitizing, a demonstration of segmenting the transcripts into thought units, and a practice trial of segmenting a portion of one of the transcripts into thought units. As described by van Someren et al. (1994), during think aloud protocol analysis "coders need to be trained in the use of the coding scheme" and the context should be considered when interpreting individual phrases (p. 128).

Following the training session, two researchers segmented 10% of the total transcripts into thought units. The percentage of agreements was calculated (agreements/agreements +

disagreements). The unitizing reliability check on 10% of 45 transcripts indicated high reliability, with an inter-rater agreement of 95.3% (Guetzkow's U = .012; Guetzkow, 1950). Due to the high inter-rater agreement and to the submission that unitizing involves subjective interpretation and contextualization (Lomard, Snyder-Duch, & Campanella Bracken, 2004), the remaining transcripts were unitized by the primary researcher, who had a thorough understanding of the research topic.

Once all 45 transcripts were unitized, the total thought units were calculated. Similar to the differences in word counts, there are clear differences in the average number of thought units generated by participants in each think aloud condition (see Table 4).

Table 4
Average Number of Thought Units across the Three Conditions

	Concurrent Condition	Retrospective Condition	Virtual Revisit Condition
	M (*SD*)	*M* (*SD*)	*M* (*SD*)
Number of Thought Units	147.40 (35.83)	30.73 (13.15)	123.27 (20.58)

Similar to the unitizing procedure, a second coder who was not involved in the research project and had no specific interest in the outcomes was trained on the coding scheme to establish inter-rater agreement. Following the training session, 10% of the transcripts were coded by the primary researcher and second coder. Disagreements were discussed and resolved until an inter-rater agreement of 97% was reached. Cohen’s Kappa was performed to determine consistency among raters and was found to be 0.98, CI (0.978-0.996). Due to the high inter-rater agreement and the assertion that researchers themselves may serve as coders (Lombard, Snyder-Duch & Campanella Bracken, 2002), the remaining transcripts were coded by the primary researcher.

Thought units were tallied to provide frequency counts for each category. These frequency counts were then transformed into percentages based on the total number of thought units across categories. Coding resulted in a total of 4,521 thought units. The frequencies of thought unit were converted to percentages for each participant. The percentage of thought units in each condition were then analyzed quantitatively. Analyzing the percentage of thought units, as opposed to the frequency of thought units, produces a more accurate representation of the cognitive processes (Rosenzweig, Krawec, & Montague, 2011). This allows for a more accurate comparison of thought units across conditions. The word counts and total number of thought units were also analyzed quantitatively. Descriptive statistics, ANOVA, and Pearson correlations were performed on the proportion of thought units, total word counts, and total number of thought units across the three conditions (concurrent, retrospective, and virtual revisit). Table 5 displays the distribution of the frequency and percentage of thought units across the three conditions.

Table 5
Frequencies and Percentages of Thought Units

Category	Concurrent Condition		Retrospective Condition		Virtual Revisit Condition		Total	
	Total	%	Total	%	Total	%	Total	%
Planning	9	0.41	20	3.96	73	4.17	102	2.26
Connecting	127	6.04	57	13.23	170	9.08	354	7.83
Reasoning	47	2.23	47	10.99	359	19.27	453	10.02
Reflecting	354	16.71	121	26.86	464	25.08	939	20.77
Evaluating Website Content	229	10.87	66	12.18	285	15.36	580	12.83
Evaluating User Experience	166	7.96	59	12.13	138	7.47	363	8.03
Diversion	136	5.97	3	0.79	56	2.99	195	4.31
Understanding	141	5.99	14	2.97	60	3.36	215	4.76
Describing Procedural Behavior	265	11.64	57	11.71	173	9.37	495	10.95
Describing Website Features	235	10.15	17	3.49	66	3.56	318	7.03
Reading	502	22.03	0.00	0.00	5	0.29	507	11.21

Results

A one-way between subjects ANOVA was performed to compare the percentage of each category verbalized by participants in each think aloud condition. The analysis revealed a significant main effect for planning, $F(2, 42) = 7.05$, $p = .002$, $\eta^2 = .251$; connecting, $F(2, 42) = 8.35$, $p = .002$, $\eta^2 = .251$; reasoning, $F(2, 42) = 22.01$, $p < .001$, $\eta^2 = .512$; and reflecting, $F(2, 42) = 5.36$, $p = .008$, $\eta^2 = .203$. ANOVA results also indicated significant main effects for describing website features, $F(2, 42) = 11.05$, $p < .001$, $\eta^2 = .345$; and for reading, $F(2, 42) = 43.81$, $p < .001$, $\eta^2 = .676$. There were no significant main effects for the other five categories. Table 6 displays the summary statistics for the main effects and the means and standard deviations for each thought unit variable. The means indicate the average proportion of each type of thought unit across participants.

Table 6
ANOVA Summary Statistics for Thought Units

	Concurrent Condition *M* *(SD)*	Retrospective Condition *M* *(SD)*	Virtual Revisit Condition *M* *(SD)*	*F*	*p*	η^2
Planning	$0.407\ (0.70)^{ab}$	$3.96\ (3.83)^{a}$	$4.17\ (3.66)^{b}$	7.05	.002*	.251
Connecting	$6.04\ (3.59)^{a}$	$13.23\ (6.53)^{a}$	9.08 (3.81)	8.35	.001*	.285
Reasoning	$2.23\ (2.46)^{a}$	$10.99\ (7.65)^{b}$	$19.27\ (9.16)^{ab}$	22.01	.000**	.512
Reflecting	$16.71\ (7.70)^{ab}$	$26.86\ (11.45)^{a}$	$25.08\ (7.48)^{b}$	5.36	.008*	.203
Evaluating Website Content	10.87 (4.96)	12.18 (8.71)	15.36 (5.27)	1.87	.166	.082
Evaluating User Experience	7.96 (6.98)	12.13 (9.03)	7.47 (3.30)	2.09	.136	.091
Diversion	5.97 (3.01)	0.79 (2.08)	2.99 (3.42)	13.65	.067	.394
Understanding	5.99 (3.50)	2.97 (4.46)	3.36 (3.90)	2.88	.067	.121
Describing Procedural Behavior	11.64 (6.21)	11.71 (9.12)	9.37 (4.91)	0.55	.582	.025
Describing Website Features	$10.15\ (4.87)^{ab}$	$3.45\ (4.83)^{a}$	$3.56\ (3.53)^{b}$	11.05	.000**	.345
Reading	$22.03\ (12.79)^{ab}$	0.00^{a}	$0.29\ (0.64)^{b}$	43.81	.000**	.676

*$p < .05$, **$p < .01$
a, b = significant post hoc comparisons
Note: $df = (2, 42)$ for all variables

Post hoc comparisons using Tukey HSD revealed that participants in the retrospective and virtual revisit conditions verbalized a significantly greater proportion of planning thought units than participants in the concurrent condition ($p < .05$). In addition, participants in the retrospective condition verbalized a significantly greater proportion of connecting thought units than participants in the concurrent condition ($p = .001$), and participants in the virtual revisit condition verbalized a significantly greater proportion of reasoning thought units than participants in the other two conditions ($p < .05$). Furthermore, post hoc tests using Tukey HSD revealed that participants in both the retrospective and virtual revisit conditions verbalized a significantly greater proportion of reflecting thought units than participants in the concurrent condition ($p < .05$). Finally, participants in the concurrent condition verbalized a significantly greater proportion of describing website features thought units and reading thought units than participants in the other two conditions ($p < .05$).

With respect to the remaining thought units, findings indicated no significant differences between the three conditions for evaluating, diversion, understanding, and describing procedural behaviors. This suggests that regardless of the type of think aloud employed, participants will verbalize a relatively equal number of thoughts related to evaluating the website, to their confusion and understanding of the web-based tools and information, and to descriptions of their own actions and online behaviors.

A one-way between subjects ANOVA was also performed to compare the frequency of words and thought units verbalized in each think aloud condition. The analysis revealed a significant main effect for both variables across the three conditions: $F(2, 42) = 130.42$, $p < .001$, $\eta^2 = .861$ for word count; and $F(2, 42) = 90.76$, $p < .001$, $\eta^2 = .812$ for number of thought units. As Table 7 shows, participants in the retrospective condition verbalized the fewest number of words ($M = 658.80$, $SD = 229.00$) and thought units ($M = 30.73$, $SD = 13.15$). Participants in the concurrent condition verbalized more than twice the number of words as participants in the retrospective condition ($M = 1676.13$, $SD = 394.89$) and more than four times the number of thought units than participants in the retrospective condition ($M = 147.40$, $SD = 35.83$). Participants in the virtual revisit condition verbalized the most number of words ($M = 2637.87$, $SD = 359.94$). However, participants in the virtual revisit condition verbalized less thought units than teachers in the concurrent condition ($M = 123.27$, $SD = 20.58$). This finding suggests that thought units verbalized by participants in the virtual revisit condition contained more words than thought units verbalized by participants in the concurrent condition.

Table 7
ANOVA Summary Statistics for Word Count and Number of Thought Units

	Concurrent Condition	Retrospective Condition	Virtual Revisit Condition			
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*F*	*p*	η^2
Word Count	1676.13 (394.89)	658.80 (229.00)	2637.87 (359.94)	130.421	.000**	.861
Number of Thought Units	147.40 (35.83)	30.73 (13.15)	123.27 (20.58)	90.756	.000**	.812

**$p < .001$
Note: $df = (2, 42)$ for all variables

Post hoc comparisons using Tukey HSD revealed significant differences in the frequency of words ($p < .01$) and the frequency of thought units ($p < .05$) between all three conditions. In contrast to the participants in the concurrent and virtual revisit conditions who were asked to verbalize their thoughts for 20 minutes, participants in the retrospective condition were not given a time constraint to verbalize their thoughts. Although an unlimited amount of time was given to participants in the retrospective condition, these participants verbalized their thoughts for an average four minutes 30 seconds. Therefore, it is not surprising that participants in the retrospective condition verbalized significantly fewer words and thought units than the other two conditions. A more interesting finding is the difference in word counts and number of thought units between the concurrent and virtual revisit conditions. Participants in the virtual revisit condition verbalized significantly more words than participants in the concurrent condition; however, the reverse is true for the number of thought units. This finding indicates that thought units produced by participants in the virtual revisit condition contained a greater number of words. This suggests that the thought units verbalized by participants in the virtual revisit condition were more complex than the thought units verbalized by participants in the concurrent condition. For example, as a participant from the concurrent condition views the homepage for the first time she describes her procedural behavior: “I'm going to have a look at the How To Videos.” A participant from the virtual revisit condition who also views the homepage for the first time goes further to provide a reason for her behavior: “I'm always interested in ways to increase my students' background knowledge of vocabulary and comprehension so I go back and forth between comprehension and vocabulary before I narrowed it down and selected vocabulary.” Both of these thought units were verbalized during the participants' initial view of the home page. However, the thought unit verbalized by the participant in the virtual revisit condition is more complex in that it provides a reason for her navigational choice. While participants in the concurrent condition verbalized on average more thought units, the verbalizations were less likely to include reasons for their decisions. Examples of thought units produced by participants in the concurrent condition include: that's interesting; let's see what that is; there's a word wall; that's like what we did in kindergarten.

Pearson correlations. Pearson correlations were computed to determine three relationships: between the cognitive processes, between the frequency of word counts and thought units, and between the cognitive processes and word counts and thought units (see Table 8). According to the results, higher order cognitive processes (planning, connecting, reasoning, and reflecting) were positively correlated to other higher order cognitive processes, and negatively correlated to lower order cognitive processes (diversion, understanding, describing website features, and reading). Similarly, lower order cognitive processes were more likely to be positively correlated to other lower order cognitive processes and negatively related to higher order cognitive processes. For instance, reading was negatively correlated to planning ($r = -.42$, $p < .01$), connecting ($r = -.39$, $p < .01$), reasoning ($r = -.50$, $p < .01$), reflecting ($r = -.55$, $p < .01$), evaluating website content ($r = -.31$, $p < .05$), and evaluating user experience ($r = -.37$, $p <.05$). In contrast, reading was positively related to diversion ($r = .46$, $p < .01$), understanding ($r = .30$, $p < .05$), and describing website features ($r = .45$, $p < .01$). These findings corroborate the above ANOVA results and suggest that participants in the concurrent condition, who were more likely to read text during the think aloud, were less likely to verbalize thought units related to higher order cognitive processes.

Pearson correlations also revealed a positive correlation between the frequency of words and the frequency of thought units ($r = .70$, $p < .01$). This finding indicates that as the number of words increased the number of thought units increased. Results also revealed a negative correlation between higher order cognitive processes and the number of thought units verbalized.

This suggests that teachers who verbalized thought units related to higher order processes were more likely to verbalize fewer thought units than teachers who verbalized thought units related to lower order processes.

Table 8
Pearson Correlations

	1	2	3	4	5	6	7	8	9	10	11
Cognitive Processes											
1. Planning	1										
2. Connecting	.31*	1									
3. Reasoning	.36*	.06	1								
4. Reflecting	.19	.38*	-.10	1							
5. Evaluating Website	-.23	.03	.03	.23	1						
6. Evaluating Experience	.05	-.08	-.09	.16	-.34*	1					
7. Diversion	-.36*	-.39**	-.22	-.46**	-.17	-.14	1				
8.Understanding	-.26	-.35*	-.21	-.29	-.34*	.05	.38*	1			
9. Describing Behaviour	-.12	-.20	.03	-.28	-.56**	-.47**	-.003	.01	1		
10. Describing Website	-.43**	-.33*	-.45**	-.31*	-.13	-.23	.39**	.23	.10	1	
11. Reading	-.42**	-.39**	-.50**	-.55**	-.31*	-.37*	.46**	.30*	.18	.45**	1

*$p < .05$, **$p < .01$

Discussion

The findings of this study indicated significant differences in participants' verbalizations across the three think aloud conditions. One possible reason for these differences relates to cognitive load. Cognitive load refers to the amount of mental effort in working memory: cognitive load "represents the load that performing a particular task imposes on the cognitive system" (Pass & Van Merrienboer, 1994, p. 122). Cognitive load may have impacted participants in the concurrent condition who were required to use a higher level of mental energy to process the given information as they simultaneously completed the website task (Ping Lim, 2004). The cognitive load on working memory may have diminished the quality of their verbalizations. This is consistent with Ericsson and Simon (1993) who state that concurrent reporting may become difficult to maintain under high cognitive load conditions. Cognitive load research also indicates that during multimedia tasks, concurrent reporting will interfere with information processing and limit the extent of the thought units (Nielsen, Clemmensen, & Yssing, 2002; van Gog et al., 2009). As the participants in the concurrent condition verbalized their thoughts while simultaneously completing the website task, the cognitive demand on their working memory increased: their brains prioritized information processing over verbalizations. Therefore, participants in the concurrent condition verbalized fewer thought units related to higher cognitive processes (Cooke, 2010).

In contrast, participants in the retrospective and virtual revisit conditions had the advantage of thinking aloud after the website task had been completed; thus, the cognitive load on their working memory was lighter during their think aloud than those in the concurrent condition. As a result of a lighter cognitive load, participants in the retrospective and virtual revisit conditions had more cognitive resources available during the think aloud; therefore, they could focus their mental energy on constructing and verbalizing complex ideas.

For example, planning is a complex cognitive activity and involves visualizing the future, producing and generating new information, and "putting elements together to form an original product" (Krathwohl, 2002, p. 215). Planning is also an indicator of teachers' intentions for practice (Clark & Peterson, 1986; Krathwohl, 2002). The planning thought units verbalized by participants in the retrospective and virtual revisit conditions contained complex information that demonstrate intentions for practice. For instance, a video on social studies and writing integration led a participant in the retrospective condition "to think of a different idea that [she] could bring into [her] class which includes doing cross-curricular work with probability and ancient civilizations and trading cards and games." Similarly, a teacher in the virtual revisit condition described how she could modify a lesson to include cross-curricular integration: "Building interviewing and report writing skills, I know that there is a lot that I can do with this in terms of reading and writing and oral and drama and themes like social justice and history and so on." These examples demonstrate complex verbalizations; teachers are planning as they use the professional development website—they are visualizing the future and beginning to generate new ideas.

Another possible reason for the findings can be drawn from research on information processing and memory recall. Information processing is enhanced when new incoming information is connected to prior knowledge and previous experiences (Mastin, 2010; Weber, Corrigan, Fornash, & Neupauer, 2003). Information processing is also enhanced when new material is interesting to the learner (Garner & Gillingham, 1991). The more deeply new information is processed (i.e., through connections and interest), the more likely it will be recalled. In contrast to participants in the concurrent condition, participants in the retrospective condition were able to process information during the website task on a much deeper level because they did not have the same cognitive demands of verbalizing their thoughts while simultaneously completing the website task. They were able to verbalize significantly more connections with past experiences and interests than participants in the concurrent condition. For instance, a participant in the retrospective condition connected components of the website to her current reading program:

> After visiting a virtual tour, I was interested in going through all the different parts of the balanced literacy and I clicked on a couple of things that interested me and that I'm working on, like comprehension skills and fluency and word building.

In addition, participants in the retrospective condition did not have access to a visual cue or memory aid. It is possible that in the absence of a visual cue, participants in the retrospective condition verbalized more connecting thought units because they recalled meaningful memories established during the website task. For instance, a participant in the retrospective condition connected website content on hand writing to a student in her class: "I saw something about hand writing which made me think about a student that I have in grade one who…it's quite a struggle for her to read what she's writing." The information that this participant recalled was meaningful to her because it directly related to the needs of a student in her classroom.

While the retrospective procedure may produce verbalizations related to meaningful connections, the retrospective procedure is limited by the fact that resulting verbalizations are based on the ability to recall information. Participants in the retrospective condition of the current study omitted most of their navigation, particularly the intermediate web-based behaviors. In contrast, participants in the virtual revisit condition had direct access to their web-based actions via the screen-capture recording. The screen recording captured participants' website navigation and acted as a visual aid during the think aloud. One possibility is that participants in the virtual revisit condition utilized the visual information as an aid to recall their navigational decisions and *why* they made them. Available cognitive resources and direct access to web-based actions allowed participants in the virtual revisit condition to extend their descriptions and clarify the reasons for their navigational choices. While the screen recording could have acted as a memory aid, it is also possible that the screen recording prompted participants to generate rationalizations for their decisions during the think aloud. In any case, the virtual revisit allowed participants to produce thorough verbalizations related to their navigational decisions and why they made them. Reasoning about behavior moves beyond simple descriptions of actions and offers rich explanations about decisions. For example, a participant from the virtual revisit condition provided a descriptive rationale for her decision to remain on a particular webpage for an extended period of time: "At this point I was trying to just read what the student wrote to get an idea of whether they were creating their own stories or whether they were doing more of a retell." Reasoning thought units provide thorough descriptions, clarifications, extensions and overall greater insight into participants' navigational choices—participants explain *why* they make particular navigational decisions.

Overall findings from this study reveal benefits and limitations to employing each type of think aloud method in the context of a professional development website. A benefit of the concurrent think aloud is that it generates direct data about the ongoing cognitive processes that occur during task performance. Since the two activities, thinking aloud and task performance occur simultaneously the verbalizations are valid forms of information—the verbalizations contain direct data about participants' thoughts in real time. The limitations of the concurrent method, however, may outweigh this benefit. The first limitation of the concurrent think aloud method is the fact that this think aloud produces fewer verbalizations related to higher order cognitive processes. Secondly, the concurrent think aloud method results in a high cognitive load on working memory.

The main benefit of employing the retrospective procedure is that this method produces verbalizations related to higher cognitive processes (planning, connecting, and reflecting). However, the limitations of employing the retrospective think aloud method are significant. Participants will omit most of their online actions and navigational decisions. Furthermore, participants will most likely have difficulty recalling their intermediate web-based actions and reasons for these decisions.

The main benefit of employing the virtual revisit procedure is that this method produces verbalizations related to higher level cognitive processes (planning, reasoning, and reflecting). Secondly, participants can rely on a visual aid to help them recall their navigational decisions and why they made them. One limitation of the virtual revisit think aloud is the time required to complete both the website task and think aloud task. In the current study, the total time for each participant in the virtual revisit condition was 40 minutes. For some researchers, this may be costly. Moreover, the time required to employ the virtual revisit think aloud may limit participant involvement.

Collectively, the findings indicate that the virtual revisit can avoid the limitations of the concurrent and retrospective procedures and provide thorough and descriptive thought units and insights into how teachers use and learn from a professional development website.

Study Limitations

There were three main limitations to this study: (1) factors that may have caused reactivity, (2) the use of one professional development website, and (3) the possibility of researcher bias. First, factors may have caused reactivity during the think aloud procedure. Reactivity occurs when task performance is altered as a result of an awareness of the study task. Reactivity may have occurred as a result of participants' awareness that they were completing a task in the presence of the primary investigator. A "motivational shift" in which the participants anticipate exposure of their think aloud protocol may have occurred when the participants were informed of the think aloud procedure (Russo, Johnson, & Stephens, 1989). Another factor that may have caused reactivity is hearing one's own voice. The additional aural stimulation may have interfered with the concurrent navigation (Russo et al., 1989). In general, reactivity was reduced as much as possible during the one-on-one meeting by: staying neutral during the task and think aloud, keeping verbal and nonverbal cues to a minimum, and providing participants in the retrospective and virtual revisit conditions with the second part of the instructions after they had completed the website task. However, the factors described above should be considered when interpreting the findings.

Secondly, the current study was context-specific and used one professional development website. Future research comparing the three think aloud conditions should be conducted with additional websites and online resources. Conducting similar studies with alternative professional development websites will enhance the credibility and transferability of the results.

Finally, it is difficult to eliminate researcher bias. Steps were taken to reduce researcher bias as much as possible: the use of relevant literature to develop the coding scheme, unitizing the transcripts based on Lincoln and Guba's (1985) procedures, using a second coder procedure, and staying as close to the data as much as possible during the analysis and interpretation of the findings. However, this limitation must be considered when interpreting the findings. To avoid researcher bias, future research could involve a team of researchers with varying backgrounds, particularly during data analysis.

Significance and Educational Implications

The significance of this research is that it compares the effectiveness of two traditional think aloud methods—concurrent and retrospective—with a think aloud method combining a retrospective procedure with screen capture technology (the virtual revisit). While think aloud research is extensive, studies that compare the effectiveness of different think aloud methodologies for understanding cognitive processes as website users navigate online resources are scarce. Based on the findings of this study, there may be potential benefits and limitations to employing each type of think aloud method. In addition, the virtual revisit think aloud, a relatively underused type of think aloud method, appears to be an effective method for examining teachers' cognitive processes as they use a professional development website. The virtual revisit think aloud method produces thorough verbalizations that incorporate reasons behind the decision-making process.

Another significance of the current study is that the participants were practicing teachers with between one and over 25 years of classroom experience. Much of the research examining teachers' beliefs and attitudes towards learning in online environments has studied preservice teachers. Studying the cognitive processes of practicing teachers has the potential to contribute to the understanding of teacher professional development and teacher cognition.

Furthermore, the virtual revisit think aloud method also has potential that transcends the specific domain in which it is applied in the current study. Virtual revisits with think aloud could be applied to examine cognitive processes of participants in research involving online learning and website use in many domains in education. If researchers provide feedback to educational website developers based on the virtual revisit methodology, improvements can be made to the design and content of their sites.

Conclusion

Teachers learn in many different aspects of practice, including their classrooms, their school communities, professional development courses, and online environments (Borko, 2004). To understand teacher learning, it must be studied within these multiple contexts and it must be studied with an effective methodology that provides rich and thorough data about the reasoning process. Gaining greater insights into teachers' cognitive processes as they navigate online environments can lead to the reconsideration of the design of online learning environments so that they "are more conducive to informal learning…so that they further develop the ability of professionals to solve problems and learn independently" (Lohman, 2006, p. 144).

References

Aranyi, G., van Schaik, P., & Barker, P. (2012). Using think-aloud and psychometrics to explore users' experience with a news website. *Interacting with Computers, 24*, 69-77. doi: 10.1016/j.intcom.2012.01.001

Barzilai, S. & Zohar, A. (2012). Epistemic thinking in action: Evaluating and integrating online sources. *Cognition and Instruction*, *30*(1), 39-85.

Boren, M.T., & Ramey, J. (2000). Thinking aloud: Reconciling theory and practice. *IEEE Transactions on Professional Communication, 43*(3), 261-278. doi:10.1177/1049732309354278

Borko, H. (2004). Professional development and teacher learning: Mapping the Terrain. *Educational Researcher, 33* (8), 3-15. doi: 10.3102/0013189X033008003

Branch, J. (2006). Using think alouds, think afters, and think togethers to research adolescents' inquiry experiences. *Alberta Journal of Educational Research, 52*(3), 148-159.

Cervetti, G.N., Kulikowich, J., Drummond, M., & Billman, A. (2012). *Influence of educative curriculum materials on teachers' learning and teaching.* Paper presented at the annual meeting of the American Educational Research Association, Vancouver, B.B. Retrieved August 8, 2014 from http://www.sandi.net/cms/lib/CA01001235/Centricity/Domain/121/R_R/DistResearch/ResultReports/585_Cervetti_G_et_al_UCBerkeley_2012B.pdf

Charalambousa, K., & Ioannou, I. (2011). The attitudes and opinions of Cypriot primary teachers about the use of the Internet for their professional development and as an educational tool. *Learning, Media and Technology*, *33*(1), 45-57.

Clark, C.M., & Peterson, P.L. (1986). Teachers' thought processes. In M.C. Wittrock, (Ed.), *Handbook of research on teaching (3rd ed.)* (pp.255-296). New York: MacMillan Publishing Company.

Cooke, L. (2010). Assessing concurrent think-aloud protocol as a usability test method: A technical communication approach. *IEEE Transactions on Professional Communication, 53*(3), 202-215. doi: 10.1109/TPC.2010.2052859

Damico, J., & Balidon, M. (2007). Examining ways readers engage with websites during think-aloud sessions. *Journal of Adolescent and Adult Literacy*, *51*(3), 254-263. doi: 10.1598/JAAL.51.3.5

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Research, 38*(3), 181-199. doi: 10.3102/0013189X08331140

Dillon, A. (2001) Beyond usability: process, outcome and affect in human-computer interactions. *Canadian Journal of Library and Information Science*, *26*(4), 57-69.

Duncan-Howell, J. (2010). Teachers making connections: Online communities as a source of professional learning. *British Journal of Educational Technology*, *41*(2), 324-340. doi: 10.1111/j.1467-8535.2009.00953.x

Ericsson, K. A., & Simon H. (1984). *Protocol analysis: Verbal reports as data.* Cambridge, MA: MIT Press.

Ericsson, K. A., & Simon, H. (1993). *Protocol analysis: Verbal reports as data* (Rev. ed.). Cambridge, MA: MIT Press.

Ericsson, K. (2002). Towards a procedure for eliciting verbal expression of non-verbal experience without reactivity: Interpreting the verbal overshadowing effect within the theoretical framework for protocol analysis. *Applied Cognitive Psychology, 16*, 981-987. doi: 10.1002/acp.925

Ericsson, K. (2003). Valid and non-reactive verbalization of thoughts during performance of tasks: Towards a solution to the central problems of introspection as a source of scientific data. *Journal of Consciousness Studies, 10*(9-10), 1-18.

Gaissmaier, W., Fific, M., & Rieskamp, J. (2010). Analyzing response times to understand decision processes. In M. Schulte-Mecklenbeck, A. Kuhberger, & R. Ranyard (Eds.), *The Handbook of Process Tracing Methods for Decision Research: A Critical Review and User's Guide*. (p. 89-114). New York, NY: Psychology Press.

Gambrell, L.B., Malloy, J.A., & Anders Mazzoni, S. (1999). Evidence-based best practices in comprehensive literacy instruction. In Mandel Marrow, L. & Gambrell, L.B. (Eds.), *Best practices in literacy instruction* (11-36). New York: The Guilford Press.

Garner, R., & Gillingham, (1991). Topic knowledge, cognitive interest, and text recall. *The Journal of Experimental Education, 59*(4), 310-319.

Guan, Z., Lee, S., Cuddihy, E., & Ramey. J. (2006). The validity of the stimulated retrospective think-aloud method as measured by eye tracking. *CHI, April 22-27.* doi: 10.1145/1124772.1124961

Guetzkow, H. (1950). Unitizing and categorizing problems in coding qualitative data. *Journal of Clinical Psychology, 6*, 47-58. doi: 10.1002/1097-4679(195001)6:1<47::AID-JCLP2270060111>3.0.CO;2-I

Horz, H., & Schnotz, W. (2010). Cognitive load in learning with multiple representations. In Plass, J.L., Moreno, R., & Brunken, R. (Ed.), *Cognitive load theory* (229-252). New York: Cambridge University Press.

Hur, J. W., & Brush, T. A. (2009). Teacher participation in online communities: Why do teachers want to participate in self-generated online communities of K-12 teachers? *Journal of Research on Technology in Education, 41*(3), 279-303.

Jaspers, M. W. (2009). A comparison of usability methods for testing interactive health technologies: Methodological aspects and empirical evidence. *International Journal of Medical Informatics, 78,* 340-353. doi: http://dx.doi.org/10.1016/j.ijmedinf.2008.10.002

Jaspers, M., Steen, T., van den Bos, C., & Geenen, M. (2004). The think aloud method: A guide to user interface design. *International Journal of Medical Infomatics, 73,* 781-795. doi: 10.1016/j.ijmedinf.2004.08.003

Kansanen, P., Tirri, K., Meri, M., Krokfors, L., Husu, J., & Jyrhama, Riitta. (2000). Teachers' pedagogical thinking: Theoretical landscapes, practical challenges. New York: NY, Peter Lange Publishing, Inc.

Kanuka, H., & Nocente, N. (2003). Exploring the effects of personality type on perceived satisfaction with web-based learning in continuing professional development. *Distance Education*, *24*(2), 227-244. doi: 10.1080/0158791032000127491

Kao, C. P., Wu, Y. T., & Tsai, C. C. (2011). Elementary school teachers' motivation toward web-based professional development, and the relationship with Internet self-efficacy and belief about web-based learning. *Teacher and Teacher Education, 27*, 406-415.

Kuusela, H., & Paul, P. (2000). A comparison of concurrent and retrospective verbal protocol analysis. *American Journal of Psychology*, *113* (3), 387-404.

Krathwohl, D. R. (2002). A revision of Bloom's Taxonomy: An overview. *Theory into Practice, 41*(4), 212-218.

Laberge, J.C., & Scialfa, C.T. (2005). Predictors of web navigation performance in a life span sample of adults. *The Journal of the Human Factors and Ergonomics Society, 47*(2), 289-302. doi: 10.1518/0018720054679470

Lewis, A., & Smith, D. (1993). Defining higher order thinking. *Theory into Practice, 32*(3), 131-137.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.

Lohman, M. C. (2006). Factors influencing teachers' engagement in informal learning activities. *Journal of Workplace Learning, 18*(3), 141-156.

Lombard, M., Snyder-Duch, J., & Campanella Bracken, C. (2002). Content analysis in mass communication assessment and reporting of intercoder reliability. *Human Communication Research, 28* (4), 587-604.

Mastin, L. (2010). *The human memory: Memory encoding*. Retrieved from http://www.human-memory.net/processes_encoding.html

Mayer, R. (2002). Cognitive theory and the design of multimedia instruction: An example of the two-way street between cognition and instruction. *New Directions for Teaching and Learning, 2002*(89), 55-71.

McDonald, L. (2012). Teacher learning, transfer and motivation: facilitating improved outcomes for students. *Journal of Educational Sciences & Psychology, 2*(2), 15-24.

McDonald, S., Edwards, H., & Zhao, T. (2012). Exploring think-alouds in usability testing: An international survey. *IEEE Transactions on Professional Communication, 55* (1), 2-19. doi: 10.1109/TPC.2011.2182569

McLoughlin, D., & Mynard, J. (2009). An analysis of higher order thinking in online discussions. *Innovations in Education and Teaching Internationally, 46*(2), 147-160. doi: 10.1080/14703290902843778

Moos, D. (2014). Setting the stage for the metacognition during hypermedia learning: What motivation constructs matter? *Computers & Education 70,* 128-137. doi: 10.1016/j.compedu.2013.08.014

Nielsen, J., Clemmensen, T., & Yssing, C. (2002). Getting access to what goes on in people's heads?-Reflections on the think-aloud technique. NordiCHI, October 19-23.

Olmsted-Hawala, E., Murphy, E., & Hawala, S. (2010). Think-aloud protocols: Analyzing three different think-aloud protocols with counts of verbalized frustrations in a usability study of an information-rich website. *IEEE International Professional Communication Conference.*

Pearson, J., Pearson, A., & Green, D. (2007). Determining the importance of key criteria in web usability. *Management Research News, 30*(11), 816-828.

Page K.L., Robson, M.J., & Uncles, M.D. (2012). Perceptions of web knowledge and usability: When sex and experience matter. *International Journal of Human-Computer Studies, 70*(12), 907-919.

Paas, F.G., & Van Merrienboer, J. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. *Journal of Educational Psychology, 86*(1), 122-133.

Ping Lim, C. (2004). Engaging learners in online learning environments. *Tech Trends, 48*(4), 16-23. doi: 10.1007/BF02763440

Pressley, M., Mohan, L., Raphael, L.M., & Fingeret, L. (2007). How does Bennett Woods Elementary School produce such high reading and writing achievement? *Journal of Educational Psychology, 99*(2), 221-240.

Ranyard, R., & Svenson, O. (2010). Verbal data and decision process analysis. In M. Schulte-Mecklenbeck, A. Kuhberger, & R. Ranyard (Eds.), *The Handbook of Process Tracing Methods for Decision Research: A Critical Review and User's Guide*. (p. 89-114). New York, NY: Psychology Press.

Rosenzweig, C., Krawec, J., & Montague, M. (2011). Metacognitive strategy use of eighth-grade students with and without learning disabilities during mathematical problem solving: A think-aloud analysis. *Journal of Learning Disabilities, 44*(6), 508-520. doi: 10.1177/0022219410378445

Russo, J., Johnson, E., & Stephens, D. (1989). The validity of verbal protocols. *Memory & Cognition, 17* (6), 759-769. doi: 10.3758/BF03202637

Tan, G. W., & Wei, K. K. (2006). An empirical study of web browsing behaviour: Towards an effective website design. *Electronic Commerce Research and Applications, 5,* 261-271.

Uppal, T. (2011). *Internet use for lesson preparation by pre-service teachers: A mixed methods study*. (Unpublished Master's thesis.) OISE, University of Toronto, Toronto, Canada.

van Gog, T., Kester, L., Nievelstein, F., Giesbers, B., & Paas, F. (2009). Uncovering cognitive processes: Different techniques that can contribute to cognitive load research and instruction. *Computers in Human Behavior, 25*, 325-331. doi: 10.1016/j.chb.2008.12.021

van Gog, T., Paas, F., van Marrienboer, J., & Witte, P. (2005). Uncovering the problem-solving process: Cued retrospective reporting versus concurrent and retrospective reporting. *Journal of Experimental Psychology: Applied, 11* (4), 237-244.

van Someren, M. W., Barnard, Y. F., & Sandberg, J. A. (1994). *The think aloud method: A practical guide to modelling cognitive processes.* London: Academic Press.

Weschke, B., & Barclay, R. (2011). Online teacher education: Exploring the impact of a reading and literacy program on student learning. *Online Learning, 15* (2), 22-43.

Zhao, T., & McDonald, S. (2010). Keep talking: An analysis of participant utterances gathered using two concurrent think-aloud methods. *Proceedings: NordiCHI,* 16-20.

Moving Beyond Smile Sheets: A Case Study on the Evaluation and Iterative Improvement of an Online Faculty Development Program

Ken-Zen Chen
National Chiao-Tung University

Patrick R. Lowenthal, Christine Bauer, Allan Heaps, Crystal Nielsen
Boise State University

Abstract

Institutions of higher education are struggling to meet the growing demand for online courses and programs, partly because many faculty lack experience teaching online. The eCampus Quality Instruction Program (eQIP) is an online faculty development program developed to train faculty to design and teach fully online courses. The purpose of this article is to describe the eQIP (one institution's multipronged approach to online faculty development), with a specific focus on how the overall success of the program is evaluated using surveys, analytics, and social network analysis. Reflections and implications for improving practice are discussed.

Keywords: online education, faculty development, instructional consultation, online course design

Chen, K., Lowenthal, P.R., Bauer, C., Heaps, A., & Nielsen, C. (2017). Moving beyond smile sheets: A case study on the evaluation and iterative improvement of an online faculty development program, *Online Learning 21*(1), 85-111. doi: 10.24059/olj.v21i1.810

Introduction

As online enrollments grew over the last decade, Boise State University was confronted with a common problem in higher education: What is the best way to design and develop high-quality online courses and support faculty as they teach? High-quality online courses begin and end with high quality faculty (Dunlap, 2005; Wilson, Ludwig-Hardman, Thornam, & Dunlap, 2004); most faculty, though, have never taken nor taught a course online before (see Bacow, Bowen, Guthrie, Lack, & Long, 2012; Tipple, 2010; Wolf, 2013). To succeed in the online classroom, faculty must possess, among other things, pedagogical, facilitative, instructional, social, managerial, assessment, and technical competencies (Baran, Correia, & Thompson, 2011). Thus, the challenge institutions face is how to develop these competencies in faculty so that they are able to develop and teach high quality online courses. Boise State created the eCampus

Quality Instruction Program (eQIP), an online faculty development program, to address this challenge. The purpose of this article is to describe the eQIP and the multipronged approach used to evaluate the eQIP.

Literature Review

Institutions across the country have used multiple approaches to address the problem of how to develop and teach high quality online courses. On one end of the continuum are universities that essentially leave it up to faculty to decide what courses they teach online and how they develop them. On the other end of the continuum are institutions, often with a for-profit business model, that have adopted a centralized "enterprise" model in which all online course development and training takes place (Herman, 2012; Lowenthal & White, 2009). Most institutions lie somewhere in between these two extremes, offering some type of support for online course development and teaching. For instance, the University of Colorado Denver takes a decentralized approach in which faculty, for the most part, decide what they want to teach online, but may be supported by CU Online (a centralized unit for faculty support) when needed; in addition to just-in-time instructional design support, CU Online also offers regular workshops and annual events to support faculty developing online courses (Lowenthal & Thomas, 2010b). On the other hand, faculty at the University of Central Florida complete a cMOOC (connectivist Mass Open Online Course) that prepares them to teach online (Moskal, Thompson, & Futch, 2015). Institutions like the University of the District of Columbia fully support faculty to develop online courses and complete a Quality Matters Standards and Peer Review for online courses to improve the quality of their online courses (Britto, Ford, & Wise, 2014; see Meyer, 2013, p.95-96 for a list of faculty development programs). There are typically three components common to models of support for online learning: (1) using instructional designers to design or support faculty as they design courses, (2) training faculty to teach online, and (3) using some type of quality control system to evaluate and in turn improve the quality of online courses (e.g., Quality Matters or the Online Consortium's 5 Pillars quality framework).

Building upon what others have already done, Boise State implemented a program called the eCampus Quality Instruction Program (eQIP) to improve the design, development, and delivery of online courses. The following paper describes the program Boise State developed and the mixed methods approach taken to evaluate and improve the program over time.

Background: The eCampus Quality Instruction Program (eQIP)

Boise State University began offering courses online in 1989. Over the years, mirroring national trends, enrollments grew. Due to this growth, as well as a desire to develop more online programs, the eCampus Center created the eCampus Quality Instruction Program (eQIP) in the fall of 2012 to improve how courses are designed and taught online at Boise State. The eQIP consists of three distinct components:

- Course Design: A course design seminar;
- Quality Assurance: A Quality Matters Peer Review, and
- Teacher Training: A teaching online seminar.

Each component is described briefly in the following paragraphs.

Course Design Seminar

To teach faculty how to design online courses and to assist them in the development of these courses, eCampus developed a 12-week online course design seminar called the eCampus Course Design and Development Seminar or eCD2S for short. Most faculty do not have the experience or skillset to design online courses (e.g. Bailey & Card, 2009; Baran, Correia, & Thompson, 2011; Marek, 2009; Moar, 2006; Schrum, Burbank, Engle, Chambers, & Glassett, 2005); faculty are content experts, many of whom have limited coursework or structured experiences in teaching and learning online (e.g., Sellani & Harrington, 2002). The eCD2S is designed specifically for faculty with no previous experience with online course design. The seminar has been iteratively improved over time (see Chen, 2014). The seminar currently consists of two distinct phases:

1. Phase 1−Design Phase (Weeks 1-5): A group of faculty, facilitated by instructional designers, work online in a Learning Management System to learn the skills necessary for online course design;

2. Phase 2−Development Phase (Week 6-12): The faculty work one-on-one with instructional designers to develop their online courses.

Over time, it became clear that even after faculty completed the seminar, many desired and benefited from additional support when designing future online courses. As a result, a shorter, condensed version of the course design seminar called the eCampus Course Development Phase or eCD was created. The eCD was designed for faculty who have completed the 12-week course design seminar (i.e., eCD2S) at Boise State. The eCD essentially focuses only on the development phase (i.e., Phase 2) of the original 12-week seminar. In the eCD, during 8 weeks, faculty work one-on-one with an instructional designer, while collaborating with colleagues, to develop a turnkey-ready online course (which we refer to as a "Master Course"). Collaborating with colleagues and instructional designers helps improve the quality of the online courses faculty design, even with faculty who already have experience designing online courses (Lowenthal & White, 2009).

Quality Matters Peer Review

After faculty develop an online course (in either the eCD2S or the eCD) and teach it for one semester, the online course is peer reviewed using the Quality Matters' framework. Quality Matters (QM) is a nationally recognized quality assurance program used to improve the design of online courses (MarylandOnline, 2011). Following QM's peer-review process, each course developed in the course design seminars is peer-reviewed by three trained QM faculty reviewers (Carter-Cram, 2014; see also MarylandOnline, 2011, for more on the QM peer-review model). After a course is peer reviewed, the faculty member who originally developed the course updates the course based upon the feedback received through the QM peer-review process. The updated course then becomes a "master" online course—that is, one that can be reused and taught by any Boise State faculty (see Lowenthal & White, 2009, for more on master courses and centralized models of course development).

Teaching Online Seminar

Teaching online is different than teaching face-to-face (Stone & Perumean-Chaney, 2011; Xu & Morris, 2007). Without proper training, experienced face-to-face faculty often struggle when moving to an online environment (Natriello, 2005). While some universities offer a face-to-face workshop on how to teach online, Boise State decided that it would be more authentic and convenient to train faculty to teach online in a fully online course. The teaching online seminar, called the eCampus Teaching Online Seminar or eTOS, is a 6-week online seminar facilitated by full-time faculty (Carter-Cram, 2014)—not instructional designers. In this seminar, faculty learn how to teach online, what it feels like to be a part of an online faculty community (Brooks, 2010; Schrum et al., 2005), how to transfer evidence-based pedagogical practices (Bailey & Card, 2009), and finally what it is like to be an online student (Fein & Logan, 2003). This seminar was designed both for faculty who have never taught online as well as faculty with years of experience teaching online (see Table 2).

While the eQIP has been offered for a few years now, this case study focuses on the faculty and evaluation strategies used during 2014. In 2014, 51 faculty took part in the course design seminars, and 65 faculty took part in the teaching online seminars (see Table 1).

Table 1
Number of eQIP participants in 2014

	Spring	Summer	Fall	Total
12-week Course Design Seminar (eCD2S)	15	10	10	35
8-week Course Development Seminar (eCD)	5	8	3	16
Teaching Online Seminar (eTOS)	26	24	15	65
Quality Matters Peer Reviews (QMPR)	9	14	15	38

Method

Meyer (2013) reviewed a series of published literature in online faculty development and criticized seven methodological flaws of the selected research. One of the flaws was that outcome measures of past faculty development research overly relies on faculty persons' "honesty and self-understanding" (p. 104) of the training. Unlike many faculty development programs that simply depend on a short post-workshop survey—sometimes pejoratively referred to as a "smile" sheet—we treat eQIP seminars more like for-credit online courses than snapshot workshops and therefore use a mixed-method approach of collecting data to find out what is working and what is not working in the program (Creswell, 2008; Creswell & Plano, 2007; Greene, 2007). Following standard practices of program evaluation (Stufflebeam & Shinkfield, 2007), empirical data were collected from participants in the course design seminar (i.e. eCD2S) and the teaching online seminar (i.e., eTOS) to evaluate the success of the eQIP and to make informed improvements of the program. This study was both evaluative and exploratory in nature. The following questions guided our inquiry:

1. How did faculty participate in the cCD2S and eTOS?
2. What were faculty perceptions of the cCD2S and eTOS?
3. How did faculty skills and dispositions change over time?
4. What concerns did faculty have with eCD2S and eTOS?

More specifically, at regular checkpoints, data from surveys (e.g., module surveys, entrance and exit surveys), learning analytics, and social network analysis was collected and analyzed to better understand faculty participation, perceptions, skills and dispositions, concerns, and thus, the overall effectiveness of the eQIP (see Table 2 for an overview of the data collected).

Table 2
Overview of Data Collection

Research Question	Seminar	Data
RQ1. Faculty Participation	eCD2S	Analytics of LMS data & Module Surveys
	eTOS	Analytics of LMS data & Module Surveys
RQ2. Faculty Perceptions	eCD2S	Exit Surveys
	eTOS	Exit Surveys
RQ3. Skills & Dispositions	eCD2S	Entrance, Midterm, & Exit Surveys
	eTOS	Entrance & Exit Surveys
RQ4. Faculty Concerns	eCD2S	Module Surveys, Midterm & Exit Surveys
	eTOS	Module Surveys & Exit Surveys

The quantitative data collected from the surveys was analyzed using descriptive statistics to reveal faculty participation, perceptions, dispositions, and concerns; the qualitative data were coded in Nvivo 10 using an open-coding method; looking for emergent themes, insights, feedback, commonalities, and differences (Merriam, 1998). Using multiple sources of data facilitated triangulation and meaning-making in the study (Lowenthal & Leech, 2009). In addition to data triangulation that was designed and collected by the first author, a methodological triangulation was conducted in turn. The synthesized findings were given feedback from the third, fourth, and fifth authors who operated the eQIP to obtain practitioner and insider perspectives. Finally, the second author joined the research team and served as the external member checker who critically reviewed and evaluated the validity and reliability of interpretations. Direct quotes supporting a finding were identified and labeled using the format of *seminar name*, *term*, and the *data source*.

Result and Discussion

The results and our reflections are presented in the following section, beginning with an analysis of faculty participation in the program, then faculty perceptions of the program, followed by how faculty skills and dispositions changed over time, and ending with faculty concerns about the program.

Faculty participation

Faculty are busier than ever before (Lucas & Murry, 2011). However, simply putting faculty development online does not magically make it easier for faculty to participate. Research suggests that scheduling conflicts and overall lack of interest prevent online faculty from engaging in faculty development (Dailey-Hebert, Mandernach, Donnelli-Sallee, & Norris, 2014). Given faculty's competing priorities and limited time, we contend that it is important for institutions, and specifically faculty developers, to analyze how much time faculty are spending in online faculty development activities as well as which parts are taking the most (or least) time.

User behavior

One way to investigate faculty participation in online faculty development is to analyze the analytics in the learning management system (e.g., Vu, Cao, & Cepero, 2014). Two common ways to analyze participation using analytics are through looking at hits per page or time logged into the system. In 2014, the average page hits per user in the course design seminar ranged from 70 to 257 hits each week or an average of 132.3 hits across all 6 weeks (see Table 3 and 4). When looking at the clicks per week, the first three weeks had more page views than the last three weeks (of the design phase). This change is likely due to the content during these weeks; weeks 1-3 focused on designing the "skeleton" of an online course, whereas weeks 4-6 focused on preparing faculty to develop the course site by outlining corresponding learning activities and assessments. The change in page views, however, simply could be the result of time; in other words, faculty might simply lose some interest after three weeks into an online faculty development seminar, or it could be the result of what else is going on at that time in the semester. Additional research needs to be conducted to uncover what might be causing this change in page views. Future research like this could help inform faculty developers of the optimal length of an online faculty development seminar given the standard ebb and flow (and therefore faculty workload) each semester.

The course design seminar is designed to take faculty 8 hours each week to complete. But this seminar is often many faculties' first experience in an online learning environment. Even though the time learners spend in any online course is bound to fluctuate some from week-to-week, we strive for these seminars to follow as accurately as possible the time estimates we make to faculty when they begin. Therefore, we regularly check the time faculty report they spend in these seminars with the time they actually spent logged into Blackboard and then make changes when needed. For instance, during the spring of 2014, faculty in the course design seminar reported spending 7 to 10 hours a week on the seminar. As a result, we simplified the seminar a little by removing some assignments. Then in the following summer and fall, we were happy to find that faculty reported spending an average of 5 to 9 hours a week in the seminar.

We regularly conduct the same type of analysis on faculty participation in the teaching online seminar (eTOS) as we do with the course design seminars. Faculty are expected to spend 6 hours each week in the 6-week teaching online seminar. Page hits for this seminar ranged between 112 and 268 each week. Similar to the frequency of page hits in the course design seminar, the average page hits decreased over time. However, whereas time logged into Blackboard fluctuated week-to-week in the course design seminar, time logged in the teaching online seminar gradually decreased over time (see Table 4). As before, more data is needed to understand this gradual decline over time. While findings like these can leave one with more questions than answers, analysis like this needs to become a regular practice in online education and especially in online faculty development where faculty continue to try to do more with less.

Table 3
Comparison of Average Page Hits to Reports of Average Time Spent

	eCD2S		eTOS	
	Average Hits	Average Time	Average Hits	Average Time
Spring 14	158.5	8.5	154.3	8.8
Summer 14	107.8	6.9	164.4	9.0
Fall 14	131.2	6.3	168.2	7.6

Table 4
Average Page Views and Time on Task

	Course Design Seminar (eCD2S) Average Views and Time					
	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6
Spring	199 (8.8hrs)	173 (7.3hrs)	195.7 (9.5hrs)	116.1 (8.6hrs)	127.1 (9.5hrs)	140 (7.4hrs)
Summer	181.5 (9.3hrs)	83.1 (6.1hrs)	132 (7.6hrs)	86.1 (6.0hrs)	93.2 (7.1hrs)	70.7 (5.0hrs)
Fall	257.4 (6.9hrs)	112.6 (5.0hrs)	106.5 (6.8hrs)	107.5 (5.9hrs)	94.5 (7.0hrs)	108.6 (6.2hrs)

	Teaching Online Seminar (eTOS) Average Views and Time					
	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6
Spring	268.8 (9.8hrs)	259.6 (9.2hrs)	192.28 (9.0hrs)	150.37 (9.1hrs)	132.96 (7.9hrs)	154.28 (7.9hrs)
Summer	234.3 (9.9hrs)	178.7 (9.3hrs)	169.4 (9.9hrs)	172 (8.5hrs)	140.8 (8.3hrs)	164.4(8.2hrs)
Fall	254.12 (10.6hrs)	203.4 (7.2hrs)	134 (8.1hrs)	165.9 (6.2hrs)	139.8 (6.5hrs)	112.2 (6.8hrs)

Discussion participation

Another way to investigate faculty participation in online faculty development is to analyze online discussions (Blignaut & Trollip, 2003). Analyzing the online discussions revealed that the facilitators in the course design seminar (who are instructional designers) posted 25% - 50% of the discussion posts, while the facilitators in the teaching online seminar (who are full-time faculty) contributed only 8%–30% of the posts (see Table 5). This finding was not a surprise; the teaching online seminar is designed intentionally to be less facilitator-driven than the course design seminar because the it focuses more on online teaching strategies (Carter-Cram & Black, 2014), which among other things, allows faculty to leverage their prior experience as educators more than designing online courses from scratch. Further, an intentional goal in the teaching online seminar is to help create a community of learners who might assist faculty throughout their online teaching careers.

Table 5
Facilitator Post Ratio of Seminars

	Course Design Seminar (eCD2S)					
	Module 1	Module 2	Module 3	Module 4	Module 5	Module 6
Spring	26%	21%	36%	31%	33%	53%
Summer	23%	30%	32%	22%	39%	44%
Fall	19%	23%	20%	19%	17%	19%
	Teaching Online Seminar (eTOS)					
	Module 1	Module 2	Module 3	Module 4	Module 5	Module 6
Spring	21%	15%	14%	17%	8%	14%
Summer	28%	23%	27%	19%	25%	23%
Fall	29%	23%	25%	23%	24%	26%

Our faculty new to online education tend to recognize the importance of student-to-student interaction and, specifically, the role online discussions play in predominantly asynchronous online courses, but they struggle with managing how to facilitate and engage in online discussions properly and how they use these discussions in their own online courses. For instance, some faculty reported how useful the discussions were in these online seminars because they *allowed insight into what others were dealing with. One respondent said,* "*I found many were similar to my own issues or ideas*" (eCD2S Spring, module survey), whereas another said that the discussions "*allowed me to ask questions, explore my hypotheses and inclinations and see what others have to say,* and therefore permitted the faculty member to "*enjoy the collaborative nature of the discussions as well"* (eTOS Spring, module survey).

At the same time, others reported that they *"struggled to find meaning in the discussion boards"* (eCD2S Fall, exit survey), and *"think this is in part because I am an independent learner by nature and DB, in general, aren't really my thing (I personally don't feel mandatory DB posts aid in my learning of the material, but I do appreciate their role in an online course environment)*" (eTOS Summer, module survey).

Analyzing online discussions and specific learners' perceptions of how their instructors use online discussions is important to the success of any discussion-based online course (Ramsay, Aman, & Pursel, 2014). Frankly, it takes time and commitment for faculty and students to become comfortable and literate with the unique style of reading and writing required to participate effectively in electronically mediated environments (Dunlap, Bose, Lowenthal, York, Atkinson, & Murtagh, 2016; Vaughan & Garrison, 2005).

Faculty Perception

The success of an online faculty development program in an institution where faculty are not forced to complete professional development depends heavily on faculty having a positive experience and finding value in the program. Therefore, from the inception of eQIP, we wanted to know whether faculty were satisfied, feeling a sense of community, and learning new skills and dispositions in the program.

Faculty Satisfaction

It is common practice to administer an end-of-course evaluation at the end of a workshop or course to determine learner satisfaction (Lowenthal, Bauer, & Chen, 2015). However, the problem with waiting until the end of the course or workshop to get learner feedback is that once the course is over, it is too late to make any changes and possibly improve the experience for the cohort learners (Dobrovolny & Lowenthal, 2011). Therefore, after each module in the course design and teaching online seminars, faculty are asked to rate their satisfaction on a 5-point Likert scale (i.e., *Please rate your level of satisfaction with this module/seminar*).

In 2014, faculty reported overall that they were satisfied with the course design and teaching online seminars (see Table 6). Faculty responses in the course design seminar ranged from 3.55 to 4.78 and from 3.67 to 4.73 in the teaching online seminar. Collecting data after each module enabled us to assess the high and low points of each seminar as well as overall trends. For instance, a closer look at the teaching online seminar weekly survey results revealed that faculty satisfaction gradually increased. Also, while Week 3 in the course design seminar in the spring was low regarding faculty satisfaction, in other semesters, faculty reported being satisfied with this week, which suggests that there is some value in not making quick changes to a workshop simply because of one bad week. Instead, faculty developers should try to collect data over time and make informed decisions based on more than one cohort of learners whenever possible.

Table 6
Faculty Satisfaction with the Course Design and Teaching Online Seminars

	Course Design Seminar (eCD2S)							
	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Mid	Exit
Spring	4.00	4.01	3.83	4.38	4.40	4.43	4.50	3.55
Summer	4.00	4.62	4.43	3.71	4.29	4.00	4.43	4.25
Fall	4.75	4.40	4.20	4.71	4.67	4.40	4.57	4.50
	Teaching Online Seminar (eTOS)							
	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Mid	Exit
Spring	3.91	4.18	4.36	4.35	4.40	4.73	--	4.00
Summer	4.07	3.94	4.19	4.71	4.25	4.28	--	4.21
Fall	3.67	4.56	4.27	4.80	4.50	4.54	--	4.34

While we found it helpful to get a sense of a learner's overall satisfaction, doing so ultimately did not help identify specifically what in a course or workshop faculty are satisfied or dissatisfied about. Therefore, we also looked to specific responses from midterm and exit surveys to get a better sense of what satisfied our learners. In the course design seminar, faculty reported extremely high satisfaction with the instructional designers (who were the facilitators). However, not all the faculty were satisfied with the layout of the course. But perhaps most interesting is that when faculty were asked whether the seminar was meeting their learning expectations, the average satisfaction score decreased from the midterm to the end of the seminar, which is when they started working one-on-one with their instructional designers to develop their own courses. This decline could be due to the change in the structure of the seminar, from one that involved interacting with fellow faculty during the first phase to one where the faculty were mostly working alone (with the guidance of an instructional designer) on their own courses. It could also be that faculty expected the instructional designers or the development experience overall to be something different than it was. Like all online learning experiences, managing the expectations of faculty taking part in future seminars might help increase their perceived learning during the second half of the seminar.

Table 7
Faculty Perceptions of the Course Design Seminar Spring, Summer, and Fall 2014

Items	Midterm (*N*=33)		End (*N*=25)	
	Mean	SD	Mean	SD
Content was scholarly-researched and sound.	4.45	0.91	4.29	0.83
Content was effective in preparing you to design your online course.	4.36	1.06	4.29	0.73
Course layout was easy to follow.	3.88	1.22	4.00	1.11
Facilitators demonstrated competence in online education.	4.75	0.44	5	0
Facilitators acted respectfully toward participants.	4.97	0.18	5	0
Facilitators created frequent interaction and prompt feedback.	4.63	0.75	4.93	.27
Seminar was meeting your learning expectations.	4.47	0.84	4.04	1.18
Seminar was a professional development community of online instructors.	4.09	1.06	4.04	0.94

Due to the short duration and already high workload in the teaching online seminar, we decided to ask the previous questions only to faculty in the eTOS exit survey. While we lose some of the value of the week-to-week data, we recognize that there can come a time where learners are surveyed too often. At the end of the teaching online seminar, faculty reported that they were *satisfied* with every item but the course layout (see Table 8). Both online seminars are conducted in Blackboard; Blackboard offers faculty the ability to create as many folders as they like. While the folders were created with what appeared to be a clear structure and numbering system to guide the learners, a few learners reported that it was not as clear as we suspected. This data, coupled with some experiments using a screen reader on the seminar, prompted the eQIP team to restructure the layout of the teaching online seminar by flattening the entire structure and therefore having fewer nested folders. We will continue to collect data to see if this alternation changes faculty perceptions of the layout of the seminars over time.

Table 8
Faculty Perceptions of the Teaching Online Seminar During 2014

Items	End (*N*=56) Mean	SD
Content was scholarly-researched and sound.	4.38	1.26
Content was effective in preparing you to teach your online course.	4.13	1.18
Course layout was easy to follow.	3.77	1.04
Facilitators demonstrated competence in online education.	4.49	0.63
Facilitators acted respectfully toward participants.	4.73	0.44
Facilitators created frequent interaction and prompt feedback.	4.05	0.91
Seminar was meeting your learning expectations.	4.20	0.80
Seminar was a professional development community of online instructors.	4.38	0.73

Perception of faculty learning community

Research has demonstrated multiple benefits of establishing a community of learners in online courses (Abdelmalak, 2015; Boettcher & Conard, 2004; Dzubinski, 2014). Therefore, when developing the eQIP, the eQIP team wanted to find a way to create an online faculty learning community to both model the power of online community as well as to engage faculty further. Community-building activities—influenced in part by the theory and practices of the Community of Inquiry (CoI) (see Bond, 2011; Dunlap & Lowenthal, 2014; Garrison, Anderson, & Archer, 2000; McElrath, & McDowell, 2008; Murdock, & Williams, 2011; Pickett, 2010; Stavredes, 2011; Vesely, Bloom, & Sherlock, 2007; Wilcoxon, 2011)—were intentionally incorporated into each of these seminars. For instance, a weekly "Dear Jen, Kim, and Patrick" letter in the teaching online seminar (see Figure 1) was posted regularly throughout the teaching online seminar to model and encourage self-disclosure and a sense of shared connection while also leveraging faculty participants' prior experiences teaching.

Community is difficult to measure but it begins and ends with interacting with others. Therefore, we created weekly social network diagrams using SNAPP (Social Networks Adapting Pedagogical Practice, http://www.snappvis.org) to better understand participants' interactions in the seminar. While student-to-student or student-to-teacher interactions are not unequivocal proof of a learning community, they are a necessary building block to establishing and maintaining a community of online learners. Visualizing these interactions (by SNAPP) enabled the facilitators to diagnose and monitor participation overtime (see also Dawson, Bakaharia, & Heathcote, 2010; McCormick, 2013). These diagrams revealed, for example, that in the teaching online seminar, interactions were facilitator-centric (see dark gray node in Figure 7) during the first week, but changed to participant-centered over time and eventually began to form what appears (when looking at the interactions) as a community of learners (see Figure 8). Supporting our theory, a participant in the teaching online seminar noted in the spring of 2014 that "not only do I feel happy and excited about my teaching as I exit eTOS, I feel a connection to many more professionals at BSU" (eTOS Spring, exit survey).

Dear Jen, Kim, and Patrick:

I took your suggestion and had my online students fill out a midterm survey about the course. When I looked at the survey results, I felt very discouraged. Students had some good things to say about the course, but they also had lots of complaints about the workload, the course organization, and the assignments. They even said they felt like the discussion board prompts were boring and didn't inspire meaningful conversation. What should I do now? Should I try to change everything they don't like about the course? Or should I just stick with what I have planned and look into making changes after the course is done?"

Discouraged in Donnelly

Figure 1. A Sample "Dear Jen, Kim, and Patrick" Letter.

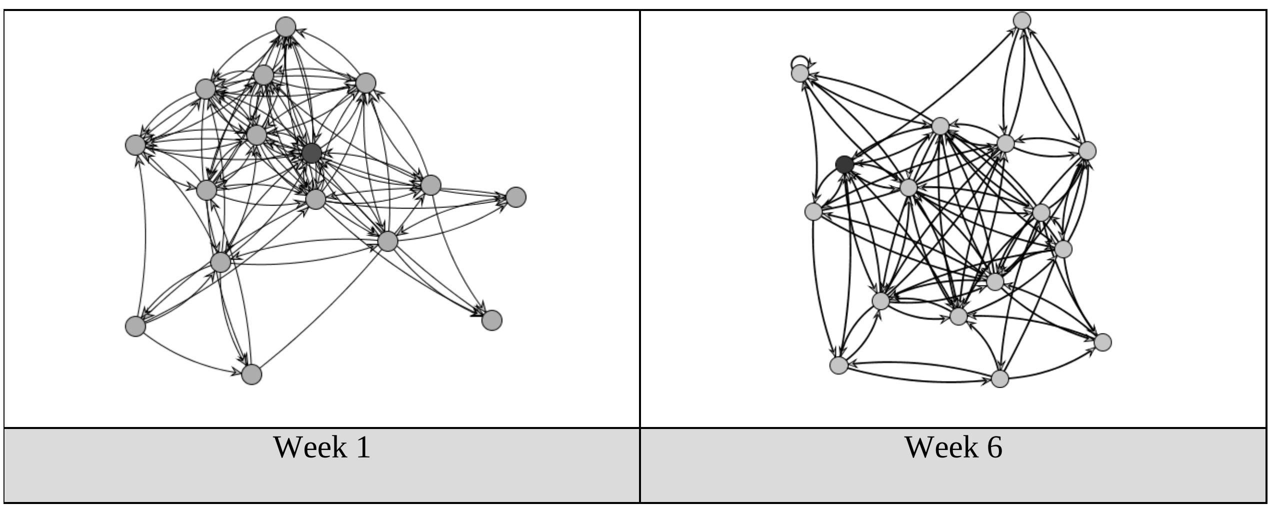


Figure 2. The Social Network Diagram of Interaction in Teaching Online Seminar (Summer)

Establishing a learning community takes time (Dzubinski, 2014); unfortunately, though, it does not always work as planned (Olson & McCracken, 2015). For instance, in the course design seminar, some faculty complained about using the discussion board as a place to "show and tell" their design documents—a strategy encouraged by some in the field (see Lowenthal & Thomas, 2010a). However, despite resistance from some faculty, other faculty, as evident in the quotes below, clearly developed a sense of community over the seminars (largely through their interactions in the discussion forums) and in turn considered adding some community building activities in their own online courses:

- “The sense of community is very important to me and I believe I gained some ground on how to promote them in my class” (eCD2S Spring, module survey).
- “The idea of ‘community’ in an online course is something that I think most of us will have to learn and adjust to in a ‘trial by fire’ sort of way” (eCD2S Spring, module survey).
- “The Discussion Board was a really helpful way to build community and share ideas about learning objectives” (eCD2S Fall, module survey).
- “I read through how the others were approaching this task made me flash on [course designs] I hadn't thought of to try myself “(eCD2S Fall, module survey).

Changes in faculty attitudes and disposition

The eQIP is designed to teach faculty how to design high-quality, fully online courses from scratch as well as how to teach effectively in this new environment. Hence, we were interested in looking at faculty attitudes, knowledge, and dispositions at the end of each seminar, as well as how they might have changed over time while participating in the online seminars. For the course design seminar, we created a survey with a series of quantitative, Likert scale questions as well as qualitative, open-ended questions to measure faculty attitudes and dispositions. The survey was administered three times in the course design seminar (i.e., beginning, middle, and end of the seminar) and two times in the teaching online seminar (i.e., the beginning and end). The results from these surveys are summarized in Table 9 and discussed along with open-ended feedback in the following sections.

Table 9

Faculty Confidence Designing and Teaching Online Courses During eQIP Seminars in 2014

Items (1=*Strongly Disagree*; 5=*Strongly Agree*)		Begin (eCD2S N=31; eTOS N=70)	Midterm (eCD2S N=33; eTOS=n/a)	Final (eCD2S N=25; eTOS N=56)
1. You are confident in identifying and describing target online student characteristics.	eCD2S	3.51	3.99	3.96
	eTOS	3.59	--	4.28
2. You are confident in analyzing differences between the online learning environment and the on-campus classroom.	eCD2S	3.56	4.04	4.18
	eTOS	3.78	--	4.35
3. You are confident in designing online assessment linked to course learning objectives and activities.	eCD2S	3.39	4.16	4.39
	eTOS	3.93	--	4.28
4. You anticipate the course evaluations for your online course compared to traditional courses will be higher.	eCD2S	2.64	--	2.75
	eTOS	4.82	--	3.30
5. You expect the workload to design and develop an online course compared to traditional one will be higher.	eCD2S	3.83	4.36	4.32
	eTOS	4.10	--	4.12
6. You expect motivating students in an online course compared to traditional one will be more difficult.	eCD2S	3.27	3.81	3.89
	eTOS	3.28	--	3.32

Faculty confidence

We were interested in how faculty confidence in designing and teaching online changed over time. We hoped that the eQIP would transform participants' view of online learning from content-oriented toward student-oriented (Vries, Grift, &, Jansen, 2014) as well as leave them with a more realistic expectation of online teaching. Table 9 shows that faculty confidence in identifying and describing target online student characteristics, analyzing differences between the online learning environment and the on-campus classroom, and designing online assessments linked to course learning objectives and activities increased over time in both seminars. Interestingly, participants in these online seminars expressed different attitudes in questions 4, 5, and 6. For instance, after faculty participated in the teaching online seminar, they anticipated lower course evaluation scores when teaching online courses; a decrease of -1.52 from the beginning to the end of the seminar (see Question 4 in Table 9). This result could be due to the fact that faculty participants read the assigned textbook about online teaching, realized the demands of highly interactive online courses in action, and read research on teaching evaluations decreasing when teaching online compared to teaching face-to-face courses (see Lowenthal et al., 2015). Then when asked if they expected the workload of designing and teaching to be higher in online courses than in face-to-face courses, faculty responses increased over the duration of both seminars—meaning that while they expected the workload to be higher, this expectation grew by the end of both seminars. However, faculty responses regarding expected workload in the teaching online seminar barely increased from the beginning to the end of the seminar, whereas faculty response in the course design seminar increased 0.49 by the end of the seminar (see Question 5 in Table 9). Finally, when asked about how difficult it will be to motivate online students, faculty responses in the teaching seminar decreased some (suggesting that they might be gaining confidence in motivating online learners), whereas faculty responses in the course design seminar increased over time (+0.62). While it is normal to lose confidence once one learns more about something (e.g., teacher preparation: students often report being more unsure about teaching after they have student taught), a goal of the eQIP is to develop our learners' confidence. Therefore, we will continue to find ways to improve these aspects of our seminars over time.

Perceived changes in online course development

Teaching is an individual activity; as such, teachers at all levels of education often feel isolated (McQuiggan, 2011). Some faculty thrive on this autonomy, while others do not. Instructional design, on the other hand (at least in large departments), is often a collaborative endeavor; therefore, we intentionally developed the course design seminar to help faculty see course design not as an individual intellectual activity (as they naturally might) but as an open collaborative activity. Open-ended feedback revealed how faculty sensed their growth. For instance, one faculty member reported how the seminar activities provided an opportunity to "think about the different scales (course vs. unit/module) of objectives. It helps to be forced to understand and articulate how they are connected with one another" (eCD2S Fall, module survey). Another reported how figuring "out how I would write directions for an online activity was a challenge and eye opener" (eCD2S Fall, module survey). Faculty also reported how these seminars helped them rethink how they teach other courses—a goal for nearly any faculty developers. For instance, one faculty member reported how she or he now saw the potential to "redesign course assessment/evaluation in my other courses, and I'm already telling colleagues outside of this seminar about it. Good stuff!" (eCD2S Fall, module survey).

Perceived changes in online teaching

Faculty who have never taught online before often hold certain assumptions about teaching online. These can range from "teaching online entails more work than teaching face-to-face" to other opinions such as "students cheat more online." The teaching online seminar is designed not only to teach faculty how to teach online but to challenge common misconceptions faculty might hold. Participants reported changes in their perception regarding the ability to effectively communicate online. For instance, one participant found "orientation videos to be personal, informative, and a great way to begin an online course," and that he or she "enjoyed watching the samples and getting ideas for [her] own" (eTOS Spring, module survey). Others discovered some of the affordances of a LMS. For instance, one participant pointed out how "the possibilities for reflection in a safe and confidential modality is important. It's different than the discussion board because the public aspect of that really forces you to be on your guard" (eTOS Spring, module survey). Finally, others specifically identified how their perceptions of online teaching changed but remained realistic. For instance, one participant stated that "I know teaching my first online class will take a lot of time; I appreciate the reminders and [facilitator]'s comment to not freak me out" (eTOS Summer, module survey). Perceived changes like these are something we hope to see (e.g. Scagnoli, Buki, & Johnson, 2009) from participants as a result of taking part in the eQIP.

Faculty Concerns

Exit survey results suggest that overall faculty considered the eQIP a valuable faculty development experience. However, our inquiry also uncovered concerns faculty had with the eQIP as well as online teaching in general.

Individualized vs. grouped training

Boise State, like many institutions, is interested in creating new online courses and programs to meet student demand as efficiently and inexpensively as possible. The eQIP was designed to train faculty to be able not only to design online courses (that they or others would teach) but to learn how to teach any online course (whether one they designed or not). Some faculty, though, had concerns with this model and felt that it was too standardized. Regarding the course design seminar, some faculty reported concerns with standardized course development. One faculty member explained that a standardized model of course development "takes longer to work within someone else's templates than your own" (eCD2S Summer, midterm survey). Some faculty explained that the eQIP-grouped "standardized" training approach does not work well for faculty, many of which work better individually (e.g., with a dedicated instructional designer) than in groups. One faculty participant captured this sentiment when he or she suggested that eCampus should embrace "the idea that faculty learners are very different from graduate students and that a personalized, one-on-one training defined by faculty priorities and focusing on personalized interaction would be much more effective" (eTOS Spring, exit survey). While we recognize that some faculty would prefer a one-on-one approach, there are not enough resources to accommodate such an approach. We are, however, continually reminding faculty that they are the subject matter experts, that these are their courses, and that they can change the course template when needed.

Novice vs. veteran faculty participants

The target audiences of the eQIP are novice online instructors who have little or no prior experience designing/developing online courses and/or teaching online at Boise State. However, more and more faculty with prior experience as either online students or online instructors at prior institutions sign up for the eQIP; they sign up because their specific department expects all new online faculty to go through the eQIP. For instance, one faculty participant reported:

> *I am beginning to think I am in the wrong course.... I have been trained in course development at the graduate level and have participated in accreditation self-studies. What I really needed was the skills in the technology of online learning, but you can't take that class until you are teaching online for Boise and you can't teach online for Boise until you take this course. (eCD2S Spring, module survey).*

Not all faculty who take part in the eQIP with prior online teaching experience are upset. A number of them are excited to learn new ways of designing and/or teaching online courses. But the majority of these faculty do want an advanced version of the eQIP as captured in the following quotes:

> *I would encourage you to consider breaking eTOS into two seminars: (1) with standard topics for instructors who have never taught online before, much like this seminar; (2) built around grouping experienced faculty with shared priorities (collect a ranked set of priorities) into small task groups to address those priorities, allowing for extensive one-on-one interaction (for example one hour per week via Google Hangouts) with instructional design consultants who have expertise in that area (eTOS Spring, exit survey).*

> *Maybe there should be a third option of course development where teachers have taught online before, but want to radically alter the design and look of their online course* (eCD2S Spring, module survey)**.**

> *One idea here might be to create a diversion plan - a set of assignments for faculty who have never taught online, and a different role set for faculty who have - so that the latter are being pushed both to review what they have (rather than plan it) and also to offer reflections on what works and not for those who have not. In other words, faculty clearly have very varied backgrounds and one size does not fit all* (eTOS Spring, module survey).

Feedback like this argues for a more flexible and individualized and/or customized version of the eQIP (eCD2S Review Meeting, August 8, 2014). This complaint is a common problem with many online courses, not just online faculty development. The eQIP team continues to consider ways to change the eQIP by incorporating individualized learning designs such as worked examples (eCD2S Review Meeting, June 4, 2014), quest-based learning and digital badges (Haskell, 2013) or competency-based learning (Sally & Louis, 2014).

Technology as prerequisite vs. just-in-time technical training

Faculty are subject-matter-experts—not necessarily technology experts (Kukulska-Hulme, 2012). At Boise State a tension exists, as it does at many other universities, about how many technology skills faculty should be expected to have (or develop) to teach online and where they should go to get help when they need it. The eQIP currently does not focus much on developing basic technical skills in faculty (e.g., uploading a document to Blackboard). Instead, the eCampus takes the stance that all faculty who wish to teach online (and participate in the eQIP) should have basic technical competences before beginning the eQIP. However, in practice, faculty do not always pay attention to this prerequisite; neither do they understand what basic technical skills are in the first place. Thus, we find a reoccurring concern with some faculty is about whether being proficient with basic technology should be a prerequisite or whether we should provide just-in-time technical training. For instance, one faculty participant suggested that there needs "to be better support during the heavy tech portions of the class." Along the same lines, another person pointed out the need for a "specific contact person [to] be identified to help troubleshoot issues" (eTOS Fall, exit survey). Aware of this continuing concern, the eQIP facilitators have tried to provide more-than-expected just-in-time technical training (such as tip sheets and tutorial videos) to compensate for the gaps in faculty technical skills. And while some faculty have praised this flexibility and support—for example, one reported how he or she "enjoyed specific skill set instruction" (Camtasia, Google Hangouts, tech skills, etc.) (eTOS Summer, exit survey)—the tension still remains about whether it is fair to expect faculty to have basic technical skills before teaching online for the first time.

Experiencing learning as an online student

Each of the eQIP seminars is conducted fully online, which allows faculty to experience firsthand what it is like to be an online student and to experience some of the common complications students face (e. g., technical difficulties with Blackboard LMS, using e-textbooks and multimedia, academic honesty, accessibility, copyright issues, student interaction, and time management [see Mason et al., 2010]). Faculty regularly report that this hidden curriculum is a valuable experience. For instance, faculty often recognize how a busy life can get in the way of completing an online course. One faculty person said, "I have more sympathy and empathy with my online students now. I get to this class after family and work responsibilities; I am trying to do enough to get by" (eTOS Spring, exit survey). Another group found themselves questioning whether they really should have deadlines over the weekends if they do not want to work over the weekends. For instance, one faculty participant in the eCD2S commented: "I am curious as to why there are Saturday deadlines and I have been considering whether I want those in my course" (eCD2S Fall, module survey). Additionally, faculty often experience technical problems for the first time from a student's perspective as captured in this quote: "I know this is more of a Bb issue but now I understand when my students say – 'Oh I didn't see that.' " (eTOS Fall, module survey).

This authentic experience, though, takes time. Some faculty reported that this authentic experience could be done in less time or with less work or that their participation in the seminar should not be graded. This concern about being treated too much like an online student was addressed in ways in the course design seminar by explaining to faculty that by experiencing discussions, assignments, quizzes, and group tools from a student's view, they would be in a better position to resolve students' issues during teaching. The eTOS facilitators, however, tried to make participants' comments and concerns teachable moments. For instance, after receiving a complaint asking "Why am I 'losing points' on a self-assessment for not doing an optional assignment? That doesn't seem fair!" (eTOS Spring, module survey), facilitators apologized for the error, which was simply an oversight, and explained that errors like this sometimes happen when adjustments are made to a course from semester to semester and that faculty should strive to double check each course when making even minor adjustments (eTOS Review Meeting, October 22, 2014). Such modeling was effective because, as one faculty member reflected in his journal entry: “Don't be afraid to admit a mistake…… I'm somewhat dyslexic and sometimes things just get confused. I tell my students that if you find something that seems odd, just tell me because it probably is and I'll fix it” (eTOS Spring, journal activity).

Workload and time competition

Time is a sensitive issue with faculty (Lowenthal, Wray, Bates, Switzer, & Stevens, 2013). One of the challenges the eQIP faces is competing with limited faculty time. For instance, even though the teaching online seminar only takes 6 hours a week for 6 weeks (the shortest of the seminars), this is still 6 hours a week that faculty have to take away from teaching, scholarship, or service. Despite our best efforts, some participants reported that tasks in seminars were “busy work.” For instance, one faculty reported in the teaching online seminar “I simply, realistically, literally(!) did not have enough time to complete while at the same time maintain my teaching, research, and personal life!!! (Wheww!)” (eTOS Summer, exit survey). This tension gets even worse with the course design seminar that takes even more time. Boise State faculty, like online faculty elsewhere, regularly report that designing an online course takes more time than preparing for a face-to-face course (see also Bento, 2011). However, some qualitative data suggest that many of these participants begin to change their perception of “busy work” by the end of the seminars. A common response from faculty looks like this: “The material in this course was overwhelming many times. At the end I can see that the structure was appropriate” (eCD2S Spring, exit survey). This change in perception is even more common in the online seminars because early in the course design seminar faculty are required to complete multiple worksheets as part of the instructional design process that they are not used to completing when designing face-to-face courses. It is usually not until the fifth or sixth week that faculty report that they “feel like things are starting to come together in a way that will allow me to best utilize my time with the design consultant” (eCD2S Spring, module survey). We have found that faculty who sign up for the course design seminar to design a course that they are signed up to teach the following semester do better than those who have a few semesters before they actually teach the course they are developing. We also began adding time estimates for each activity in the LMS for both seminars to help faculty manage their own time. However, providing accurate time estimations for all learners is nearly impossible; therefore, some faculty who worked slower than others actually were bothered when time estimates did not match the time it took them to complete a project.

Conclusion

A successful online faculty development program must include pedagogical support, technology support, and design and development support (Baran & Correia, 2014) that overcome obstacles about time, expertise, and motivation of faculty (Henning, 2012). This study addressed the implementation of an online faculty development program at a mid-size, metropolitan research university. We investigated faculty participation and perceptions by user pattern, discussion participation, satisfaction, learning community, attitude and disposition, as well as their concerns. The results of our inquiry suggest that while there are still ways we can improve each seminar, overall, the eQIP is preparing Boise State faculty to design, develop, and teach online courses. In 2014, in fact, 51 online courses were developed as a result of the eQIP, 100 faculty members completed both seminars, and 38 courses went through QMPR. With the faculty intent to change teaching practice, eQIP impacts both online and face-to-face teaching. In the words of one faculty participant:

> This course is a worthy endeavor for both online and face to face teachers. I will use the materials that I made for the online course in face-to-face classes as well. I plan to incorporate online components into my face-to-face classes in the future (eCD2S Spring, exit survey).

Future improvement based on lessons learned

Online faculty development is a rapidly changing field and its offerings require continuous evolutions (Meyer & Murrell, 2014). Based on the findings of this study, the eCampus Center should continue improving the seminars to be as engaging as possible. Our continued data collection efforts reveal the following four potential areas of improvement.

1. Provide individualized seminar processes

Several faculty comments indicated that having all participants go through the faculty development seminars at the same speed failed to serve faculty from various disciplinary backgrounds. A more directional and customized faculty development process may better meet faculty's needs. For example, a discipline-based repository for online course will be helpful and flexible for faculty who are not beginners in online education. A faculty development program like the eQIP does not have to be 100% online to be authentic, relevant, motivating, engaging, and useful for faculty (Sorinola, Thistlethwaite, Davies, & Peile, 2015).

2. Empower faculty ownership of the seminars

Experienced faculty should be identified as soon as possible and encouraged to become mentors to novices in the seminar (Xu & Morris, 2007). Seminar topics should be chosen from faculty's teaching experience to promote faculty ownership of the seminars.

3. Condense course content

In spring 2014, the eTOS team invited a new facilitator to lead the seminar. His external perspective helped reorganize and remove any redundant information. This significantly lowered faculty's workload in the teaching online seminar.

4. Provide tenure-related incentives

The eCampus Center should negotiate at the University level to provide faculty supports and recognitions in online teaching. For example, to include the eQIP participation as one element in faculty promotion, allow online course developers to count development time toward semester teaching loads, and create online teaching fellow programs to recognize faculty contribution to online education.

Limitations

The purpose of this study was to describe one institution's effort to develop high-quality online courses and high-quality online teachers. Therefore, as a case study within a highly-individualized context, the results of this study offer insights to peer institutions with similar backgrounds, but may not be generalizable to a larger population (Stake, 1995; Yin, 2009). But even as a case study, the results could have been strengthened by including additional data such as a faculty focus group; a design of pre- and post-seminar faculty survey, participants' retention rates, and student outcomes in the high-quality courses developed, and faculty end-of-course evaluations teaching these courses. In the future, concerted and coordinated research efforts across peer institutions may advance the study of faculty development beyond the level of sharing best practices in the field (Meyer, 2013).

References

Abdelmalak, M. M. M. (2015). Web 2.0 technologies and building online learning communities: Students' perspectives. *Online Learning Journal, 19,* 1-20.

Bacow, L. S., Bowen, W. G., Guthrie, K. M., Lack, K. A., & Long, M. P. (2012). *Barriers to adoption of online learning systems in US higher education*. New York, NY: ITHA. Retrieved from http://sites.tufts.edu/strategicplan/files/2012/10/barriers-to-adoption-of-online-learning-systems-in-us-higher-education1.pdf

Bailey, C. J., & Card, K. A. (2009). Effective pedagogical practices for online teaching: Perception of experienced instructors. *Internet and Higher Education, 12*, 152-155.

Baran, E., & Correia, A. (2014). A professional development framework for online teaching. *TechTrends, 58*(5), 96-102.

Baran, E., Correia, A., & Thompson, A. (2011). Transforming online teaching practice: Critical analysis of the literature on the roles and competencies of online teachers. *Distance Education, 32,* 421-439. doi:10.1080/01587919.2011.610293

Benton, C. J. (2011). Asynchronous learning and faculty development: Evolving college-level online instruction and empowered learning. *International Journal of Information & Communication Technology Education, 7*(1), 89-96. doi:10.4018/jicte.2011010109

Blignaut, A., & Trollip, S. R. (2003). Measuring faculty participation in asynchronous discussion forums. *Journal of Education for Business, 78,* 347-353.

Boettcher, J., & Conrad, R. (2004). *Faculty guide for moving teaching and learning to the Web* (2nd ed.). Phoenix, AZ: League for innovation in the Community College.

Bond, T. (Producer). (2011, August 17). *Community building: Putting it all together - Web 2.0.* [Video File] Retrieved from https://www.youtube.com/watch?v=dxoH0n875BM

Britto, M., Ford, C., & Wise, J., (2014). Three institutions, three approaches, one goal: Addressing quality assurance in online learning. *Journal of Asynchronous Learning Network, 17*, 11-24.

Brooks, C. F. (2010). Toward 'hybridised' faculty development for the twenty-first century: Blending online communities of practice and face-to-face meetings in instructional and professional support programmes. *Innovations in Education & Teaching International, 47*(3), 261-270. doi:10.1080/14703297.2010.498177

Carter-Cram, K. (2014). Game on! Teaching foreign language online. In M. Orleans (Ed.), *Cases on critical and qualitative perspectives in online higher education* (pp. 208-226). Hershey, PA: IGI Global. doi:10.4018/978-1-4666-5051-0.ch011

Carter-Cram, K., & Black, J. (2014, January). *Professional development for effective online teaching*. Poster presented at the 3rd Annual Great Ideas Symposium, Center for Teaching & Learning, Boise ID.

Chen, C. C., & Shaw, R. S. (2006). Online synchronous vs. asynchronous software training through the behavioral modeling approach: A longitudinal field experiment. *International Journal of Distance Education Technologies, 4*(4), 88-102.

Chen. (2014, March). *Supporting faculty toward high-quality online course design: Boise State eQIP Program.* Poster session presented at the conference of Diversity, Learning, and Student Success: Policy, Practice, and Privilege. Chicago, IL: Association of American Colleges and Universities.

Creswell, J. W. (2008). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Thousand Oaks, CA: Sage.

Creswell, J. W., & Plano, V. L. (2007). *Designing and conducting mixed methods research*. Thousand Oaks, CA: Sage.

Dabbagh, N., & Bannan-Ritland, B. (2005). *Online learning: Concepts, strategies, and application*. Upper Saddle River, NJ: Prentice Hall.

Dailey-Hebert, A., Mandernach, B. J., Donnelli-Sallee, E., & Norris, V. R. (2014). Expectations, motivations, and barriers to professional development: Perspectives from adjunct instructors teaching online. *Journal of Faculty Development, 28*, 67-82.

Dawson, S., Bakaharia, A., & Heathcote, E. (2010). *SNAPP: Realizing the affordances of real-time SNA within networked learning environments*. Paper presented at the 7th International Conference on Networked Learning, Aalborg, Denmark.

Dobrovolny, J. & Lowenthal, P. R. (2011). Using reflection surveys to improve teaching and learning. In P. R. Lowenthal, D. Thomas, A. Thai, B. Yuhnke, M. Edwards, & C. Gasell. (Eds.). *The CU Online Handbook, 2011* (pp. 97-113). Raleigh, NC: Lulu.

Dunlap, J. (2005). Workload reduction in online courses: Getting some shuteye. *Performance Improvement, 44*(5), 18-25.

Dunlap, J., Bose, D., Lowenthal, P. R., York, C. S., Atkinson, M., & Murtagh, J. (2016). What sunshine is to flowers: A literature review on the use of emoticons to support online learning. To appear in *Emotions, Design, Learning and Technology*. 163-182, Amsterdam, The Netherlands: Elsevier.

Dunlap, J.C., & Lowenthal, P.R. (2014). The power of presence: Our quest for the right mix of social presence in online courses. In A. A. Piña & A. P. Mizell (Eds.), *Real life distance education: Case studies in practice* (pp. 41-66). Greenwich, CT: Information Age.

Dzubinski, L. M. (2014). Teaching presence: Co-creating a multi-national online learning community in an asynchronous classroom. *Journal of Asynchronous Learning Network, 18,* 1-16.

Fein, A. D., & Logan, M. C. (2003). Preparing instructors for online instruction. *New Directions for Adult and Continuing Education, 100*, 45-55.

Garrison, D. R., Anderson, T., & Archer, W. (2000). Critical inquiry in a text-based environment: Computer conferencing in higher education model. *The Internet and Higher Education, 2*(2-3), 87-105.

Greene, J. C. (2007). *Mixed methods in social inquiry*. San Francisco: Jossey-Bass.

Haskell, C. (2013). *Understanding quest-based learning*. [white paper]. Retrieved from http://works.bepress.com/chris_haskell/21/

Henning, T. (2012). Writing professor as adult learner: An autoethnography of online professional development. *Journal of Asynchronous Learning Networks, 16*, 9-26.

Herman, J. H. (2012). Faculty development programs: The frequency and variety of professional development programs available to online instructors. *Journal of Asynchronous Learning Networks, 16,* 87-106.

Hislop, G. W., & Ellis, H. J. C. (2004). A study of faculty effort in online teaching. *Internet and Higher Education, 7*, 15-31.

Kukulska-Hulme, A. (2012). How should the higher education workforce adapt to advancements in technology for teaching and learning? *Internet and Higher Education, 15,* 247-254.

Lowenthal, Bauer, & Chen (2015). Student perceptions of online learning: An analysis of online course evaluations. *American Journal of Distance Education, 29*(2), 85-97.

Lowenthal, & Leech, N. (2009). Mixed research and online learning: Strategies for improvement. In T. T. Kidd (Ed.), *Online education and adult learning: New frontiers for teaching practices* (pp. 202-211). Hershey, PA: IGI Global.

Lowenthal, & Thomas, D. (2010a). Death to the digital Dropbox: Rethinking student privacy and public performance. *EDUCAUSE Quarterly*, *33*(3). Retrieved from http://www.educause.edu/ero/article/death-digital-dropbox-rethinking-student-privacy-and-public-performance

Lowenthal, & Thomas, D. (2010b). Digital campfires: Innovations in helping faculty explore the online learning wildness. *Journal of Online Learning and Teaching, 6,* 665-672.

Lowenthal, & White, J. W. (2009). Enterprise model. In P. Rogers, G. Berg, J. Boettcher, C. Howard, L. Justice, & K. Schenk (Eds.), *Encyclopedia of distance and online learning* (2nd ed., pp. 932-936). Hershey, PA: IGI Global.

Lowenthal, Wray, M. L., Bates, B., Switzer, T., & Stevens, E. (2013). Examining faculty motivation to participate in faculty development. *International Journal of University Teaching and Faculty Development, 3*, 149-164.

Lucas, C. J., & Murry, J. W. (2011). *New faculty: A practical guide for academic beginners.* New York, NY: Palgrave Macmillan.

Marek, K. (2009). Learning to teach online: Creating a culture of support for faculty. *Journal of Education for Library & Information Science, 50*(4), 275-292.

Maor, D. (2006). Using reflective diagrams in professional development with university lecturers: A developmental tool in online teaching. *Internet and Higher Education, 9,* 133-145.

MarylandOnline (2011). *Quality Matters™ rubric workbook for higher education (2011-2013 Edition).* Annapolis, MD: MarylandOnline.

Mason, J., Hickman, C., Dyer, A., Koproske, C., Fry, G., & Taha, M. (2010). *Engaging faculty in online education: Rightsizing incentives and optimizing support*. Washington, DC: Education Advisory Board.

McCormick, J. (2013). Visualizing Interaction: Pilot investigation of a discourse analytics tool for online discussion. *Bulletin of the IEEE Technical Committee on Learning Technology, 15*(2), 10-13.

McElrath, E., & McDowell, K. (2008). Pedagogical strategies for building community in graduate level distance education course. *Journal of Online Learning and Teaching, 4*(1), 117-127.

McQuiggan, C. A. (2011). *Preparing to teaching online as transformative faculty development.* Unpublished Doctoral Dissertation, Penn State University, State College, PA.

Merriam, S. B. (1998). *Qualitative research and case study applications in education* (2nd ed.). San Francisco, CA: Jossey-Bass.

Meyer, K. A. (2013). An analysis of the research of faculty development for online teaching and identification of new directions. *Journal of Asynchronous Learning Network, 17*, 93-112.

Meyer, K. A., & Murrell, V. S. (2014). A national study of training content and activities for faculty development for online teaching. *Journal of Asynchronous Learning Network, 18.* Retrieved from http://files.eric.ed.gov/fulltext/EJ1030527.pdf

Moskal, P., Thompson, K., & Futch, L. (2015). Enrollment, engagement, and satisfaction in the Blendkit faculty development open, online course. *Online Learning*, *19*(4). Retrieved from http://olj.onlinelearningconsortium.org/index.php/olj/article/view/555/173.

Murdock, J. L., & Williams, A. M. (2011). Creating an online learning community: Is it possible? *Innovative Higher Education*, *36*(5), 305-315.

Natriello, G. (2005). Modest changes, revolutionary possibilities: Distance learning and the future of education. *Teacher College Record, 110,* 105-138.

Olson, J. S., & McCracken, F. E. (2015). Is it worth the effort? The impact of incorporating synchronous lectures into an online course. *Online Learning Journal, 19,* 1-12.

Pickett, A. (Producer). (2010, August 16). *Social presence.* [Video File] Retrieved from https://www.youtube.com/watch?v=P_Fbo63Gi_k&list=PL936ED0CDBD9B4C06

Ragan, L. C. (1999). Good teaching is good teaching. An emerging set of guiding principles and practices for the design and development of distance education. *Cause/Effect*, *22*(1), 20-24. Ramsay, C. M., Aman, D. D., & Pursel, B. K. (2014). Blogging pragmatics and pedagogy: An adventure in faculty development. *Education and Information Technologies*, *19*, 425-440.

Sally, M. J., & Louis, S. (2014). Principles for developing competency-based education programs. *Change*, *46*(2), 12-19. DOI: 10.1080/00091383.2014.896705

Salmon, G. (2012). *E-moderating: The key to online teaching and learning*. New York, NY: Routledge.

Scagnoli, N. I., Buki, L. P., & Johnson, S. D. (2009). The influence of online teaching on face-to-face teaching practices. *Journal of Asynchronous Learning Networks, 13,* 115-128.

Schrum, L., Burbank, M. D., Engle, J., Chambers, J. A., & Glassett, K. F. (2005). Post-secondary educators' professional development: Investigation of an online approach to enhancing teaching and learning. *Internet and Higher Education, 8,* 279-289.

Sellani, R. J., & Harrington, W. (2002). Addressing administrator/faculty conflict in an academic online environment. *Internet and Higher Education, 5*, 131-145.

Sorinola, O. O., Thistlethwaite, J., Davies, D., & Peile, E. (2015). Faculty development for educators: A realist evaluation. *Advances in Health Sciences Education: Theory and Practice*, *20*, 385-401.

Stake, R. (1995). *The art of case study research*. London, UK: SAGE.

Stavredes, T. (2011). *Effective online teaching: Foundations and strategies for student success.* San Francisco, CA: Jossey-Bass.

Stone, M. T., & Perumean-Chaney, S. (2011). The benefits of online teaching for traditional classroom pedagogy: A case study for improving face-to-face instruction. *Journal of Online Learning and Teaching, 7*, 393-400.

Stufflebeam, D. L., Shinkfield, A. J. (2007). *Evaluation theory, models, and applications.* San Francisco, CA: Jossey-Bass.

Tipple, R. (2010). Effective leadership of online adjunct faculty. *Online Journal of Distance Learning Administration*, *13*(1). Retrieved from http://www.westga.edu/~distance/ojdla/spring131/tipple131.html

Vaughan, N., & Garrison, D. R. (2005). Creating cognitive presence in a blended faculty development community. Internet and Higher Education, 8, 1-12.

Vesely, P., Bloom, L., & Sherlock, J. (2007). Key elements of building online community: Comparing faculty and student perceptions. *Journal of Online Learning and Teaching, 3*(3), 234-246.

Vries, S., Grift, W. J. C. M., &, Jansen, E. P. W. A. (2014). How teachers' beliefs about learning and teaching relate to their continuing professional development. *Teacher and Teaching: Theory and Practice, 20,* 338-357.

Vu, P., Cao, V., & Cepero, J. (2014). Factors driving learner success in online professional development. *The International Review of Research in Open and Distance Learning, 15*(3), 120-139. Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/1714/2978

Wilcoxon, K. (2011, October 3). Building an online learning community. *Learning Solutions*. Retrieved from http://www.learningsolutionsmag.com/articles/761/building-an-online-learning-community

Wilson, B. G., Ludwig-Hardman, S., Thornam, C. L., & Dunlap, J. (2004). Bounded community: Designing and facilitating learning communities in formal courses. *The International Review of Research in Open and Distance Learning, 5*(3). Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/204/286

Wolf, E. A. (2013). *A study of motivation of online adjunct college undergraduate faculty* (Doctoral dissertation, University of Phoenix).

Xu, H., & Morris, L. (2007). Collaborative course development for online courses. *Innovative Higher Education, 32*(1), 35-47. doi:10.1007/s10755-006-9033-5

Yin, R. K. (2009). *Case study research: Design and methods* (4th ed.). London: SAGE.

SECTION II:
Integrating Accessibility into Online Higher Education

Reading Between the Lines: Accessing Information via *YouTube*'s Automatic Captioning

Chad Smith and Tamby Allman
Texas Woman's University

Samantha Crocker
Weatherford Regional Day School Program for the Deaf

Abstract

This study and discussion center upon the use of YouTube's automatic captioning feature with college-age adult readers. The study required 75 participants with college experience to view brief middle school science videos with automatic captioning on *YouTube* and answer comprehension questions based on material presented auditorily and/or through the automatic captions. Participants were divided into groups and presented with the captioned videos with or without sound. The videos, which all focused on the solar system, contained low and high instances of errors within the captions. The research found that comprehension of the automatic caption text varied significantly based on how the participants viewed the videos, with significantly more errors in comprehension for the group that viewed the high error video with automatic captioning only.

Keywords: captioning, access, accessibility, online learning, distance education, deaf, hard of hearing

Smith, C., Crocker, S. & Allman, T. (2017) Reading between the lines: Accessing information via *YouTube*'s automatic captioning, *Online Learning 21*(1), 115-131. doi: 10.24059/olj.v21i1.823

Introduction

Since *Google*'s acquisition of *YouTube* in 2006, both web giants have been working on developing a captioning method to make web-based video content accessible for deaf and hard of hearing users. *YouTube* currently offers users who post videos the option—which YouTube strongly encourages—to add subtitles and captions to their video. https://support.google.com/youtube/answer/2734796?rd=1 (3 Play Media, 2014). Also available to users is an automatic captioning feature. The automatic captioning feature is based on speech-recognition technology that employs a complex statistical model for the probability of specific sounds, words, and word combinations occurring within a language. According to *Google*'s *YouTube* Help site, automatic captioning is available in 10 languages worldwide (*Google*, 2015).

While attending class, and preparing university courses that include students who are deaf and/or hard of hearing, the authors noted the volume of errors in the automatic captioning present in several of the videos posted and viewed in *YouTube*. After some discussion, the authors decided to conduct a preliminary investigation into the overall effectiveness of the automatic captioning tool. To determine the consistency of *YouTube*'s automatic captioning feature of online videos, 50 videos targeted at a middle-school audience were viewed. In each of the videos, a variety of errors and error types were documented. Though some error types were more prominent than others, each error type plays a role in the overall comprehension of the video content. Errors were divided into 11 categories during the review and included addition of words, deletion of words, coherent miscues, incoherent miscues, spelling, incomprehensible phrases, word condensing, homonyms, approximation (content errors), speed, and visual readability. Videos with audio of 1) non-native English speakers with accents, 2) young children's voices, 3) voices that contained mumbling or computer-generated/mechanical sounds were harder to understand and had more captioning errors. The speed of the videos and timing of the captioned content also proved problematic. Readability of the captioned content and aesthetics of the captioning were also noted for each video.

The results of the initial project were gathered by watching the first two minutes of the 50 videos focusing on the solar system with the automatic captioning feature enabled and with content from the 8^{th} grade Texas Essential Knowledge and Skills (TEKS) goals. The frequency and type of errors were documented along with the quality, speed, and type of video. The data were then put into an *Excel* spreadsheet in preparation for the data analysis. The errors that were documented from each phrase were categorized as follows: additions, deletions, coherent miscues (full-phrase), incoherent miscues (full-phrase), miscues of a single word, word condensing, homonyms, approximations, and morphemes. These categories were established to set up guidelines for what was to be considered an error. This initial evaluation of *YouTube's* content prompted the current study.

Review of Literature

Effects of Captioning

The discussion of captioning audio-visual material must focus on more than simply attempting to present textual representations of audio content on the video. Simply putting text on the screen is insufficient for providing equitable access to the audio content. Successful captioning has been an appropriate supplement to video-based materials and has even been shown to be useful as a foreign language instructional tool when used with videos containing native speaker accents (Dabhi, 2004). Captioning in different formats, including keyword captioning, in which students view the video with partial captioning using only pre-selected keywords while listening to a video at the same time, has proven effective with users of video-based content, especially when the complexity of the video content is beyond the reading level of the viewer (Ruan, 2015). In such instances, the captioned content can help clarify the viewer's understanding of the video content presented. Lewis & Jackson (2001) found that the script comprehension for captioned videos for students who were deaf or hard of hearing was greater than the comprehension of script in other text forms and increased comprehension beyond the identified reading levels of students. Verbatim captions that are paced to the natural rate of delivery provide access to complete conversational exchanges including both the audio and visual information and allow viewers to comprehend both explicit and implicit information.

There is an advantage to both deaf and hearing students in terms of comprehension when video and captions are presented (Lewis & Jackson, 2001). Advantages of captioned video include facilitating novel vocabulary identification and overall comprehension (Winke, Gass, & Sydorenko, 2010). For second language learners, captioning aids in form-mapping, the process of connecting spoken and written vocabulary, by not having to focus auditorily on word meaning, and instead focus on printed form to connect it with meaning.

Gass, & Sydorenko, 2010). Information presented verbally and visually is integrated as it is stored in memory (Sadoski & Paivo, 2004). Johnson-Glenberg (2000) reports that the recall of the linguistic information will stimulate retrieval of the visual information and vice versa. However, even when captioned material contains enhanced or expanded content, it often goes unused by educators despite feedback from students indicating that using captioning while viewing video content would be appealing (Steinson & Stevenson, 2015).

Successful Caption Use

There are a number of issues involved with successful captioning of audio-video content. Two issues affecting the overall captioned experience include speed and formatting. Jensema & McCann (1995) found that the "safe speed" for word content displayed in captioned material was approximately 120-140 words per minute. Unfortunately, captioned material can be presented at speeds exceeding 200 words per minute.

Formatting can make the message delivery problematic as well. Closed-captioning versus automatic captioning and placement of the text on the video content can also play significant roles in how the captioned content is understood. Closed-captioning involves embedded textual content (by an author/programmer) that is timed to present with the audio content synchronously. Closed captioning can be presented live, as with television programming, or post-production, as with cinema movie content, and is not visible until the user activates their decoding systems and displays the captions on their screens. New television sets and video display monitors sold in the United States must be equipped with built-in caption decoder chips. Schools, colleges, libraries and other recipients of federal financial assistance are required under section 508 of the Rehabilitation Act to make their communication accessible to and usable by persons with disabilities (National Association of the Deaf, 2002). Automatic captioning involves the presentation of textual material based solely on the success of speech-recognition software. The captioned content for automatic captioning is not a permanent component of the video, and often varies in accuracy based on the speech delivery of the audio content.

Despite the numbers of individuals with disabilities using social media options and information and communication technologies (ICTs) today, many individuals still struggle with issues related to accessibility (Seale, Georgeson, Mamas, & Swaim, 2015; Asuncion, Budd, Fichten, Nguyen, Barile, & Amsel, 2012; Fichten, Asuncion, & Scapin, 2014). In response to increased population with hearing loss who use Internet-based technologies and social media, *YouTube* has created an automatic captioning feature. The feature attempts to approximate text-based representations of the speech audio content within videos through the use of speech recognition software. Sadly, researchers and companies, including *Google*, have recognized that despite the vast resources of *Google* and *YouTube*, the automatic captioning can fail to accurately convey the intended message (Barton, Bradbrook, & Broome, 2015; Johnson, 2014).

No current research focusing on the success or comprehension of material with the automatic captioning feature of *YouTube* was found. This mirrors historical trends regarding the use of captioning. In fact, according to Cambra, Silvestre, and Leal (2009), the study of deaf individuals' use of closed-captioning on television has not been a research priority in the field of deaf education. In a review of the literature of closed-captioning on television, no research on the influence of errors within the captioning was located. As captioning for online content is relatively new, the issue of Web-based captioning has not become a priority in deaf education either.

Successful use of captioned video material requires a significant cognitive load visually even when the captioned material is presented appropriately. Cognitive load can be explained as a complex theory that attempts to quantify the burden that performing a specific task imposes on the cognitive system of a learner (Paas & van Merriënboer, 1994; Paas, Tuovinen, Tabbers, & Van Gerven, 2003). In relation to this study, cognitive load occurs when a viewer relying on captioned material to access the content presented auditorily must attempt to access both the captioned content and the video-based content simultaneously. The brain is taxed visually and the two modalities compete for delivering the material to the brain via the same cognitive space. That cognitive competition can add stress to the situation making comprehension more difficult, especially for those with limited reading proficiency.

The primary focus for this study is based in part on the assertion by Cambra, Silvestre, and Leal (2009) that reading comprehension and reading speed influence comprehension of closed-captioned text of video content. Because of the complexity of successfully navigating captioned video material, comprehension of such material requires individuals with successful reading skills. The proposed research seeks to evaluate the comprehensibility of *YouTube*'s automatic captioning feature based on the reading abilities of college-educated adult readers. The research question for this study is, "Can college-level adult readers understand basic concepts from science videos posted on *YouTube* with auto-captioned text containing errors?"

Methodology

Participants

Participation in the study was open to students, staff, and faculty at a doctoral-granting, public university in the southwestern US (Table 1). To locate participants with successful reading skills, participants had to have had *some* college experience in order to participate in the study. The researchers anticipated that participant ages would vary and range from 18 to 60+ years. All genders of adult students, staff, and faculty at the university were allowed to participate. Study participation was dependent on reading ability.

The study sought to determine whether individuals with college-level reading abilities could understand the printed messages of content delivered via the *YouTube* automatic captioning feature. Participants were required to have the ability to read Web-based automatic captioning and simple sentences and questions, as well as write their own answers to the questions about the videos.

Participants for the research study were from the university community where the research was conducted. Researchers emailed an invitation to university students, faculty, and staff on the primary campus of the university regarding participation in the study. Participants were invited to participate in the study using one of the on-campus computer classrooms on the university campus. Participants completed a consent form before participation was permitted.

The research team hosted and monitored the study participants in a university computer classroom/lab using an Internet-connected computer and paper questionnaire. Demographic information was collected at the beginning of the classroom portion of the research (Table 1). The vast majority of participants were female (94%), with English as a first language (81.2%), and were hearing (95%). Only three individuals self-identified as deaf or hard of hearing. Participants watched one of three middle-school videos about the solar system and answered basic questions about the material presented via one of the three viewing options. Using a pen and a paper questionnaire, participants answered a series of questions for one of three video options on the solar system. The video selection, caption availability, and sound availability were chosen at random for each participant and the sound was turned off or on accordingly. The research occurred over a two-week period. Participants were given the option of selecting a date they wished to participate. Participants did not know which video or under which viewing conditions they would be watching the video until they sat down at the computer. Each of the computer stations was randomly set to view one of the videos under a specific viewing option. The questionnaires were numbered to indicate which video and viewing condition. The computers were set up to view the corresponding video under the specific viewing condition. Students self-selected where they sat upon entering the classroom.

Originally, participants were to view one of three middle school science videos on the solar system labeled as having captions with “few” errors, “moderate” errors, or “high” errors. The error types used in determining the groupings of “few-error, moderate-error, and high-error” status were identical to those used at the initial stages of the project to determine the types of errors present in the automatic captioning. Videos were selected from the initial evaluation of 50 videos. Videos were categorized into one of the three groups, based on the number of errors present. To select the videos for the study, the authors used the videos with the highest, median, and lowest numbers of errors. Videos were then re-evaluated to determine which in each group contained standard American English speech from a human (not digitized or robotic speech) with limited to no accents so that the only influence on the automatic captioning was the quality of the speech recognition software. Due to limited participation, the authors focused the initial viewing sessions on the videos with “few” and “high” errors. No participants viewed the videos with “moderate” instances of errors. Participants for each session viewed the videos similarly under one of the following conditions: 1) sound without captioning, 2) sound with automatic captioning, or 3) automatic captioning without sound (Appendix A). Determination of which participant sessions were given a particular viewing method were determined by random selection. Participants who seated themselves at computers where the sound was enabled for the video viewing were provided a new pair of earphones to use throughout the experience. Participants were free to take the earphones with them once their participation was complete. Sessions were arranged so that approximately the same number of participants viewed each video option.

Each of the *YouTube* videos were watched online during each viewing. The Web addresses for the videos are listed at the end of this document and in the attached video questionnaires (Appendix B). Participants viewed the pre-determined video on a university computer. Participants were free to watch the video and answer the questions as they wished. They could watch the video entirely and then go back and answer questions or they could answer questions while watching the video. While participants were only allowed to participate one time in the study, there were no limitations on the number of times they could watch the video or pause and go back during that participation. The researchers had the *YouTube* site open to the appropriate video upon participants' arrival to the computer lab.

Results

Participants

Frequencies and percentages for the demographic variables are displayed in Table 1. The largest percentage of participants were in group 2 (20.3%) and the majority of participants were in low error groups (52.7%). In addition, the largest percentage of participants were in the sound and caption group (35.1%). Finally, the majority of participants were female (94.0%), reported that English was their first language (81.2%), and had identified themselves as hearing (95.1%).

Means and standard deviations for the continuous variables are displayed in Table 2.

Participants' ages ranged from 18 to 59 (M = 22.55, SD = 6.14) and the number of years of college experience ranged from 0<1 to 11 (M = 3.21, SD = 1.83). The number of correct responses by participants ranged from 1 to 12 (M = 8.76, SD = 2.88)

Table 1. *Frequencies and Percentages of Categorical Demographic Variables*

		n	%
Total Groups			
	Group 1 (Low Error, Sound No Caption)	11	14.9
	Group 2 (Low Error, Sound and Caption)	15	20.3
	Group 3 (Low Error, Caption No Sound)	13	17.6
	Group 7 (High Error, Sound No Caption)	13	17.6
	Group 8 (High Error, Sound and Caption)	11	14.9
	Group 9 (High Error, Sound and Caption)	11	14.9
Error Groups			
	Low Error	39	52.7
	High Error	35	47.3

Caption Group		
Sound No Caption	24	32.4
Sound and Caption	26	35.1
Caption No Sound	24	32.4
Gender		
Female	63	94.0
Male	3	4.50
Transgendered	1	1.50
English as First Language		
No	13	18.8
Yes	56	81.2
Hearing Status		
Deaf	1	1.6
Hard of Hearing	2	3.3
Hearing	58	95.1

Table 2. *Means and Standard Deviations of Continuous Variables*

	n	*M*	*SD*	Min	Max
Age	69	22.55	6.14	18.00	59.00
College Experience (Years)	69	3.21	1.83	0.00	11.00
Number Correct	72	8.76	2.88	1.00	12.00

Table 3 *Means and Standard Deviations for the Number of Correct Responses by Gender and Group*

	Low Error		Higher Error		Total	
	n	Mean	n	Mean	n	Mean
Sound No Caption	11	10.45 [a]	13	8.85 [b, y]	24	9.58
		(1.13)		(1.68)		(1.64)
Sound and Caption	15	10.07	11	9.00 [y]	26	9.62
		(1.10)		(2.90)		(2.08)
Caption No Sound	13	9.69 [a]	9	2.78 [b, x]	22	6.86
		(21.45)		(1.72)		(3.80)
Total	39	10.05	33	7.24	72	8.76
		(1.23)		(3.48)		(2.88)

Note. Standard deviations are shown in parenthesis below means. [a b] rows with different superscripts differed significantly, [x y] columns with different superscripts differed significantly

Data Analysis

A 2 (error: low vs. high) x 3 (caption: sound no caption vs. sound and caption vs. caption no sound) two-way analysis of variance (ANOVA) was conducted to examine the effect of error group and caption group on the number of correct responses.

There was a statistically significant interaction between error group and caption group, F (2,66) = 19.69, $p < .001$, *partial eta squared* = .374.

Simple main effects analyses revealed that participants in the low error group who watched the video with caption but no sound answered a significantly greater number of questions correctly (*M* = 9.69, *SD* = 1.44) than participants in the high error group who watched the video with caption but no sound (*M* = 2.78, *SD* = 1.72), $p < .001$ (Figure 1). In addition, participants in the lower error group who watched the video with sound but no caption answered a significantly greater number of questions correctly (*M* = 10.45, *SD* = 1.13) than participants in the high error group who watched the video with sound but no caption (*M* = 8.85, *SD* = 1.68).

Simple main effects analyses also revealed that participants in the high error group who watched the video with caption but no sound answered a significantly fewer number of questions correctly (*M* = 2.78, *SD* = 1.72) than participants in the high error group who watched the video with sound and caption (*M* = 9.00, *SD* = 2.90) and participants in the high error group who watched the video with sound but no captions (*M* = 8.85, *SD* = 1.68) (Figure 2).

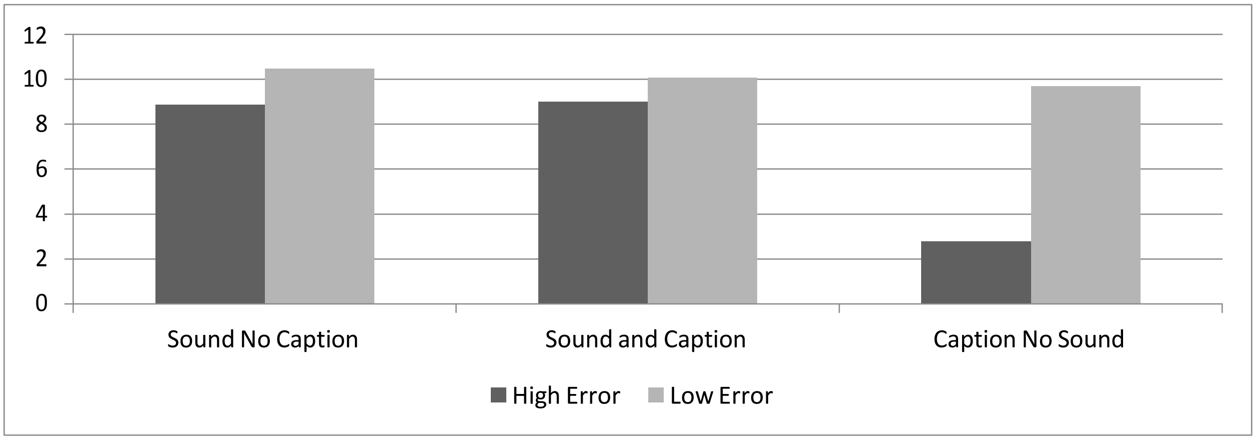
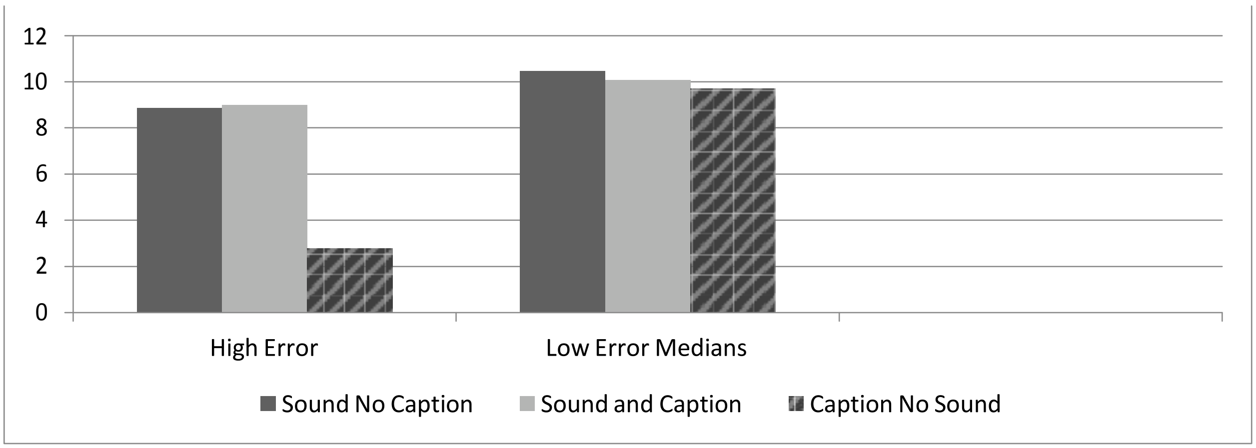


Figure 1: Comparison of comprehension errors between high and low captioning error



groups based on viewing options

Figure 2: Comparison comprehension low error frequency versus high error frequency based on viewing option

For the low error group, the three caption groups did not significantly differ in total correct scores. For the high error group, the caption no sounds group had significantly fewer correct scores than the sound no caption and the sound and caption groups (Figures 1 & 2).

Limitations and Future Directions

This study was foundational research into the effectiveness of web-based automatic captioning for successful adult readers. Limitations to the study included a predominantly female participant group, stemming primarily from the make-up of the university's student body. Another limitation of the study was the use of successful readers. Future research should include the use of automatic captioning with struggling readers such as school-age deaf and hard of hearing students who demonstrate a variety of reading levels and who would typically be exposed to educational content similar to the content presented in this study. The extent to which such readers struggle with problematic automatic captioning should also be evaluated. While this research explored the abilities of individuals with native English reading experiences, future research should also include participants with limited English abilities to determine whether even accurate automatic captioning is problematic and the extent to which it is troublesome.

Discussion

There are several points of discussion based on the results of the study. Overall, the results indicated that the number of questions participants were able to answer correctly did not depend only on the type of error or caption group in which they were placed, they also depended on the interaction between the two types of groups. Participants in the high error group who watched a video with caption but no sound answered a significantly fewer number of questions correctly than participants in the high error group who watched a video with sound but no caption or both sound and caption.

The results clearly demonstrate that when auto-captions are presented accurately, regardless of whether sound is present, typical adult readers are able to comprehend the messages being delivered via text. Conversely, as assumed, when automatic-captions are presented inaccurately, containing significant numbers of errors, and when no audio content is available, even hearing, college-educated adult readers are unable to comprehend the messages being delivered. These results are significant in that participants in this study all have reading experience at the postsecondary level and were unable to accurately perceive automatic-captioned messages with high errors that were delivered without sound. If adult readers with the ability to interact with print at the college level are unable to successfully navigate such automatic-captioned content, expecting school-age students who are deaf or hard of hearing with varied reading abilities to perform any better would be inappropriate as caption comprehension is highly and positively correlated with grade level (Lewis & Jackson, 2001).

Captioning is critical for video to be accessible to students who are deaf or hard of hearing. However, the focus must be more than simply putting text on a screen, as appropriate captioning is imperative for successful comprehension. *YouTube* Teacher was created to help K-12 teachers use educational video in their classrooms to support learning and engage and inspire learners. *YouTube* for Schools allows schools that opt-in to access thousands of educational videos (Buzzetto-More, 2014). Clearly the use of multimedia and the presentation of verbal and visual information will continue to be a common and recommended model of instruction in classrooms. Teachers need to become adept at using captioning accurately. As indicated earlier, teachers can manually caption their videos using several of the tools available in *YouTube's* video manager.

Accurate captioning is equally important for access at the university level. Faculty are using and recommending the use of online videos such as those found on *YouTube* (Moran, Seaman, & Tinti-Kane, 2011; Tan & Pearce, 2011). Betts, Cohen, Veit, Alphin, Broadus and Allen (2014) identified inaccessibility to videos and voice-over PowerPoint Presentations because they do not have captions as one of the greatest challenges for an online student with a hearing loss. Such recommendations carry the weight of assuring that videos for courses are accessible to all students. Relying on the automatic captioning feature of *YouTube* will be insufficient to provide student users who require the captioned text for comprehension.

Captioning is vital but needs to be studied. We are not suggesting getting rid of it—only finding out what works best. Further discussion regarding the automatic captioning of web-based video content should center on several issues, including mandating which web-based content should be permitted to employ automatic captioning features and how to improve upon the infrastructure of automatic captioning and speech recognition platforms.

References

Asuncion, J. V., Budd, J., Fichten, C. S., Nguyen, M. N., Barile, M., & Amsel, R. (2012). Social media use by students with disabilities. *Academic Exchange Quarterly*, *16*(1), 30-35.

Betts, K., Cohen, A.H., Veit, D.P., Alphin, H.C., Broadus, C., & Allen, D. (2014). Strategies to increase online student success for students with disabilities. *Journal of Asynchronous Learning*, 17(3), 50-65.

Barton, F., Bradbrook, G., & Broome, G. (2015). *Digital accessibility: A report for citizens online*. Edinburgh, Scotland.

Buzzetto-More, N. (2015) An examination of undergraduate student's perceptions and predilections of the use of YouTube in the teaching and learning process, Interdisciplinary Journal of e-learning and learning objects. Retrieved from https://www.questia.com/library/journal/1G1-418467094/an-examination-of-undergraduate-student-s-perceptions.

Cambra, C., Silvestre, N., & Leal, A. (2009). Comprehension of television messages by deaf students at various levels of education. *American Annals of the Deaf*, *153*(5), 425-434.

Dahbi, M. (2004). English and Arabic after 9/11. *The Modern Language Journal*, *88*(4), 628-630.

Fichten, C. S., Asuncion, J., & Scapin, R. (2014). Digital Technology, Learning, and Postsecondary Students with Disabilities: Where We've Been and Where We're Going. *Journal of Postsecondary Education and Disability*, *27*(4), 369-379.

Google. (2015). Auto captioning [*YouTube* Help: Edit videos & setting: Caption]. Retrieved from https://support.*Google*.com/*YouTube* /answer/3038280?hl=en

Jensema, C. & McCann, R. (1996). Presentation Speed and Vocabulary in Closed Captioned Television. *American Annals of the Deaf*, 141(4).

Johnson, A. (2014). Video Captioning Policy and Compliance at the University of Minnesota Duluth. Unpublished Master's Thesis, University of Minnesota – Duluth, USA.

Johnson-Glenberg, M. (2000). Training reading comprehension in adequate decoders/poor comprehenders: Verbal and visual strategies. *Journal of Educational Psychology*, 92(4), 772-782.

Lewis, M. & Jackson, D. (2001). Television literacy: Comprehension of program content using closed captions for the Deaf. *Journal of Deaf Studies and Deaf Education*, 43-53.

Moran, M., Seaman, J., & Tinti-Kane, H. (2011). Teaching, Learning, and Sharing: How Today's Higher Education Faculty Use Social Media. *Babson Survey Research Group.*

Paas, F., & van Merriënboer, J. J. G. (1994a). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review, 6*, 51–71.

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational psychologist*, *38*(1), 63-71.

Ruan, X. (2015). The Effect of Caption Modes on EFL Students' Video Comprehension. *Journal of Language Teaching and Research, 6,* 2, 397-404.

Sadoski, M., & Paivio, A. (2004) A dual-coding theoretical model of reading. In R. B. Ruddell & N. N. Unrau (Eds.), *Theoretical models and processes of reading* (5th ed., pp. 1329-1362). Newark, DE: International Reading Association.

Seale, J., Georgeson, J., Mamas, C., & Swain, J. (2015). Not the right kind of 'digital capital'? An examination of the complex relationship between disabled students, their technologies and higher education institutions. *Computers & Education*, *82*, 118-128.

Stinson, M. S., & Stevenson, S. (2013). Effects of expanded and standard captions on deaf college students' comprehension of educational videos, American Annals of the Deaf, 158(4), 453-467.

Tan, E., & Pearce, N. (2012). Open education videos in the classroom: exploring the opportunities and barriers to the use of YouTube in teaching introductory sociology. *Research in Learning Technology, 19*.

Texas Educational Agency (2010). Chapter 112. Texas Essential Knowledge and Skills for Science, Subchapter B. Middle School. Retrieved from http://ritter.tea.state.tx.us/rules/tac/chapter112/ch112b.html#112.20

Winke, P., Gass, S., & Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology*, *14*(1), 65-86.

3 Play Media (2014). Future of accessibility and video captions according to google and YouTube. Retrieved from http://www.3playmedia.com/2014/01/10/future-accessibility-video-captions-according-google-youtube/

APPENDIX A: VIDEOS & VIEWING CONDITIONS

Group #	Description
	VIDEO ONE LOCATION: http://www.*YouT ube* .com/watch?v=B1AXbpYndGc
1	Video 1(low error) sound no caption
2	Video 1(low error) sound and caption
3	Video 1(low error) caption no sound
	***VIDEO TWO LOCATION**: http://www.*You Tube* .com/watch?v=RJ0JCg3S7xQ
4	Video 2(medium error) sound no caption
5	Video 2(medium error) sound and caption
6	Video 2(medium error) caption no sound
	VIDEO THREE LOCATION: http://www.*Yo uTube* .com/watch?v=tDnawSj64jg
7	video 3(high error) sound no caption
8	Video 3(high error) sound and caption
9	Video 3(high error) caption no sound

*No participants viewed this group of videos. Originally, the medium-error videos were to be viewed. When it became apparent that the number of participants were going to be limited, the authors chose to focus attention on the high-error and low-error videos.

APPENDIX B: VIDEO VIEWING QUESTIONNAIRE EXAMPLE

Video Questionnaire: GROUP 1

"Naked Science: Birth of the Solar System"
http://www.youtube.com/watch?v=B1AXbpYndGc

1. A dense clump of water formed what?
2. When a star reaches 18 million degrees Fahrenheit what kicks in?
3. When was our star(the sun) born?
4. What fuses together to form helium?
5. What is the first type of light made by our sun?
6. Was the solar system's birth peaceful?
7. Where was the sun born?
8. An entire universe was supposed to be created from what?
9. A big explosion, that caused the creation of the universe, is known as?
10. There are how many naturally occurring chemical elements?
11. What are two elements that planets are made of?
12. Hydrogen and Helium fuse to make what?

YouTube Automatic captioning Participant # _

YouTube Automatic captioning Participant # _

Gender: M F T Is English your first language? Y N Age: _
# of years college experience? _ Hearing status: Hearing Deaf Hard of Hearing

YouTube Automatic captioning Participant # _

SECTION III:
Students, Community, and Online Learning

Online Learning Integrity Approaches: Current Practices and Future Solutions

Anita Lee-Post and Holly Hapke
University of Kentucky

Abstract

The primary objective of this paper is to help institutions respond to the stipulation of the Higher Education Opportunity Act of 2008 by adopting cost-effective academic integrity solutions without compromising the convenience and flexibility of online learning. Current user authentication solutions such as user ID and password, security questions, voice recognition, or fingerprint identification are not infallible and may violate students' rights to privacy or cause undue interruptions to their efforts in performing assessment tasks. Existing authentication solutions are evaluated for their cost effectiveness in preventing fraud and cheating while ensuring learner identity and honesty. Emerging technologies in the form of biometrics, surveillance systems and predictive analytics are also examined to provide insights into the future of e-authentication for ensuring the academic integrity of online learning.

Keywords: academic integrity, online education, authentication, higher education opportunity act, academic misconduct

Lee-Post, A. & Hapke, H (2017). Online learning integrity approaches: Current practices and future solutions, *Online Learning 21*(1),135-145. doi: 10.24059/olj.v21i1.843

Introduction

The number of students taking at least one online course has been growing at a rate faster than that of the overall higher education student body since 2003, reaching over seven million in 2013 (Allen & Seaman, 2015). Students enjoy the flexibility to learn anywhere, anytime, and anyplace at their own convenience and preference. On the other hand, online education gives higher education institutions a means to increase student access with the potential to reduce costs and increase productivity. Despite the growing popularity and acceptance of online education, there is concern about its rigor and quality. A 2013 Gallup poll survey found that 49% of Americans believed that employers did not perceive an online degree as positively as a traditional one. In addition, 45% of Americans thought online education provided less rigorous testing and grading that could be trusted than the traditional classroom-based counterpart (Saad, Busteed, & Ogisi, 2013). To determine if our students' perception of academic integrity corresponded, we administered a survey to juniors and seniors in an online undergraduate course in Operations Management (n=167). We found that while nearly all students indicated they have not had someone else take an exam for them, over 45% regarded cheating in an online class as easy and 30% would cheat if given an opportunity.

There, we felt a need to address the lack of trust in online education, and an examination of its academic integrity solutions was in order. A review of current and emerging approaches to online learning integrity will be presented in this paper. The effectiveness of these approaches will then be assessed to provide insights into best practices and future solutions that may ensure the academic integrity of online learning. Here we use the term approach to denote a broad category or strategy. A specific implementation of an approach is called a solution or practice.

Background

Academic integrity is defined as a commitment to six core values, namely, honesty, trust, fairness, respect, responsibility, and courage, in all aspects of scholarly practices, even in the face of adversity (Fishman, 2012). The six core values serve to guide behavior that is congruent with the values. An investigation of the extent of academic integrity is being practiced in online education should therefore involve an examination of the values and behaviors of the institution, faculty, and students against a set standard. However, the broad nature of such investigation is beyond the scope of this paper. Thus, we narrow our focus to the institution level and adopt the Higher Education Opportunity Act of 2008 as the minimum standard against which approaches to online learning integrity are assessed.

The Higher Education Opportunity Act (2008) states that "Institutions that offer distance education must have processes through which the institution establishes that the student who registers in a distance education or correspondence education course or program is the same student who participates in and completes the program and receives academic credit." While the Act does not reflect all six core values of academic integrity, it asks institutions to provide assurance that a process is in place to authenticate learners in a virtual environment to ensure a registered student is the one who is actually doing the course work. This implies that institutions need to have a way to (1) create and maintain a virtual learning environment that only registered learners can access; (2) monitor and track registered learners' learning activities; (3) detect and deter academic integrity misconduct in general, and impersonation, in particular. Simply put, institutions are to put in place effective learner authentication solutions to prevent fraud and cheating while ensuring learner identity and honesty.

Literature Review

We conducted a literature review with the goal of identifying relevant research articles on online learning integrity solutions. Keywords including "online education," "online learning," "cheating," "academic dishonesty," "academic integrity," "authentication," "Higher Education Opportunity Act," "technology," and "technological solution" were used to search the Google Scholar, Academic Search Complete, Web of Science, and ERIC databases. Articles that were not from academic peer-reviewed outlets (e.g., periodicals, blogs) were excluded, resulting in twenty key articles. Relevant articles cited by the key articles are included to give a final set of

34 papers that form the basis of our discussion on current and future solutions for online learning integrity.

Existing online learning integrity approaches can be divided broadly into two types: prevention and enforcement. Prevention approaches are proactive strategies that stop misconduct from happening in the first place. Jones (2009) advocates the use of an honor code and authenticity statement to ensure students understand and commit to institutional values of character and integrity. The honor code provides a clear definition of academic integrity and the consequences of non-compliance, whereas an authenticity statement is a signed declaration from students acknowledging that the work is genuinely their own. In an online environment, students can be reminded of the honor code periodically and/or required to submit an authenticity statement when submitting the course work. Mcallister and Watkins (2012) suggest seven ways that an online course can be redesigned to develop students' self-regulation skills to refrain from engaging in academic misconduct. Their seven course design recommendations are: (1) use extensive calendaring to promote task planning and time management; (2) monitor ongoing stream of work instead of exams; (3) randomize exam questions to individualize an exam for each student; (4) discuss academic integrity to create awareness and commitment; (5) allow asynchronous learning to decouple student progress; (6) track student submissions to identify potential inconsistencies; (7) provide prompt feedback to facilitate a student's assessment of progress.

These prevention approaches are supported by the cognitive development theory which posits that the knowledge of academic integrity will compel an individual to act accordingly (Kohlberg, 1984). These approaches are also in line with the view of Chickering and Reese (1993) that integrity is one of the seven developmental tasks for optimal student growth and success. For the prevention approaches to be effective, an institutional culture of academic integrity needs to be developed. It requires an institution to (1) articulate clearly what constitutes academic integrity; (2) gain faculty commitment to honor and enforce integrity practices; (3) develop students' integrity and self-regulation skills; (4) develop an academic integrity system to measure, monitor, and track academic integrity development.

Enforcement approaches, on the other hand, are defensive strategies that detect academic misconduct. Software such as TurnItIn can be used to detect plagiarism for written assignments and class discussion (Heckler, 2013; Moten et al., 2013). Browser lock-down software such as Respondus can be used to control a testing environment that prevents students from printing, copying, screen-sharing, screen-capturing, going to another website, or accessing other applications while taking a test (Sewell et al., 2010). In addition, authentication solutions can be used to confirm the identity, authenticity, and presence of a student engaging in online learning activities. Authentication solutions range from the basic user ID and password to biometric schemes to video monitoring.

The first line of defense in user authentication is to allow only registered users to access the online learning systems. This is usually done by confirming the identity of the user based on the user's knowledge of unique facts about himself or herself. A user ID and password scheme is the most commonly used knowledge-based authentication solution. Other knowledge-based authentication solutions include challenging or security questions (Ullah et al., 2012; McNabb, 2010).

While knowledge-based authentication solutions are simple and easy to use, they cannot prevent collusion and impersonation. A strong authentication solution uses the user's biometrics (who the user is or what the user does distinctively) such as fingerprint, face, iris, voice, signature, and keystroke to confirm both the identity and authenticity of the user (i.e., it is really you?) (Rabuzin et al., 2006). However, biometric-based authentication solutions require the use of special devices to read and match a user's characteristics. There are also concerns about data security and privacy issues in dealing with sensitive data on users. In addition, user characteristics such as face, signature, and keystroke require complex technology and training overhead.

Biometric-based authentication solutions can only prevent impersonation at initial login. To ensure that the user stays put after the initial login, a next level of solution called continuous or presence authentication is needed. Presence authentication solutions are of particular relevance in authenticating users taking online examinations. Video monitoring and/or recording via webcam is a commonly used presence authentication solution (Apampa et al., 2010). Once again, additional devices for video recording and sophisticated software for analyzing video footage are needed. In addition, institutions need to have data security and privacy control measures in place to safeguard sensitive user-specific data from being stolen or lost.

Another presence authentication solution is proctoring. Both face-to-face and virtual proctoring can be viable solutions to authenticating users taking high stakes examinations. Face-to-face proctoring requires students to physically go to a testing center to take a test at a specific time (Larson & Sung, 2009; Shapley, 2000). Virtual proctoring usually is arranged with a third-party provider such as ProctorU (www.proctoru.com), RemoteProctor (www.remoteproctor.com), and SmarterProctoring (www.smarterprocoring.com) (Dunn et al., 2010). Depending on the level of authentication solutions needed, it costs from less than $10 to over $100 for each proctored examination. For example, RemoteProctor charges an annual fee of $30 and an equipment fee of $125 to use fingerprints for student identification, and video surveillance and recording systems for continuous authentication (Rodchua et al., 2011).

Assessment of existing approaches

In tables 1 and 2 we evaluate the online learning integrity approaches for their cost effectiveness with respect to the stipulation of the Higher Education Opportunity Act. Costs from the perspective of the institution, faculty, and students are considered. They include loss of flexibility, inconvenience, privacy concerns, security concerns, third-party involvement, extra technological requirements, extra costs, and extra effort. Effectiveness is measured as the extent to which user authentication can be confirmed. A summary of the assessment of prevention approaches and enforcement approaches are provided in the appendices (see below).

For prevention approaches, such as honor code, authentication statement, and course re-design, the extra effort put in is worthy of the benefits gained if a culture of academic integrity is developed at the institution, faculty, and student levels. However, culture is difficult if not impossible to measure objectively. As such, prevention approaches alone may not be able to satisfy the stipulation of the Higher Education Opportunity Act as the honor code or authentication statement are not solid evidence of user authentication.

For enforcement approaches, knowledge-based and biometric authentication solutions require minimal effort and extra technologies to confirm user identity and authentication at log in. However, they are not able to prevent impersonation and collusion. In order to provide a satisfactory assurance that the registered user is the one completing the coursework, a more expensive presence authentication solution will need to be adopted.

Emerging online integrity solutions

As biometric technologies become more accurate and less costly, an authentication solution based on a unique aspect of who the user is and/or what the user does surely will replace the simplistic username and password scheme as a stronger proof of user identity, authenticity, and presence. Among the different biometric-based authentication solutions, fingerprinting is the most mature and proven technology for such purpose (Yang et al., 2011; Ratha et al., 2001). Indeed, fingerprint biometrics has already been incorporated in Apple's iPhone 5 for user identification and authentication. It is only a matter of time before a computer's input device will have a built-in fingerprint reader. As learners use such devices to interact with the virtual learning environment, their fingerprint biometrics can be examined in a continuous fashion to perform presence authentication in a non-intrusive manner.

A unimodal biometric-based authentication solution is not without its vulnerabilities and limitations. Collusion cannot be prevented if a biometrically authenticated user has someone's help in taking an exam. In addition, fingerprint biometrics will not be administrable for a student lacking this feature because of physical impairment. A multi-modal scheme for user authentication that involves surveillance technologies is therefore necessary. A bimodal scheme such as video monitoring can be used in conjunction with biometric authentication to prevent collusion. Such a scheme is less intrusive and more effective than having to re-authenticate the user when suspicious behavior is detected. A tri-modal scheme such as browser tracking and/or lock-down can also be added to video monitoring and biometric authentication to further assure that the student does not have access to unauthorized resources while taking a test. Biometric authentication adaptations or special accommodations can be made for students with disabilities. In any case, further advancement in biometric and surveillance technologies will provide institutions with more cost-effective options for online learning integrity assurance.

Predicative analytics is another area of technological advancement that holds promise in the development of next generation online integrity solutions. As students interact with the virtual learning environment, a wide variety of data such as their physical location, devices used, access patterns, learning progress, performance, etc. can be collected. These data can be mined for integrity promotion purposes. For example, student-course interaction data can produce useful information about a student's level of engagement with the course, and generate low performance and/or procrastination warnings to steer at-risk students onto a path of success. These data can also be mined for integrity enforcement purposes. Unusual or suspicious activities (e.g., students who did not do their coursework and yet have a perfect score on an exam) can be identified from the data collected so that attention can be dedicated to investigate situations of significant integrity concerns. Predictive analytics, with its ability to extract information from data to predict trends and patterns of behavior, will be well suited in this regard.

Conclusion

We conducted a review of current approaches to online learning integrity. Existing approaches are assessed in accordance with the Higher Education Opportunity Act. Emerging technological solutions based on biometrics, surveillance, and predictive analytics are discussed. Although our review is far from exhaustive, it does provide a comprehensive overview of the cost effectiveness of different online learning integrity solutions. Institutions seeking conformance to the Higher Education Opportunity Act are urged to put in place a user authentication solution that can verify a learner's identity, authenticity, and presence. With the rapid pace of technological advancement, educational institutions will be able to implement cost-effective academic integrity solutions that are powered by sophisticated but affordable authentication hardware and software. An integrity solution that incorporates both prevention and enforcement approaches to adequately address the issues of academic integrity beyond user authentication will become a reality in the foreseeable future.

Acknowledgement

This study was presented in the 8th Annual Emerging Technologies for Online Learning International Symposium, April 22-24, 2015, with the support of the eLearning Innovation Initiative Grant at the University of Kentucky.

References

Allen, I.E., & Seaman, J. (2015). Grade Level: Tracking Online Education in the United States. Babson Survey Research Group and Quahog Research Group, LLC. Retrieved from: http://www.onlinelearningsurvey.com/reports/gradelevel.pdf

Apampa, K.M., Wills, G., & Argles, D. (2010). User Security Issues in Summative E-Assessment Security. *International Journal of Digital Society*, 1(2), 135-147.

Bailie, J.L., & Jortberg, M.A. (2009). Online Learner Authentication: Verifying the Identity of Online Users. *Journal of Online Learning and Teaching*, 5(2), 197-207.

Baron, J., & Crooks, S.M. (2005). Academic Integrity in Web Based Distance Education. *TechTrends*, 49(2), 40-45.

Bedford, W., Gregg, J., & Clinton, S. (2009). Implementing Technology to Prevent Online Cheating: A Case Study at a Small Southern Regional University. *Journal of Online Learning and Teaching*, 5(2), 230-238.

Bedford, W., Gregg, J., & Clinton. S. (2011). Preventing Online Cheating with Technology: A Pilot Study of Remote Proctor and an Update on Its Use. *Journal of Higher Education Theory and Practice*, 11(2), 41-58.

Caldwell, C. (2009). A Ten-Step Model for Academic Integrity: A Positive Approach for Business Schools. *Journal of Business Ethics*, 92(1), 1-13.

Chickering, A.W., & Reisser, I. (1993). *Education and Identity*, 2nd Edition. San Francisco: Jossey-Bass.

Chiesl, N. (2007). Pragmatic Methods to Reduce Dishonesty in Web-Based Courses. *The Quarterly Review of Distance Education*, 8(3), 203-211.

Dunn, T.P., Meine, M.F., & McCarley, J. (2010). The Remote Proctor: An Innovative Technological Solution for Online Course Integrity. *The International Journal of Technology, Knowledge and Society*, 6(1), 1-7.

Farcasin, M., & Chan-tin, E. (2015). Why We Hate IT: Two Surveys on Pre-generated and Expiring Passwords in an Academic Setting. *Security and Communication Networks*, 8, 2361-2372.

Fishman, T. (2012). The Fundamental Values of Academic Integrity. The International Center for Academic Integrity. Retrieved from: http://www.academicintegrity.org/icai/assets/Revised_FV_2014.pdf

Hart, L., & Morgan, L. (2009). Strategies for Online Test Security. *Nurse Educator*, 34(6), 249-253.

Heckler, N.C., Rice, M., and Bryan, C.H. (2013). TurnItIn Systems: A Deterrent to Plagiarism in College Classrooms. *Journal of Research on Technology in Education*, 45(3), 229-248.

Kirkpatrick, K. (2015). Technology Brings Online Education in Line with Campus Programs. *The Communications of the ACM*, 58(12), 17-19.

Kitahara, R.T., & Westfall, F. (2007). Promoting Academic Integrity in Online Distance Learning Courses. *Journal of Online Learning and Teaching*, 3(3), 265-276.

Kolberg, L. (1984). *Essays in Moral Development: The Psychology of Moral Development*, Vol. 2, New York: New York.

Inglesant, P.G., & Sasse, M.A. (2010). The True Cost of Unusable Password Policies: Password Use in the Wild. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 383-392.

Jones, I.M. (2009). Cyber-Plagiarism: Different Method Same Song. *Journal of Legal, Ethical and Regulatory Issues*, 12(1), 89-100.

Just, M., & Aspinall, D. (2009). Personal Choice and Challenge Questions: A Security and

Usability Assessment. *Proceedings of the Fifth Symposium on Usable Privacy and Security*, July 15-17.

Larson, D.K., & Sung, C. (2009). Comparing Student Performance: Online Versus Blended Versus Face-To-Face. *Journal of Asynchronous Learning Networks*, 12(1), 31-42.

LoSchiavo F.M., and Shatz, M.A. (2011). The Impact of an Honor Code on Cheating in Online Courses. *Journal of Online Learning and Teaching*, 7(2), 179-184.

Mastin, D.F., Peszka, J., & Lilly, D.R. (2009). Online Academic Integrity. *Teaching of Psychology*, 36, 174-178.

Mcallister, C., & Watkins, P. (2012). Increasing Academic Integrity in Online Classes by Fostering the Development of Self-Regulated Learning Skills. *The Clearing House*, 85, 96-101.

McNabb, L. (2010). An Update on Student Authentication: Implementation in Context. *Continuing Higher Education Review*, 74, 43-52.

McNabb, L. & Olmstead, A. (2009). Communities of Integrity in Online Courses: Faculty Member Beliefs and Strategies. *Journal of Online Learning and Teaching*, 5(2), 208-221.

Moini, A., & Madni, A.M. (2009). Leveraging Biometrics for User Authentication in Online Learning: A Systems Perspective. *IEEE Systems Journal*, 3(4), 469-476.

Moten, J., Fitterer, A., & Brazier, E. (2013). Examining Online College Cyber Cheating Methods and Prevention Measures. *Electronic Journal of e-Learning*, 11(2), 139-146.

Rabuzin, K., Baca, M., & Sajko, M. (2006). E-Learning: Biometrics as a Security Factor. *International Multi-Conference on Computing in the Global Information Technology*. 64-69.

Ratha, N.K., Connell, J.H., & Bolle, R.M. (2001). Enhancing Security and Privacy in Biometrics-based Authentication Systems. *IBM Systems Journal*, 40(3), 614-634.

Rodchua, S., Yiadom-Boakye, G., & Woolsey, R. (2011). Student Verification System for Online Assessments: Bolstering Quality and Integrity of Distance Learning. *Journal of Industrial Technology*, 27(3), 2-8.

Rowe, N.C. (2004). Cheating in Online Student Assessment: Beyond Plagiarism. *Online Journal of Distance Learning Administration*, 7(2). Retrieved from: http://www.westga.edu/~distance/ojdla/summer72/rowe72.html

Saad, L., Busteed, B., & Ogisi, M. (2013). In U.S., Online Education Rated Best for Vale and

Options. A Gallup Poll Survey. Retrieved from:: http://www.gallup.com/poll/165425/online-education-rated-best-value-options.aspx

Sewell, J.P., Frith, K.H., & Colvin, M.M. (2010). Online Assessment Strategies: A Primer. *Journal of Online Learning and Teaching*, 6(1), 297-305.

Shapley, P. (2000). On-line Education to Develop Complex Reasoning Skills in Organic Chemistry. *Journal of Asynchronous Learning Networks*, 4(2), 43-52.

Shay, R., Komanduri, S., Kelley, P.G., Leon, P.G., Mazurek, M.L., Bauer, L., Christin, N., & Cranor, L.F. (2010). Encountering Stronger Password Requirements: User Attitudes and Behaviors. *Proceedings of the Sixth Symposium on Usable Privacy and Security*. July 14-16.

The Higher Education Opportunity Act (2008). Retrieved from: http://www.gpo.gov/fdsys/pkg/PLAW-110publ315/pdf/PLAW-110publ315.pdf

Ullah, A., Xiao, H., Lilley, M., & Barker, T. (2012). Using Challenging Questions for Student Authentication in Online Examination. *International Journal for Infonomics*, 5(3/4), 631-639.

Vandehey, M., Diekhoff, G., & LaBeff, E. (2007). College Cheating: A Twenty-Year Follow-Up and the Addition of an Honor Code. *Journal of College Student Development*, 48(4), 468-480.

Yang, J., Xiong, N., Vasilakos, A.V., Fang, Z., Park, D., Xu, X., Yoon, S., Xie, S., & Yang, Y. (2011). A Fingerprint Recognition Scheme Based on Assembling Invariant Moments for Cloud Computing. Communications," *IEEE Systems Journal*, 5(4), 574-583.

Appendices

Table 1. The Cost Effectiveness of Prevention Approaches to Online Learning Integrity

Integrity solution	Student costs	Faculty costs	Institution costs	Effectiveness
Honor code	Annoyed with frequent reminder of the code. (Vandehey et al., 2007)	Extra work in reminding students about the code. (Chiesl, 2007)	Extra work in enforcing the code consistently. (Caldwell, 2009; Baron and Crooks, 2005)	A weak evidence of students' commitment to honor the code. No preventing of impersonation. (LoSchiavo & Shatz, 2011; Hart & Morgan, 2009; Kitahara and Westfall, 2007)
Authenticity statement	Annoyed with frequent signing of statements. (Vandehey et al., 2007)	Extra work in preparing and collecting the statement. (Caldwell, 2009)	Extra work in enforcing the statement consistently. (Caldwell, 2009)	A weak evidence of students' honesty. No preventing of impersonation. (Hart & Morgan, 2009; Mastin et al., 2009)
Course re-design	None (Caldwell, 2009; Chiesl, 2007)	Extra work in re-designing and delivering the course. (Hart & Morgan, 2009; McNabb and Olmstead, 2009)	Extra work in enforcing the solution consistently. (Caldwell, 2009)	A weak assurance of integrity. No preventing of impersonation. (Hart & Morgan, 2009; Rowe, 2004)
User id and password	Annoyed with frequent updates of a strong password. (Farcasin and Chan-tin, 2015)	None (Shay et al., 2010; Inglesant and Sasse, 2010)	Extra work to securely store, match, and update a user's id and password. (Shay et al., 2010; Inglesant and Sasse, 2010)	A strong evidence of user identity confirmation. No preventing of impersonation. (Ullah et al., 2012; Bailie & Jortberg, 2009)
Challenging or security questions	Annoyed with frequent questionings. (Hart & Morgan, 2009; Just & Aspinall 2009)	None (Just & Aspinall 2009)	Extra work to securely store, match and update a user's challenging questions. (Baili & Jortberg, 2009)	A strong evidence of user identity confirmation. No preventing of impersonation. (Ullah et al., 2012; Bailie & Jortberg, 2009)

Table 2. The Cost Effectiveness of Enforcement Approaches to Online Learning Integrity

Integrity solution	Student costs	Faculty costs	Institution costs	Effectiveness
Biometrics	Extra device to read biometrics. Privacy concerns. (Ullah et al., 2012; Rodchua et al., 2011; Bailie & Jortberg, 2009)	None (Bedford et al., 2011)	Extra work to securely store and match a user's biometrics. (Bailie & Jortberg, 2009)	A strong evidence of user identity and authenticity confirmation. No preventing of impersonation after login. (Bedford et al., 2011; Dunn et al., 2010)
Biometrics re-authentication	Extra device to read biometrics. Privacy concerns. Annoyed with frequent re-authentications. (Apamap et al., 2010)	None (Bedford et al., 2011)	Extra work to securely store and match a user's biometrics. Extra work to process a random re-authentication. (Moini and Madni, 2009)	Prevention of impersonation. No prevention of collusion. (Apampa et al., 2010; Moini & Madni, 2009)
Video monitoring	Extra device to record video. Privacy concerns. (Rodchua et al., 2011; Bedford et al., 2009; Hart & Morgan, 2009)	Extra work to analyze video footage. (Apampa et al., 2010; Bedford et al., 2009)	Extra work and costs to securely store and retrieve a user's video footage. (Bedford et al., 2011)	Prevention of impersonation. Prevention of collusion. (Bedford et al., 2011)
Face-to-face proctoring	Extra effort to be physically present at an agreed time and place. Extra cost for taking proctored exams. (McNabb, 2010; Bailie & Jortberg, 2009; Hart & Morgan, 2009)	Extra work to arrange for proctoring. (Bailie & Jortberg, 2009)	Extra work and cost to provide a testing center or endorse a trustworthy third party provider. (Bailie & Jortberg, 2009)	Prevention of impersonation and collusion only if the proctor is trustworthy. (Kirkpatrick, 2015)
Virtual proctoring	Extra cost for taking proctored exams. Extra cost for proctoring equipment. Privacy concerns. (Kirkpatrick, 2015; Rodchua et al., 2011)	Extra work to arrange for proctoring. (Kirkpatrick, 2015)	Extra work and cost to provide a proctoring center or endorse a trustworthy third party provider. (Kirkpatrick, 2015)	Prevention of impersonation and collusion only if the provider is trustworthy. (Kirkpatrick, 2015; Bedford et al., 2011)

Examining the Effect of Proctoring on Online Test Scores

Helaine Mary Alessio, Nancy Malay, Karsten Maurer, A. John Bailer, and Beth Rubin
Miami University

Abstract

Online education continues to grow, bringing opportunities and challenges for students and instructors. One challenge is the perception that academic integrity associated with online tests is compromised due to undetected cheating that yields artificially higher grades. To address these concerns, proctoring software has been developed to address and prevent academic dishonesty. The purpose of this study was to compare online test results from proctored versus unproctored online tests. Test performance of 147 students enrolled in multiple sections of an online course were compared using linear mixed effects models with nearly half the students having no proctoring and the remainder required to use online proctoring software. Students scored, on average, 17 points lower [95% CI: 14, 20] and used significantly less time in online tests that used proctoring software versus unproctored tests. Significant grade disparity and different time usage occurred on different exams, both across and within sections of the same course where some students used test proctoring software and others did not. Implications and suggestions for incorporating strategic interventions to address integrity, addressing disparate test scores, and validating student knowledge in online classes are discussed.

Keywords: online education, academic integrity, online testing, proctoring software, online course grades

Alessio, H. M., Malay, N., Maurer, K., Bailer, A. J., & Rubin, B. (2017). Examining the effect of proctoring on online test scores, *Online Learning, 21* (1), 146-161. doi: 10.24059/olj.v21i1.885

Introduction

A recent analysis of Integrated Postsecondary Education Data System (IPEDS) data stated that about 5.3 million students, representing more than 25% of total college enrollment, took at least one online class in 2013 (Allen & Seaman, 2015). The increased popularity of online classes presents benefits and challenges to students, faculty, and academic institutions. Geographic locations and time zones no longer present obstacles for students to enroll in a class since online classes can be delivered nearly anywhere in the world with an internet connection.

This provides students an opportunity to advance in their studies while working, traveling, and attending to family responsibilities. In this paper, data is presented for a popular online elective class with an initial enrollment of 180 students that compares proctored and unproctored tests within and across class sections. The effect of proctoring was large enough to suggest an impact on test time and scores.

Literature Review

The credibility of online classes faces criticism due to the distance between students and instructors that may contribute to breaches in integrity (Moten, Fitterer, Brazier, Leonard, & Brown, 2013). Researchers contend that online programs must address student integrity; the use of proctoring software is one way to do so, to try to assure that students are being fairly and effectively evaluated. Moten and colleagues explained that in online courses, students work in relative autonomy and anonymity and instructors may not be certain who is taking exams or how best to validate learning (2013). In addition, Berkey and Halfond (2015) have examined the sensitive subject of cheating in online courses, and found an alarming 84% of 141 students who responded to their survey agreed that student dishonesty in online test taking was a significant issue. Yet, less than half the students surveyed indicated they had ever used proctoring software in online tests.

In a study by King, Guyette, and Piotrowski (2009), 73% of 121 undergraduate students surveyed felt it was easier to cheat online compared to a traditional face-to-face classroom. When asked if they were likely to cheat, a survey of 635 students found that nearly one out of three would consider cheating in any environment and students indicated that they were more than four times as likely to cheat in an online class (Watson & Sottile, 2010). However, the same survey found no significant differences in student descriptions of cheating behavior in online and face-to-face classes (Watson & Sottile, 2010).

Many studies that address the prevalence of cheating on line vs. to face-to-face classes, many of these studies relied on student self-reports (Guyette & Piotrowski, 2009; Stuber-McEwen, Wisely, & Hoggatt, 2009; Etter, Cramer, & Finn, 2007; Watson & Sottile, 2010). Research focusing on actual student behavior has found conflicting results. For example, Ladyshewsky (2015) analyzed graduate student test scores and found no difference between the test scores in unproctored online tests when compared to face-to-face, proctored tests. Similarly, Yates and Beaudrie (2009) found no differences in course grades between community college students who took monitored versus unmonitored exams. Beck (2014) extended this work to examine scores on specific tests, where steps to reduce cheating such as randomizing the order of questions, having a single question on each page, and only allowing forward progress through the tests were used. Beck also found no differences between undergraduate student grades on monitored versus unmonitored tests (2014).

Other studies have found rampant cheating. For example, one large-scale study of cheating in online courses and work tasks found that between 26% and 34% of students cheated by looking up answers online, as did 20% of contract employees (Corrigan-Gibbs, Gupta, Northcutt, Cuttrell & Thiess, 2015). This innovative study used multiple techniques to identify cheating, including: 1) planting a fake resource that appeared in Google search engines when the exact wording of the question was entered; 2) expert analysis of wording, comparing student responses to one another as well as to common website language focusing on idiosyncratic language; and 3) tracking of IP addresses. However, unlike a typical university class, both samples involved a degree of anonymity: the class was a massive open online course aimed at undergraduate engineering students in India, and the contract employees were identified and assigned the work through a crowdsourcing work platform.

In summary, when clear-cut differences in test scores occur in separate sections of the same course or when a test is taken under contrasting conditions, questions arise about potential underlying reasons for grade disparities. There are various strategies for addressing integrity during online tests, and the use of proctoring software is one of them (Berkey & Halfond, 2015).

Proctoring software involves two major elements. First, it activates the camera on a computer, and records the student taking the exam. This enables faculty to observe the students' behavior and identify activities that may indicate cheating such as talking to others or looking up information in books. Second, it either limits the students' ability to use their computers for other tasks by eliminating the ability to engage in activities such as copy-pasting, printing and searching the internet, or it records everything that students do on their computers, or both. Limiting students' abilities to use other tools or resources is referred to as "locking down" the computer or browser. Recordings of exams can be reviewed by the professor or teaching assistants; alternatively, they can be reviewed by employees of the proctoring vendor, either simultaneous to the exam or afterward, who mark points in the exam when possible violations of exam rules are identified.

The purpose of this study was to compare test performance of students enrolled in multiple sections of the same online class where four of the nine sections used proctoring software for at least one of their tests and the other five course sections never proctored tests. We also compared student scores in the same section with and without the use of proctoring software.

Methods

This study examined the effect of proctoring tests in an online undergraduate course, Medical Terminology (KNH 209), at Miami University, a public university located in southwestern Ohio with approximately 17,000 students. Medical Terminology is a lower level undergraduate elective class, with no pre-requisites. All university students enrolled as full or part time students are eligible to take the class. It satisfies requirements toward graduation in virtually all academic divisions. Twenty students enrolled in each of nine sections of this course, totaling 180 undergraduates with the following majors: accountancy, athletic training, biochemistry, biology, economics, English, finance, public health, media studies, kinesiology, mechanical engineering, microbiology, nutrition, political science, psychology, Spanish, and speech pathology/audiology, sport leadership and management, communication, supply chain management, and zoology. All nine instructors agreed to use common exam formats that apply concepts from WCET's best practice for online education, including timed tests, random questions from a common question pool, and responses that are in randomized order (WCET, 2009).

Of the nine sections of this course, four used proctoring software. Three instructors selected a few of the tests to be proctored using Software Secure (http://www.softwaresecure.com/), a remote proctoring software that videotapes the student in their surroundings, blocks some unauthorized activities on the computer, and records students' desktops during the test. Software Secure uses live proctors, who review the recordings after the exam and identify likely situations of cheating. Two proctors, certified by the vendor, review every test. The tool also requires students to scan the room in which they are taking their exam. One instructor had all of the tests proctored using Respondus Monitor (http://www.respondus.com/products/monitor/index.shtml), which utilizes both locking down the browser and videotaping the student taking the test.

Following the completion of the tests, videos from Software Secure were reviewed by the company to detect rule violations or suspicious activity. The instructor for the course received feedback of the review and could watch the videos at each point of a potential breach to confirm if a violation occurred. Respondus Monitor generates a set of thumbnails of the full video recording that can be reviewed and flagged by the instructor for potential violations. The instructor can click on each thumbnail to watch that segment of the full video recording of the student taking the quiz. Five instructors did not use proctoring software options, while one instructor used only Lockdown Browser (no video recording or review) for half of the tests.

Students in all nine sections were informed that tests were to be taken by themselves with no notes or other resources allowed during the test. Students in the sections that were proctored were not certain of the exact test(s) throughout the course that would be proctored prior to the start of the test. Tests varied in terms of time limits, number of questions, and proctoring, but all covered similar material, and questions were randomly drawn from a shared question bank. Table 1 provides a summary of the nine class sections and indicates which of the tests in each section were proctored.

Table 1
Four Quiz Conditions (P=proctored, U=unproctored, U/L=unproctored/lockdown) for KNH 209/Medical Terminology sections A through I in January 2015

Quiz	**Section A**	**Section B**	**Section C**	**Section D**	**Section E**	**Section F**	**Section G**	**Section H**	**Section I**
1	U	U	P	U	U	U	U	U	U
2	U	P	P	U	U	U	U	P	P
3	U	U	P	U/L	U	U	U	U	U
4	U	P	P	U/L	U	U	U	U	U

U: Unproctored
U/L: Unproctored and lockdown only, no video monitoring
P: Proctored with video monitoring (Software Secure or Respondus Monitor)

Table 2 reports the number of students who were proctored or unproctored on each quiz. Student enrollments were tracked in all sections. Following the conclusion of the course, all students were contacted about the use of their data in class with all identifiers removed, and were provided an opportunity to have their data omitted from analyses.

Table 2

Total number of quizzes that were Proctored and Unproctored in 9 sections of KNH 209/Medical Terminology in January 2015

Quiz 1	One quiz (n=14) was proctored; eight quizzes (n=148) were unproctored
Quiz 2	Three quizzes (n=48) were proctored; six quizzes (n=109) were unproctored
Quiz 3	Two quizzes (n=31) were proctored; seven quizzes (n=129) were unproctored
Quiz 4	Two quizzes (32) were proctored; seven quizzes (n=130) were unproctored

Of the initial 180 students enrolled, 22 dropped the course. Of the 158 students who completed the course, 11 did not complete all tests. The anonymized data from the 147 students who consented and had completed all four tests were then used in a statistical analysis to assess the effect of proctoring on test scores and percentage of allotted time taken.

Data Analysis

The impact of proctoring on student quiz performance was evaluated using a linear mixed effects model (Verbeke & Molenberghs, 1997; Montgomery, 2013). A realistic assumption can be made that responses from tests taken by the same student or students with the same instructor may be related; thus linear mixed effects models are used to allow for these relationships to be reflected in the correlation structure of our analysis. First we aimed to model the test score percentages to assess the effect of proctoring. Due to a concern that the difficulty of the four exams may not be uniform in difficulty of material and that the number of questions per test may have effect on test scores, we consider these as covariates in the modeling. The model selection based on the Bayesian Information Criterion (BIC) confirms the importance of accounting for these factors. The selection process yielded a model with fixed effects for tests, proctoring administration and number of questions on the test, and random effects for sections and for students within sections. The linear mixed effects model for test score percentage was parameterized as:

Model Equation 1.

$$Score_{ijk} = Test_k + \beta_L I_{ijk}(Lockdown) + \beta_P I_{ijk}(Proctored) + \beta_Q NumQ_{ik} + \delta_i + \gamma_{ij} + \epsilon_{ijk}$$

where we model the score of the k^{th} test for the j^{th} student in the i^{th} section using a cell means parameterization of test averages, μ_k, which use non-proctored exams with 20 questions as the baseline. The model terms associated with the fixed effects are defined as:

$Test_k$	Average score on test *k* with no proctoring software and 20 questions (baseline)
β_L	Additive change to baseline score when Lockdown (no video) used on test
β_P	Additive change to baseline score when video proctoring used on test
$I_{ijk}(\,.)$	Indicator function for use of proctoring software in test *k* for student *j* in section *i*
β_Q	Additive change to baseline score for every additional questions above the baseline
$NumQ_{ik}$	The number of questions beyond than the baseline of 20 on test *k* in section *i*.

The model terms i(section), ij(student) and εijk (error) are nested random effects that are specified such that:

$$Cov(Score_{ijk}, Score_{lmn}) = \begin{cases} 0, & \text{if sections } i \neq l \text{, student } j \neq m \text{, and test } k \neq n \\ \sigma^2, & \text{if sections } i \neq l \text{, student } j \neq m \text{, and test } k = n \\ \sigma^2 + \sigma_\gamma^2, & \text{if sections } i \neq l \text{, student } j = m \text{, and test } k = n \\ \sigma^2 + \sigma_\gamma^2 + \sigma_\delta^2, & \text{if sections } i = l \text{, student } j = m \text{, and test } k = n \end{cases}$$

It was also speculated that academic dishonesty on online tests may manifest as longer times taken on the tests due to the extra time spent searching through prohibited reference materials. To explore the impact of proctoring software on the time taken to complete the tests we fit a linear mixed effects model to the percentage of allotted time used. Note that the metric used in modeling differences in time usage was the percentage of allotted time used by the student; this is to maintain a consistent interpretation with different numbers of questions and time allowed across the sections. Model selection and diagnostics were run in the same fashion as in the model for test scores, and the model covariates and random effects for the selected model turn out to follow an identical structure to those in Equation (1) above. The model for percentage of time taken follows the form,

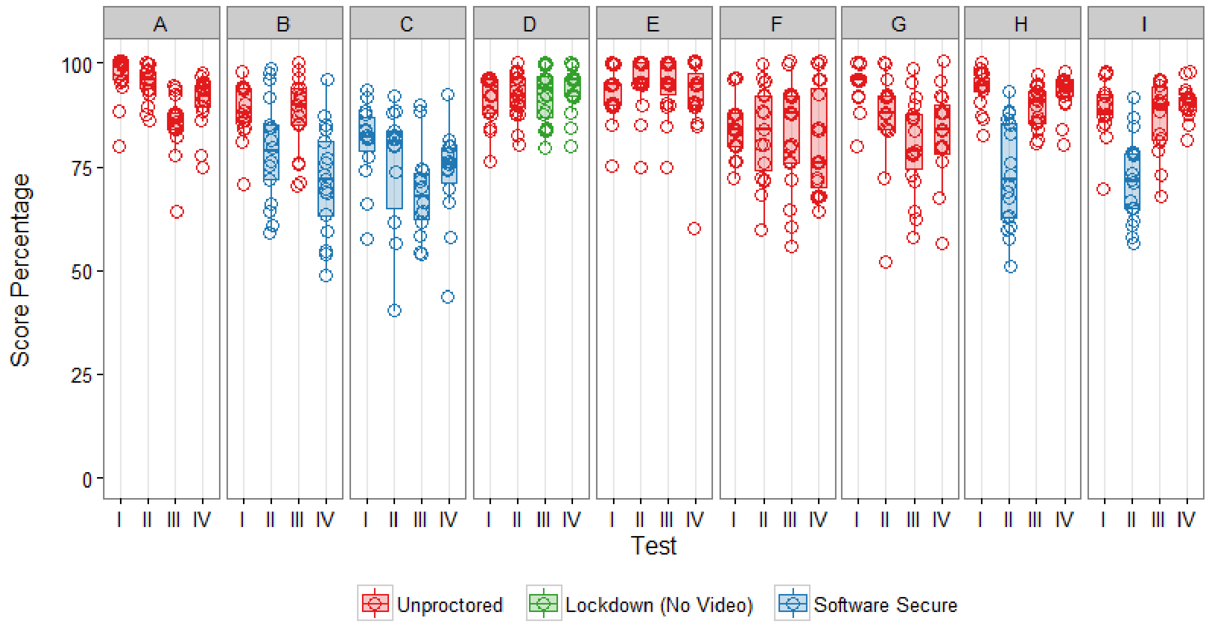
Model Equation 2.

$$Time_{ijk} = Test_k + \beta_L I_{ijk}(Lockdown) + \beta_P I_{ijk}(Proctored) + \beta_Q NumQ_{ik} + \delta_i + \gamma_{ij} + \epsilon_{ijk} \ . \ (2)$$

Data cleaning, data summaries, visual graphics and linear mixed models and diagnostic tools were created using the R software using the dplyr (Wickham & Francois, 2015), ggplot2 (Wickham, 2009) and nlme (Pinheiro, Bates, DebRoy, Sarkar & R Core Team, 2015) packages.

Results

Figure 1 visually presents the scores and times taken on tests within each class section and is colored to emphasize the proctoring status of each test group. A test was considered proctored when it included videotaping. We see that there are noticeable differences in proctored and unproctored exams, primarily that proctored exams seem to have lower scores and take a larger percentage of the allotted time. The average test scores for proctored tests was 74.3% (SD=12.3) and 89.4% (SD=9.0) for unproctored tests. The average percentage of allotted time taken on proctored tests was 20.4% (SD=13.9) and unproctored tests was 41.2% (SD=14.1); showing that students took approximately half the amount of time taking proctored test compared with unproctored tests. Note that unproctored tests with lockdown only (no video monitor) had an average score of 93.2% (SD = 5.9) and took an average of 40.0% (SD=10.1) of the time allotted; quite comparable in test scores and time used with the unproctored tests. See Table 3 for a full listing of statistics for test scores and percent of allotted time used within proctoring groups.



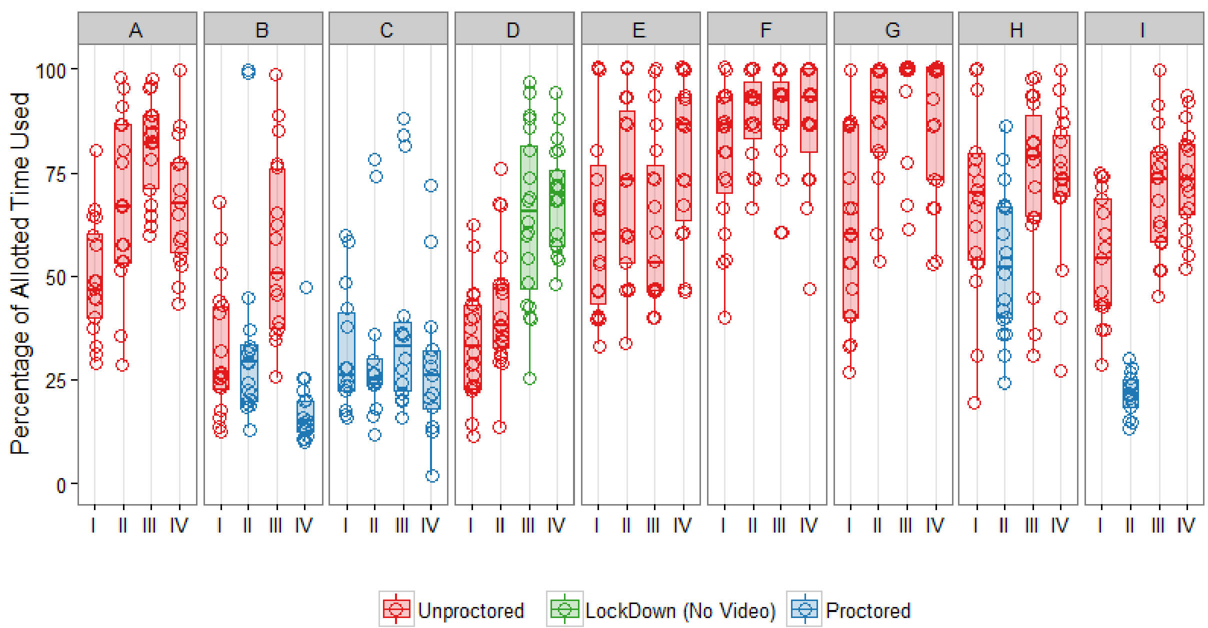


Figure 1. Test scores (%) and time used (% of allotted) in nine Sections (A-I), colored by proctoring status. Proctored tests (Blue) tended to score lower and take less time than unproctored tests (Red). Tests with Lockdown (Green) behaves similar to unproctored sections.

Table 3
Proctor Status and Average Test scores, Percent Time Used, and Number of Total Tests and Students

Proctor status	Average + [SD] Test Score (% correct)	Average + [SD] Percent Time Used (% of time given)	Number of Tests	Number of Students
Unproctored	89.4 [12.3]	41.2 [14.1]	471	147
Proctored *with* video monitor	74.3 [5.9]	20.4 [13.9]	125	66
Lockdown (no video monitor)	93.2 [9.0]	40.0 [10.1]	40	20

We turn to the fitted models discussed in the analysis section above to assess the significance of the proctoring related difference seen in the visual and numerical exploration. Table 4 shows the summary of the effects of proctoring on test scores, as estimated from the linear mixed effects model. There did not appear to be any trend or extreme outliers in the residuals, hence, the use of this model appears to be justified. The baseline means for tests 1 through 4, unproctored tests with 20 questions, were: 89.7, 87.8, 83.4, and 84.8, respectively. This accounts for general differences in difficulty, where the first two tests were less difficult than the last two tests. The differences in the test scores in the model are statistically significant ($p<0.05$). Tests proctored with the Software Secure video monitoring were found to have significantly lower test scores than unproctored test scores. The video proctored tests were found to score 17.2 percentage points (95% CI: [4.8, 19.6]) *lower* than unproctored tests. This implies a significant, and substantial, decrease in scores under video proctoring, after controlling for differences in test difficulty and number of questions. Tests that used only Lockdown (no video) were found to have a score 7.4 percentage points (95% CI: [3.9, 11.2]) *higher* than unproctored tests, after controlling for differences in test difficulty and number of questions. While this implies students taking a test using only Lockdown (no video) have a significant improvement in scores, there is only one section that implemented this technology; thus confounding the effect of lockdown and instructor.

Table 4
Fitted Coefficients and Variance Estimates for Linear Mixed Effects Model for Test Score Percentages, as Parameterized in Model Equation

Fixed Effect	Model Term	Estimate	95% CI
Test 1 baseline (unproctored/20 questions)	$Test_1$	89.77	(86.86 , 92.68)
Test 2 baseline (unproctored/20 questions)	$Test_2$	87.80	(84.46 , 91.14)
Test 3 baseline (unproctored/20 questions)	$Test_3$	83.37	(80.12 , 86.62)
Test 4 baseline (unproctored/20 questions)	$Test_4$	84.76	(81.32 , 88.20)
Lockdown (no video) effect	β_L	7.54	(3.92 , 11.15)

Proctored (Software Secure or Respondus Monitor) effect	β_P	-17.23	(-19.62 , -14.83)
Additional questions effect	β_Q	0.13	(0.05 , 0.21)
Random Effect	Variance Term	Variance Estimate	Percentage of Total Variance
Section	σ^2_δ	10.1	11.2 %
Student	σ^2_γ	28.3	31.3 %
Residual error	σ^2	52.0	57.5 %

Not only did proctoring of tests affect test scores but there was evidence that proctoring also affected how long students took to finish. Table 3 shows that when students were unproctored, including only Lockdown (no video), they used much more of their available time than if the test was proctored with video. Table 5 contains the effects from the linear mixed effects model for the percentage of allotted time taken. The baseline tests, with 20 unproctored questions, show that students tended to take more time to complete the later exams. There was no significant effect of number of questions on the percentage of allotted time taken, indicating that the time per questions was sufficiently similar to allow comparison across sections. Lastly, the proctored group took an estimated 30.5 percent less of the time allotted (95% CI: [25.4 , 35.7]) in completing their exams than the unproctored students.

Table 5

Fitted Coefficients and Variance Estimates for Linear Mixed Effects Model for Percentage of Allotted Time Taken on Tests, as Parameterized in Model Equation 2

Fixed Effect	Model Term	Estimate	95% CI
Test 1 baseline (unproctored/20 questions)	$Test_1$	56.10	(46.05 , 66.16)
Test 2 baseline (unproctored/20 questions)	$Test_2$	69.67	(58.87 , 80.48)
Test 3 baseline (unproctored/20 questions)	$Test_3$	71.43	(60.80 , 82.06)
Test 4 baseline (unproctored/20 questions)	$Test_4$	70.50	(59.50 , 81.49)
Lockdown (no video) effect	β_L	19.75	(11.95 , 27.57)
Proctored (Software Secure or Respondus Monitor) effect	β_P	-30.53	(-35.69 , -25.36)
Additional questions effect	β_Q	-0.01	(-0.20 , 0.18)

Random Effect	Variance Term	Variance Estimate	Percentage of Total Variance
Section	σ^2_δ	184.51	34.2 %
Student	σ^2_γ	130.33	24.2 %
Residual Error	σ^2	224.17	41.6 %

The results of the linear mixed effects models for test scores and percentage of allotted time used show us that unproctored tests had significantly higher scores and took significantly more time than proctored test, while controlling for test ordering and number of questions. We see the dramatic difference in testing behavior in Figure 2 which contains the scatterplot of test scores and percentage of allotted time taken, colored by proctoring status. These finding are consistent with the suspicion that academic dishonesty, in the form of students searching through prohibited reference materials during the test, is more prevalent on unproctored exams.

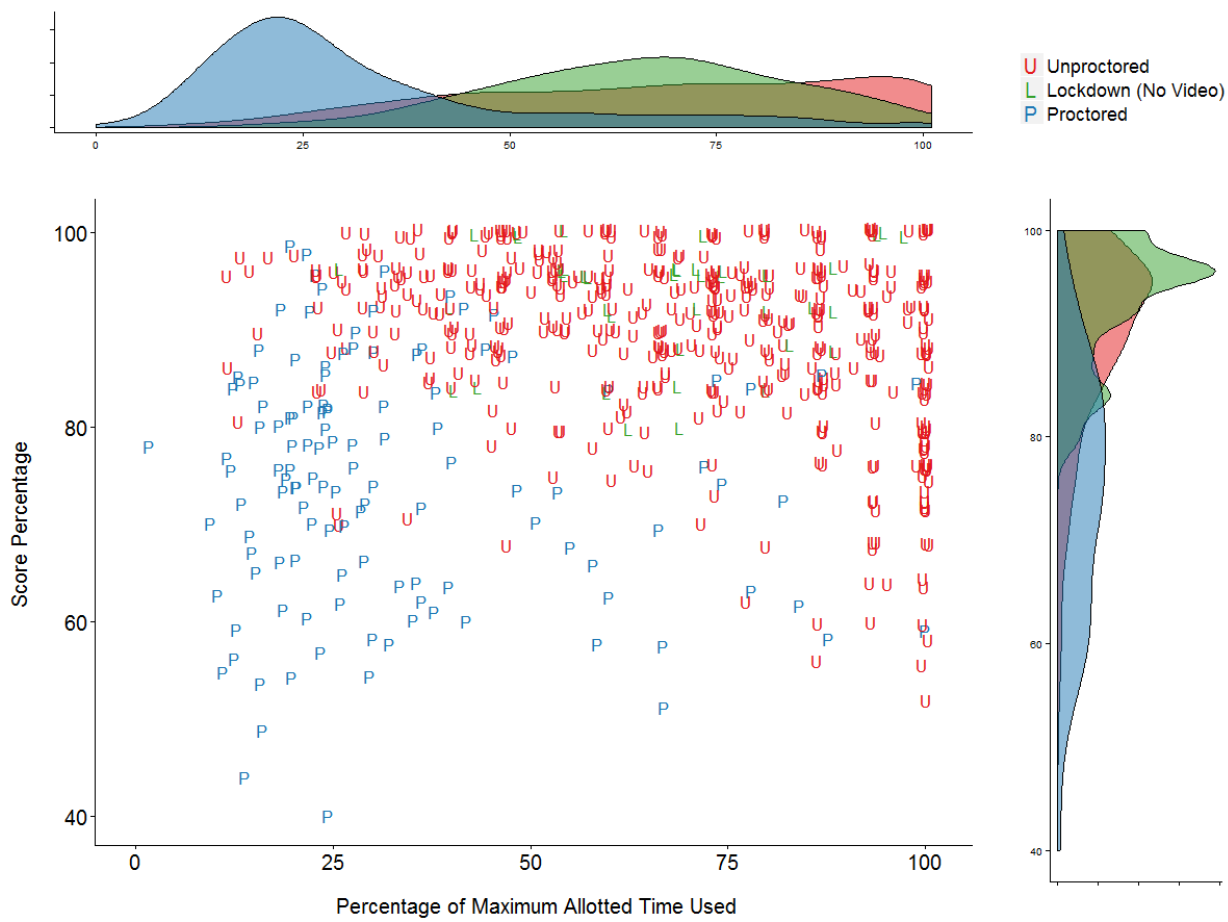


Figure 2. Plot of test score (%) vs. amount of time used (% max) for all sections combined. Points correspond to students in sections that were proctored using Software Secure or Respondus Monitor with Lockdown (Blue P), used Lockdown alone (Green L) or unproctored

(Red U). Once again we see that the tests behaviors are similar for lockdown and unproctored tests.

Discussion

These results indicate clear and significant grade disparities in comparing test scores when the students took online tests that were proctored with video monitoring versus unproctored or unproctored with Lockdown (no video). Proctored test (with video) scores are significantly lower than unproctored test scores. The model fit also shows the same results as the difference between proctored and unproctored test scores is between 14 and 20 points, the difference of one or two letter grades. This difference occurred in students between multiple sections of the same course as well as within sections when the same students took tests proctored versus unproctored.

Test scores are not the only component factoring into student grades, as forum postings, case studies, homework assignments, blogs, and other types of work all contributed to the final grade in this course. Nevertheless, the striking difference in scores from proctored versus unproctored tests appeared to factor significantly into final grades as evidenced by the different final grade distributions. Sixty three percent of all students in sections with only unproctored tests earned an A, whereas 17% of all students in sections with proctored tests earned an A.

Another concern is the difference in attrition between the sections that offered proctored versus unproctored tests. Only seven of the 100 students initially enrolled in sections with unproctored tests dropped the class compared with 15 of 80 students initially enrolled in sections with proctored tests who dropped. Although no inquiries were made as to why students dropped the class, more than twice as many students in the proctored group dropped compared with the unproctored group.

Bunk, Li, Smidt, Bidetti, and Malize (2015) explored faculty perceptions in explaining negative attitudes toward online classes. While proctoring and academic honesty were not directly mentioned, faculty did express concern about compromised educational quality in online classes. In a study on student and faculty views of academic dishonesty and online learning, both faculty and students agreed that it would be easier to cheat in online classes (Kennedy, Nowak, Raghuraman, Thomas, & Davis, 2000). Methods suggested by faculty to counter cheating, included supervised final exams counting for a high percentage of the course grade, changing assignments each semester, using personalized assignments, verification software, and using open-book exams. Proctoring software was not mentioned by faculty, because it was not commonly available at the time.

A study by Spaulding (2009) did not provide compelling evidence for an increased prevalence of academic dishonesty in online vs. traditional classes, which may lead many faculty to underestimate the frequency of academic dishonesty in their classes. Given this perception, Hard, Conway, and Moran (2006) reported that faculty members who perceive academic dishonesty as rare do not actively work to prevent it. Investigating academic integrity is complicated, whether in traditional or face-to-face learning and testing environments. Student perception of cheating online may be different than in a face-to-face situation (Rains, et al.,

2011) even when instructions clearly state otherwise. The potential for academic dishonesty (Corrigan-Gibbs et al., 2015; Jones, Blankenship, & Hollier, 2013; Moten, et al., 2013) and the perception that cheating occurs more frequently in online classes (Grijalva, Nowell, & Kerkvliet, 2006; Raines et al, 2011) present challenges to all stakeholders. Much research argues that cheating is prevalent in online courses, but few studies measure actual cheating behavior. Some found evidence of significant cheating in online tests (Corrigan-Gibbs, et al., 2015), while others did not (Ladyshewsky, 2015). The current study did not assess cheating behavior. Instead, it compared test scores when students used proctoring software with those that were unproctored. Disparate test grades imply that cheating likely occurred when student tests were unproctored, especially given the large and statistically significant grade difference of 17 points, representing an average difference of two letter grades between scores on tests when proctoring software was used versus when it was not.

This study provides substantive evidence of disparate test results in online courses, as indicated by significantly higher scores both within classes and across class sections on unproctored versus proctored online tests. After controlling for the effects of test difficulty and student and teacher differences, students taking proctored online exams scored approximately 17 points lower out of 100 when compared to unproctored students. The different scores approximated an average test grade of A to A- on unproctored tests and C to C- on proctored tests. Furthermore, when unproctored, students took significantly more time to complete tests. It is possible that students used the extra time to look up answers, despite the application of testing best practices of providing limited time, randomized selection of items, and instructions stating that using resources during a test was not allowed.

This potential for academic dishonesty cannot be ignored (Harbin & Humphrey, 2013). Previous research on student perceptions about whether they felt they might cheat in online versus face-to-face test conditions have been inconsistent, however, it appears that in this current study, the finding by Watson and Sottile (2010), where students indicated they would be more than four times likely to cheat in an online class, seemed to have occurred, with the grade distribution indicating that students taking unproctored online tests were four times more likely to receive a grade of A compared with students who took proctored online tests. Concerns about the integrity of online courses due to cheating and fraud have reached the popular press (Newton, 2015). There are real consequences for students who cheat, who may not learn critical content for thinking, problem solving, and foundational information required for upper level course work. Additionally, the reputation of faculty and institutions and student learning are compromised when acts of cheating are not addressed. Faculty and institutions will need to confront the likelihood that breaches in academic honesty occur in all class formats. In online classes, in particular, proactive interventions that include proctoring software with video monitoring may deter cheating and protect academic integrity.

Limitations of this Study

It is important to consider potential limitations to the generalizability of these results. This was a class of medical terminology, which requires that technical terms be memorized and accurately applied, and where assessment included multiple choice tests. It is not clear that the size of the effect would be as large with courses that do not involve timed, closed-ended tests. In addition, these were classes populated by traditional students in a Midwestern, public university. Again, the potential for generalizability to other populations may be low.

Conclusions and Future Studies

Students enrolled in online courses in which at least one online test was proctored with video monitoring scored on average 17 points lower than students enrolled in the same courses with no test proctoring. The effect of *proctoring with video* is large enough to suggest that an impact on test scores exists, with the likelihood that when unproctored, students may resort to academic dishonesty by using resources that were explicitly forbidden during the test. The effect of proctoring with video shows a potential effect on the percentage of test time used to take the test, with proctoring resulting in less time compared with unproctored tests, where students took significantly more time to complete the test. Additionally, lockdown software without video monitoring, did not have a similar impact as proctoring software that used video monitoring. Proctoring with video monitoring significantly negatively impacts online test grades, probably because it deters cheating, and its use is important to assure academic integrity through similar test taking conditions in similar courses when using online tests.

It would be interesting to replicate this study or use a randomized design in other courses and at other universities. In addition, the different proctoring tools themselves could be examined. As online test proctoring becomes more common, faculty and students may learn about advantages and disadvantages of different vendors and systems. For example, it may be fruitful to examine possible differences between vendors that employ human proctors as opposed to fully-automated proctoring systems. While future research may affect the proctoring choices, these results point to the need for proctoring software to contribute to the integrity of online testing.

Acknowledgements: The authors thank the STA 660 data practicum class at Miami University, and, in particular, Daniel Garbinsky, Mitch Beebe, Sambed Adhikari, and Travis Sellers, for data analyses on this project.

Correspondence concerning this article should be addressed to Beth Rubin, Assistant Provost for eLearning, Miami University, Oxford Ohio 45056. Contact: rubinb@miamioh.edu

References

Allen, I.E., and Seaman, J. (2015). Grade level: Tracking online education in the United States. Babson Survey Research Group and Quahog Research Group, LLC. Last access on February 3, 2016: http://www.onlinelearningsurvey.com

Beck, V. (2014). Testing a model to predict online cheating: Much ado about nothing. *Active Learning in Higher Education, 15(1),* 65-75.

Bunk, J., Li, R., Smidt, E., Bidetti, C., and Malize, B. (2015. Understanding faculty attitudes about distance education: The importance of excitement and fear. *Online Learning, 19(4), 1-11.*

Berkey, D., and Halfond, J. (2015). Cheating, student authentication and proctoring in online programs. *New England Journal of Higher Education*, July 20. Retrieved from http://www.nebhe.org/thejournal/cheating-student-authentication-and-proctoring-in-online-programs

Corrigan-Gibbs, H., Gupta, N., Northcutt, C., Cutrell, E., and Thies, W. (2015). Deterring cheating in online environments. *ACM Transactions on Computer-Human Interaction, 22*(6), Article 28. Last access on February 3, 2016: DOI: http://dx.doi.org/10.1145/2810239

Etter, S., Cramer, J.J., and Finn, S. (2007). Origins of academic dishonesty: Ethical orientations and personality factors associated with attitudes about cheating with information technology. *Journal of Research on Technology in Education*, *39(2),* 133-155.

Grijalva, T.C., Nowell, C., and Kerkvliet, J. (2006). Academic honesty and online courses. *College Student Journal*, *27(3)*, 180-185.

Harbin, J. L., and Humphrey, P. (2013). Online cheating: The case of the emperor's clothing, elephant in the room, and the 800 lb. gorilla. *Journal of Academic and Business Ethics*, *7*, 1-6.

Hard, S.F., Conway, J., and Moran, A.C. (2006). Faculty and college student beliefs about the frequency of student academic misconduct. *The Journal of Higher Education, 77(6), 1058-1080.*

Jones, I.S., Blankenship, D., and Hollier, G. (2013). Am I cheating? An analysis of online students' perceptions of their behaviors and attitudes. *Psychology Research, 3*(5), 261-269.

Kennedy, K., Nowak, S., Raghuraman, R., Thomas, J., Davis, S.F. (2000). *College Student Journal, 34(2), 309*.

King, C.G., Guyette, R. W., & Piotrowski, C. (2009). Online exams and cheating: An empirical analysis of business students' views. *The Journal of Educators Online, 6, 1*, 1-11.

Ladyshewsky, R.K. (2015). Post-graduate student performance in 'supervised in-class' versus 'unsupervised online' multiple choice tests: implications for cheating and test security. *Assessment and Evaluation in Higher Education*, 40(7), 883-897. DOI: 10.1080/02602938.2014.956683

Montgomery, D. (2013). Experiments with Random Factors. In *Design and Analysis of Experiments* (8th ed.). New York: Wiley.

Moten Jr., J., Fitterer, A., Brazier, E., Leonard, J., and Brown, A. (2013). Examining online college cyber cheating methods and prevention methods. *The Electronic Journal of eLearning, 11, (2),* 139-146.

Newton, D. (2015). Cheating in online classes is now big business. *The Atlantic, 4*. Retrieved from http://www.theatlantic.com/education/archive/2015/11/cheating-through-online-courses/413770

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., and R Core Team (2015). *nlme: Linear and Nonlinear Mixed Effects Models.* R package version 3.1. http://CRAN.R-project.org/package=nlme

Raines, D.A., Ricci, P., Brown, S.L., Eggenberger, T., Hindle, T., and Schiff, M. (2011). Cheating in online courses: The student definition. *The Journal of Effective Teaching, 11*(1), 80-89.

SAS Institute Inc. 2012. *Base SAS® 9.4* Cary, NC: SAS Institute Inc.
Spaulding, M. 2009. Perceptions of academic honesty in online vs. face-to-face classrooms. *Journal of Interactive Online Learning, 8(3), 183-198.*

Stuber-McEwen, D., Wisely, P., and Hoggatt, S. (2009). Point, click, and cheat: Frequency and type of academic dishonesty in the virtual classroom. *Online Journal of Distance Learning Administration,* 12(2). Retrieved from http://www.westga.edu/~distance/ojdla/fall123/stuber123.html

Verbeke, G. and Molenbergh, G. (1997). *Linear Mixed Models in Practice: A SAS-Oriented Approach.* New York, NY: Springer.

Watson, G., and Sottile, J. (2010). Cheating in the digital age: Do students cheat more in online courses? *Online Journal of Distance Learning Administration, 13*(1) Retrieved from http://www.westga.edu/~distance/ojdla/spring131/watson131.html

WCET (2009). *Best practice strategies to promote academic integrity in online education.* WCET, UT TeleCampus, and Instructional Technology Council. Retrieved from http://wcet.wiche.edu/sites/default/files/docs/resources/Best-Practices-Promote-Academic-Integrity-2009.pdf

Wickham,H, and Francois, R. (2015). *dplyr: A grammar of data manipulation*. R package version 0.4.3. Retrieved from. http://CRAN.R-project.org/package=dplyr

Wickham, H. (2009) *ggplot2: Elegant graphics for data analysis.* New York, NY: Springer.

Yates, R.W., and Beaudrie, B. (2009). The impact of online assessment on grades in community college distance education mathematics courses. *American Journal of Distance Education, 23*, 62–70.

SECTION IV:
Students, Community, and Online Learning

Creating a Community of Inquiry in Large-Enrollment Online Courses:
An Exploratory Study on the Effect of Protocols within Online Discussions
Baiyun Chen, Aimee deNoyelles, Janet Zydney, Kerry Patton

Exploring Small Group Analysis of Instructional Design Cases in
Online Learning Environments
Jesus Trespalacios

Utilization of an Educational Web Based Mobile aApp for Acquisition and
Transfer of Critical Anatomical Knowledge, Thereby Increasing Classroom and
Laboratory Preparedness in Veterinary Students
Kevin Hannon

A Critical Review of the Use of Wenger's Community of Practice (CoP)
Theoretical Framework in Online and Blended Learning Research, 2000-2014
Sedef Uzuner Smith, Suzanne Hayes, Peter Shea

Institutional Factors for Supporting Electronic Learning Communities
Jayme N. Linton

Adapting for Scalability: Automating the Video Assessment of Instructional Learning
Amy M Roberts, Jennifer LoCasale-Crouch, Bridget K Hamre, Jordan M Buckrop

Creating a Community of Inquiry in Large-Enrollment Online Courses: An Exploratory Study on the Effect of Protocols within Online Discussions

Baiyun Chen & Aimee deNoyelles
University of Central Florida

Kerry Patton & Janet Zydney
University of Cincinnati

Abstract

It can be difficult to foster focused and effective communication in online discussions within large classes. Implementing protocols is a strategy that may help students communicate more effectively, facilitate their learning process, and improve the quality of their work within online discussions. In this exploratory research study, a protocol was developed and improved over two iterations in a very large undergraduate video-streaming business course (N_1=412; N_2=450). The discussion instructions were consolidated and adjusted, and design elements such as a grading rubric, exemplary student samples, and due date reminders were added in the second iteration. There were higher perceptions of social, cognitive, and teaching presences in the second iteration, as well as significantly more group cognition within the discussion measured through a Community of Inquiry coding template. Findings suggest that protocols are a potentially useful strategy to manage online discussions in large classes.

Keywords: Community of Inquiry; large classes; online discussions; protocols

Chen, B., deNoyelles, A., Patton, K., & Zydney, J. (2017). Creating a community of inquiry in large-enrollment online courses: An exploratory study on the effect of protocols within online discussions, *Online Learning 21*(1), 165-188. doi: 10.24059/ olj.v21i1.816

Introduction

Asynchronous discussions play an important role in online learning by providing a space for instructors and students to form a community, to engage in dialogue about the course content, and to co-construct knowledge (Gao, Zhang, & Franklin, 2013). Because of the asynchronous nature of the discussions, participants have more time to think before responding, and the act of writing elicits the formation of new knowledge and ideas (Hew, Cheung & Ng, 2010). In addition, discussion forums provide a permanent space for the participants to return to their original contributions, promoting reflection (Hew et al., 2010) and self-assessment (Gao et al., 2013).

Despite these affordances, it can be challenging to create and sustain focused, in-depth online discussions (Gao et al., 2013). In a literature review, Hew et al. (2010) identified several reasons why students do not fully participate in online discussions. Students will not fully participate when they do not see the purpose of the discussion, do not understand what to contribute, do not receive responses to their posting, and cannot make sense of the discussion due to the structure of the online forum. Therefore, online discussions must be structured in a way that clearly communicates their purpose and student expectations, encourages students to co-construct knowledge, and facilitates meaningful discussion. Past research has specifically identified the establishment of clear communication protocols and requirements for participating as vital for a successful discussion (Brannon & Essex, 2001; Darabi, Liang, Survavanshi, & Yurekli, 2013; Makitalo, Weinberger, Hakkinen, Jarvela, & Fischer, 2005). An increasingly common constraint, however, is class size. While an optimal online class size is between 12 to 16 students for effective communication (Orellana, 2006; Tomei, 2006), online courses can sometimes hold hundreds of students. Even when sectioned into groups, it can be difficult for instructors to effectively facilitate online discussions due to logistical and grading issues.

The purpose of this paper is to explore the use of a discussion strategy called a protocol, which may potentially address the challenge of stimulating a productive discussion in a large online class. Protocols explicitly structure meaningful conversations with the purpose of stimulating student problem solving, reflection, and elicitation of support from others (McDonald, Zydney, Dichter, & McDonald, 2012). Protocols establish a well-defined goal, clear roles, and set rules for interactions, and specific deadlines for posting. A notable example of a protocol is called the Tuning protocol (McDonald et al., 2012). The goal of this protocol is to improve ("tune") a particular work in progress. For this protocol, a participant shares the work in progress and has the opportunity to ask the others for relevant feedback. Others listen about the work in progress, and then ask clarifying and probing questions before supplying focused feedback. Finally, the participant reflects on the feedback and brainstorms ways to improve the work. This participant also listens and provides feedback to the others. In this way, protocols sharpen communication, enhance collective thinking, and build knowledge. The Tuning protocol exhibits the four core characteristics of a protocol that set it apart from other structured approaches (McDonald et al., 2012). First, all participants have a voice and play a designated role. Second, all participants engage in different and varied ways. Third, all participants have the dual roles of both creating as well as reading text. Finally, protocols foster trust since the norms are well established.

Protocols first started within face-to-face environments, but are now being explored in the online environment (McDonald et al., 2012). This study explored the use of protocols within online discussions in a large enrollment college course over a two-semester period. The explicit structure of a protocol may keep students in a large online class on track and focused on a goal, sustain the community, and eliminate the need for constant facilitation from the instructor. At the conclusion of the paper, specific enhancements that enable a protocol-based discussion to logistically work in a large online class are proposed.

Theoretical Background: Community of Inquiry

The Community of Inquiry (CoI) framework has been used as one of the premier theories in the last two decades to conceptualize community in many online discussion research studies (see CoI website). The framework proposes three essential presences that contribute to a successful educational experience: social, cognitive, and teaching (Garrison, Anderson & Archer, 2000; Garrison, Anderson, & Archer, 2010). Social presence is the ability of learners to project themselves socially and emotionally, being perceived as "real" people in mediated communication (Garrison & Arbaugh, 2007; Rourke, Anderson, Garrison & Archer, 1999). Cognitive presence refers to the extent to which learners are able to construct and confirm meaning through sustained reflection and discourse (Garrison, Anderson & Archer, 2001 & Garrison & Arbaugh, 2007). Finally, teaching presence is defined as the "design, facilitation, and direction of cognitive and social processes for the purpose of realizing personally meaningful and educationally worthwhile learning outcomes" (Anderson, Rourke, Garrison, & Archer, 2001, p. 5).

Using these presences as a framework, an effective online discussion will encourage students to not only communicate on a social level but also engage in an academic conversation, while being purposely directed to achieve learning outcomes.

Literature Review

Class Size

There is no standardized definition of what constitutes a "large class." Maringe and Sing (2014) define large class size as "any class where the numbers of students pose both perceived and real challenges in the delivery of quality and equal learning opportunities to all students, in that classroom" (p. 763). The researchers adopt Maringe and Sing's definition for purposes of this paper and regard our target classes of over 400 students as large.

In the literature, class size was found to be a factor in students' interactions in online courses (Hewitt & Brett, 2007; Orellana, 2006). The larger the class size is, the less likely students will actively engage in online activities (Rocca, 2010). Consequently, class size also negatively affects students' social engagement as well as creates information overload for students (Hewitt & Brett, 2007). Past research has found that to maintain sufficient instructor-student interactions, the ideal online class size was between 12 to 16 students to achieve the highest level of interactions (Orellana, 2006; Tomei, 2006). Due to student enrollment and administrative pressure, however, class sizes have continued to increase in American colleges in recent years. In this study's school, a regular general education course can be over one hundred students.

A search via Google Scholar and the Education Research Information Center (ERIC) database using the keywords "large class" and "online discussion" was conducted. Unfortunately, only a few journal articles were found on the topic of facilitating discussions in large online classes (see Yang, 2007). For large classes, it is recommended to section students into small groups to achieve high quality interactivity (Kim, 2013). Hew and Cheung (2011) found that as discussion group size increased up to 10 students, the level of knowledge construction also increased. Although prior research does identify strategies that are effective in eliciting knowledge in online discussions (see deNoyelles, Zydney, & Chen, 2014), these studies often do not sample large enrollment online classes. Thus, more research is needed to determine how to meet the challenge instructors face in creating an effective online discussion activity in large undergraduate classes.

Online Discussions and Protocol-based Discussions

Research about online discussions in large classes is rare (Eyitayo, 2005). One study found that small group, online discussions can help students develop their critical thinking skills with skillful facilitation in a class of 133 students (Yang, 2007). The instructors divided the class into smaller groups and assigned five teaching assistants to be group facilitators who used Socratic questions. Unfortunately, most instructors who teach large classes are not equipped with multiple teaching assistants to assist with discussion facilitation. Given the constraints, is it still possible to use online discussions as a teaching/engagement strategy in large online classes?

In a review of literature, deNoyelles, Zydney & Chen (2014) identified discussion strategies associated with higher levels of the CoI presences. One effective strategy noted was the use of protocols; however, empirical research on protocols in online discussions is scarce. One study that compared a protocol and a non-protocol online discussion in two small sections of a graduate course found that the three CoI presences exhibited by students were more evenly distributed in protocol-based discussions (Zydney, deNoyelles & Seo, 2012). The protocol promoted more shared group cognition, rather than individual monologues. It also significantly increased students' opportunity to participate in the instructional design of the course, making teaching presence a shared responsibility between teachers and students. The findings indicated that the shared responsibility reduced the burden on the instructor for doing all the facilitation and enabled the instructor more time to diagnose misconceptions and inject knowledge when necessary.

Skillful facilitation is especially difficult in large online classes. Considering the benefits discovered from prior research on protocols, the researchers expected that protocol-based discussions could benefit large online classes where extensive instructor facilitation is not feasible. Therefore, our research focused on how to utilize one specific discussion strategy, the protocol, in order to improve students’ perception and enactment of CoI in online discussions within large classes.

Methodology

This study examined what enhancements are needed to enable a protocol-based discussion to logistically work in a large class, and whether the enhancements made in the second design iteration of the protocol-based online discussions improved the CoI for a large enrollment course. Two research questions were posed:

(1) How did students in a large enrollment online course perceive the enhancements made to the protocol-based discussions?
(2) How did the enhancements made to the protocol-based discussions impact the elements that contribute to a CoI?

The Context of the Study

Online discussion protocols were developed and implemented in a large undergraduate business course (GEB3113) over two semesters (Iteration 1 and 2) in 2014, taught by the same instructor at the University of Central Florida. The course was classified as "video streaming," meaning that face-to-face attendance was optional, with the sessions being streamed to an online audience. Iteration 1 included 412 participants and Iteration 2 included 450. In both iterations, students regularly engaged in graded online discussions in groups of approximately 10 students (called "Bazinga circles"), resulting in at least 40 groups per semester.

Three protocol-based discussions were implemented in the course, each focused on written class assignments, with the goal of encouraging students to reflect and give meaningful feedback about each other's work. The protocol featured in this study was adapted from the Tuning protocol described in the introduction (McDonald et al., 2012), and was the first one offered in the course. The piece of work in question was the business model, which is essentially a description of an innovative business concept. Included in the business model are elements such as customer relationships, revenue streams, and key resources.

For both iterations, the online discussion protocol was divided into three basic parts:

- Part A: Students posted their written assignment (business model), and then asked group members to consider at least one aspect in which they desired constructive feedback.
- Part B: One week after Part A, students chose at least one person to whom to reply and provide feedback.
- Part C: One week after Part B, students reflected on the next steps for further developing their business model based on peer feedback.
- Once the discussion concluded, the students submitted the business model as a formal assignment for grading.

Based on feedback from the students and instructor, problem areas with the discussion were identified and enhancements were made to the online discussion protocol used, resulting in Iteration 2 of the discussion protocol (Appendix A).

Seven improvements were made after the first design iteration: 1) The discussion structure was simplified. In Iteration 1, Part A was in one discussion forum (i.e., a discussion board page), while Parts B and C were in a separate forum. As this resulted in confusion about how to interact with each other, in Iteration 2, Parts A and B were consolidated into one forum to encourage smoother peer feedback, with Part C (personal reflection) submitted as part of the final written Business Model assignment. 2) The feedback prompts were simplified. In Iteration 1, students were told in part B to specify their peer feedback to be "warm" (describe what is working with the model), "cool" (consider aspects of work that need improvement) and/or "hard" (ask deeper questions that get at the larger aspects of the work). In an effort to simplify the instructions, in Iteration 2, this classification was omitted and students were only asked to provide feedback based on the first person's post. Instead of including complicated directions, the warm, cool, and hard feedback was modeled in example discussion posts. 3) To simplify the focus of the discussion, the discussion and the Business Model assignment instructions were separated. 4) Due dates for the multiple parts of the discussion were added to the course calendar in Iteration 2 to provide reminders to students. 5) Exemplary discussion samples were provided from past semesters in order to display good examples. 6) A rubric was added to explain how the discussion would be graded. 7) The point-value of the discussion went from 10 points to 30 points in order to encourage more careful work.

Data Collection and Preparation

With the approval from the Institutional Research Board, the research team collected two types of data, survey and discussion posts, over the two iterations (Table 1) in the spring and fall semesters of 2014. Based on the CoI framework, researchers created a survey instrument (Arbaugh et al., 2008) to measure the quality of online teaching and a coding template (Garrison et al., 2000) to analyze indicators of cognitive presence, social presence, and teaching presence in students' discussion activities. The researchers used these two instruments in our research to evaluate students' perceptions and discussions. At the end of each semester, a survey was distributed that asked students to assess the protocol discussions used in the course. This survey included two parts: a quantitative section that utilized the CoI instrument (Appendix B), which assessed students' perceived levels of social, cognitive and teaching presences with reference to the protocol-based discussions, and a qualitative section that included open-ended questions to assess students' feedback on the discussions. In addition to this survey, students' discussions were analyzed for the social, cognitive, and teaching presences. The triangulation of this data enabled us to examine CoI from two main perspectives: the closed-ended questions were directed mainly at the instructor role in fostering a CoI, the coding of the group discussions was focused on the students' contribution to CoI in the discussions, and the open-ended questions reflected both the instructor and student involvement.

Table 1
Data Source Types, Analyses, and Participant Number

Data Source	Analysis	Iteration 1	Iteration 2
Semester		Spring	Fall
Survey Closed-ended	Descriptive & Independent Samples T-Test	*N*=394	*N*=446
Survey Open-ended I	Content Analysis	*N*=394	*N*=446
Survey Open-ended II	Content Analysis	*N*=394	*N*=446
Discussion posts	Content Analysis & Fisher's Exact Test	10 (One selected group)	11 (One selected group)

Analyses

For Research Question 1, the research team analyzed the open-ended survey results to evaluate students' perceptions towards the enhancements made to the protocol-based discussions. For Research Question 2, the researchers analyzed both the quantitative survey results and the discussion posts to assess if the enhancements made to the protocol-based discussions impacted the CoI elements.

Open-ended survey analysis. Students in both iterations were given open-ended questions as part of the course evaluation. In total, 840 students responded with comments on the protocol-based discussions. One open-ended question asked them to provide feedback about the discussions. In Iteration 2, an additional question asked students how the discussion contributed to their performance on the written assignment. A conventional content analysis was applied to the first open-ended survey question. Conventional content analysis emphasizes becoming immersed in the data so that new insights can emerge (Hsieh & Shannon, 2005). For the initial analysis of the first open-ended survey question, one member of the research team who was unfamiliar with the course design of either iteration independently open-coded to reveal emerging themes, noting anything in relation to the main research questions. Examples of initial codes generated include group size, group composure, clarity of instructions, amount of feedback, type of feedback, and usefulness of feedback, among others. Codes were then organized into the following categories: instructions, group formation, improving learning, examples, and feedback. The second open-ended survey question was then analyzed, with the team member continuing with open coding, and filing them under the initial codes and categories when appropriate. Examples of emerging themes include clarity of instructions and purposeful group formation. Finally, a comparative qualitative analysis between the two iterations was performed on students' responses. The differences among the themes between the two iterations were then highlighted. In addition to allowing the themes to emerge from the data, the research team also discussed the themes in connection with the three presences of the CoI framework within the Discussion section (Merriam, 2001).

Quantitative survey analysis. For the quantitative survey results, frequencies were used to analyze information regarding the level of presences. Independent samples t-tests were conducted to investigate the differences in perceptions between two iterations.

Discussion post analysis. An in-depth analysis of the discussions of one group from Iteration 1 and 2 was conducted. The discussion posts were downloaded, cleared of student names and identifying information, and stored on a secure server as randomly numbered groups. To prepare for the coding process, one member of the research team with expertise in the modified CoI coding system (see Table 2) devised in an earlier study (Zydney et al., 2012) trained two other members. The main difference between the original CoI coding system developed by Garrison et al. (2000) and the modified one is that it separated the "individual and group categories within cognitive presence to distinguish between participants answering questions as a monologue unrelated to other participants' postings versus participants interactively discussing the topic with one another" (Zydney et al., 2012, p. 81). For this process, each complete thought was the unit of analysis and was coded for cognitive, social, and teaching presence as defined by this model. Each unit of analysis was coded for the highest level of cognitive presence and could have multiple social or teaching indicators. To establish the coding process among the research team, a practice group of posts was selected and coded by two members of the research team. The third observed to understand the process and provide input on any cases about which the original coders did not agree. A random group of posts from each iteration was then coded independently by two researchers who later met to resolve discrepancies. A total of 366 ratings was assigned with 47 discrepancies for an inter-rater reliability of 87%. All discrepancies were resolved for a 100% agreement.

After the discussion posts were coded, frequencies and percentages for each category by individual and group presence were calculated to analyze the differences in the group between the two iterations. In addition, the Fisher's Exact test (2-sided) was used to identify statistically significant results. This nonparametric test was chosen because of the small sample size. Following this analysis, further qualitative investigations were done to identify differences between protocol iterations. Examples of student posts were used to illustrate any variances.

Table 2.
CoI Coding Template (adapted from Garrison et al. (2000))

Elements	Categories	Indicators	Example
Cognitive Presence	Triggering Event	A new or related topic is raised	NA[a]
	Exploration - Individual	Ideas, experiences, and prior knowledge are repeated and/or described	"Obstacles I am anticipating include funding the business and establishing brand awareness and loyalty."
	Exploration - Group	Ideas are exchanged among the group, questions are asked to confirm understanding	"I like your idea to make a Bookstore/Cyber Cafe that would sell college books under retail value."
	Integration - Individual	Prior knowledge and/or experience is related to the text; texts are related to each other	"I've talked to many friends and family members which has caused me to want my product to be a value based product instead of cost driven."
	Integration - Group	Ideas of others are connected and/or expanded upon; counterpoints raised; improvements are suggested	"While I agree with his premise that word of mouth advertising is great, it's not the least bit practical for a business that's just starting out; you must advertise through other mediums first to get your name out there."

	Resolution - Individual	Applying ideas to personal future, theory to practice	"Many large groups have the pain of paying far too much for too many rooms to fit all of their members in... My hotel chain will solve these problems by... putting the large groups into larger rooms for far less cost to them."
	Resolution - Group	Applying what has been discussed	"The response I got from X completely turned my original idea around. At first, my plan was having the customers get picked up from home by a professional driver who takes them for a ride. This idea proved to be risky because it would be incredibly tricky to get an insurance company to cover injuries that can possibly occur on public roads. With having a closed course, I can stay within certain parameters to severely lessen liability on my business."
Social Presence	Emotional Expression	Including emoticons, humor, and statements which exhibit emotion	"Feel free to give me a ring when and if you need any graphic design or photography work done, D, ha-ha."
	Open Communication	Complimenting someone, expressing oneself freely, sharing stories, including names	"XX, I think the concept of bringing the gym to the customer is a great idea."
	Group Cohesion	Encouraging collaboration, community building, helping and supporting	"I hope this helped a little!"
Teaching Presence	Instructional design and organization	Identify and prompt discussion topics	NA[a]

Facilitating discourse	Asking others for feedback, encouraging the direction of the discussion	"If anyone has any feedback or ideas they would like to share on reducing costs, I would greatly appreciate it!"
Direct instruction	Explain content in an authoritative way, provide support for claims, providing resources	NA[a]

[a]NA - not applicable because it did not show up in the groups' discussions.

Results

A number of themes developed for the first open-ended survey responses including the clarity of instructions, purposeful group formation, improving learning and gaining new perspectives, encouraging community and collaboration, providing examples and eliciting effective feedback. An overall comparison of open-ended responses between iterations yielded a variety of similarities and differences among the six themes.

Similar themes between iterations. Overall, responses across iterations were found to be similar for clarity of instructions, purposeful group formation, improving learning and gaining new perspectives and encouraging community and collaboration.

Clarity of instructions. The responses concerning the instructions and access to the discussions were overwhelmingly positive with almost all students reporting clear, easy-to-use instructions as one student illustrated, "The discussions were very accessible and easy to reply to other posts. I thought the instructions were spot on and I had no difficulties." A very small amount of the responses suggested a variety of changes. A few students commented that the instructions were too long, with one student suggesting that, "The instructions were too complex that it caused confusion… If the instructions were simplified then they would be easier to follow." A few students suggested that it was initially unclear where or when to post responses, or the proper place was difficult to find among all the other groups. A small number of students also requested reminders of dates that responses were due, or suggested that some due dates were unclear. One student suggested that, "I believe the instructions could have been described in a way for it to be easier to understand along with the due dates."

Purposeful group formation. Many suggestions focused on group formation, however, the suggestions varied greatly. Some students wanted a larger group for more feedback, while others asked for smaller groups for more personal communication. For example, one student suggested that future discussions should include "larger Bazinga circles [discussion groups] to ensure more replies and comments being shared." In contrast, another student suggested that, "maybe by making groups smaller you could improve this experience." Some students chose to focus on group composition by suggesting that groups include students of similar interests, while other suggestions asked for groups to include students of varied interest. To illustrate, one student wrote, “I would have liked Bazinga groups to be picked because of similarities in the field even if it fell down to just separating it into products and services. Or what the industry is.”

Improving learning and gaining new perspectives. Students also reported that their performances on the class assignments were improved because the structured discussion allowed them to get multiple perspectives, iron out details, and identify problems. For example, one student commented that, “The structured discussions were very easy to understand and helped me learn more about my ideas from different perspectives.” In addition, giving feedback to other students was reportedly helpful by providing peer examples to compare to their own work. Overall, students felt that the protocol discussion helped them receive a better grade on the final project. One student summarized this idea by writing, “I really liked the discussions. This helped wrap everything up that the research team had learned. The end discussions, which I felt I was very prepared for, was thanks to the Bazinga discussions.”

Encouraging community and collaboration. Finally, students reported that the group discussion made them feel better about their work and interacting with the class. One student reported that “the discussions gave me a feeling of belonging and collaborating with other students,” while another commented that, “I was able to interact with my peers more than in any other online course I’ve taken.” With class sizes in the 350-450 range, it is easy for students to feel disconnected, especially in an online setting. However, the protocol discussions may have helped students feel connected, with one student stating “I found the discussions helpful! They definitely helped me develop a sense of community within the class.” The protocol discussions were also reported to increase collaboration. As one student wrote, “I thought the discussions were great and made me feel like I was actually a part of the class. Being in such large classes is tough and having a small group collaborations helped immensely.”

Different themes between iterations. Responses concerning provided examples and effective feedback were considerably different between groups.

Providing examples. Although the overall response to clarity of instructions was similar between iterations, one difference emerged. Students in both iterations consistently responded that instructions were clear and easy to follow. Additionally, the students in Iteration 2 often elaborated that the provided examples were helpful. One student suggested

> The instructions were very clear and concise. I love that you guys give us clear instructions AND an example of what you are somewhat looking for. Makes working on the assignment a lot easier and is very reassuring to know I am on the right track.

Eliciting effective feedback. Responses concerning effective feedback were considerably different between groups. The students in Iteration 2 more often offered positive responses regarding effective feedback, while the students in Iteration 1 compiled more negative responses in this area. Feedback from students in Iteration 1 more often identified the negative aspects of feedback including lack of feedback, ineffective last-minute feedback, or feedback that was minimal. For example, a student from Iteration 1 responded that

> I felt as though I didn’t receive adequate feedback through the circle discussions, and I never met any of the members in my group. I posted several assignments and asked for

> feedback but didn't receive any. It was clear from the feedback that I did receive that some students did actually take the assignment seriously and attempt to provide me with a different perspective to aid me, but others were short and frankly quite useless.

Students in Iteration 2 more often responded that the feedback was helpful and provided insight. As one student from Iteration 2 responded, "By using feedback from my Bazinga group members I was better able to come up with and expand on my business concept after taking into account the various ideas and contributions that my group members gave me. Very helpful!"

Additional themes related to additional question for iteration 2. In Iteration 2, an additional question was added to the course survey to provide further insight into how the protocol-based discussions influenced the performance on the Business Model assignment. A number of additional themes developed including guiding student progress, utilizing peer examples, and revising and refining ideas.

Guiding student progress. Students felt that the protocol discussion helped them to keep pace with the course schedule and assignments. Students reported being forced to start assignments early, to understand material in order to be able to participate in discussions, and to keep up to date with responses. As a result, students had more time to reflect and modify their work before turning in the final copy. One student wrote, "The structured (protocol) discussions, like the Business Model assignment, helped me with time management. Since assignments were due first in the discussion it allowed me to schedule my time more efficiently."

***Utilizing peer examples*.** In addition, students felt that the discussions allowed them to see examples of other student's work, and provided a basis to compare their own work. As one student suggested, "I loved the structured (protocol) discussions! It was great to see many examples of what my peers were doing, to help improve my own understanding of the assignement and better my performanace."

***Revising and refining ideas*.** The students reported that the protocol discussions helped to improve performance by allowing them to revise and refine their ideas. The discussions allowed students to share their ideas, receive feedback from their peers, and adjust the assignments before submission. For example, a student responded

> They helped me in a number of ways, but the most is when I was missing something or didn't go into detail about something, the people in my Bazinga group caught it and helped me score very high on these assignments. They would also tell me things like 'I love your idea but have you thought about adding this?' and with things like that it really helped me to refine my idea and made it stronger than I ever thought possible!

Research Question 2: How Did the Enhancements Made to the Protocol-Based Discussions Impact the Elements that Contribute to a CoI?

A summary of the quantitative survey results and the quantitative and qualitative analysis of the discussion posts indicate that the enhancements made to the protocol-based discussions positively influenced elements of the CoI. In particular, the cognitive presence among students was significantly higher in Iteration 2 than in Iteration 1.

Summary of quantitative survey results. Exploratory analyses (Table 3) showed that for both iterations, all three presences of the CoI were perceived at high levels. However, students' perceptions of teaching, social, and cognitive presences were significantly higher in Iteration 2.

Table 3.
Means and Standard Deviations of the Three Presences in the Two Iterations

	Iteration 1 (Spring 2014)			Iteration 2 (Fall 2014)		
	N	Mean	Std. Deviation	*N*	Mean	Std. Deviation
Teaching Presence	394	4.04	0.60	446	4.27	0.51
Social Presence	394	3.79	0.72	446	4.10	0.63
Cognitive Presence	394	3.74	0.76	446	4.07	0.64

There were statistically significant differences in all three presences between the two iterations (Table 4). These results suggest that the improved discussion protocol in Iteration 2 did have an effect on students' perception of online community. Specifically, when the discussion protocol is enhanced, students perceived a higher sense of teaching presence, social presence and cognitive presence.

Table 4.
Mean Differences and T-Value of the Three Presences Between the Two Iterations

	Mean Difference	*t*	*df*	Sig. (2-tailed)
Teaching Presence	.23	6.05	770	<.001***
Social Presence	.31	6.60	788	<.001***
Cognitive Presence	.33	6.79	773	<.001***

Note: *** $p<.001$.

Discussion post analysis. Overall, the students' discussions from Iteration 2 presented a higher frequency of social presence than the discussions from Iteration 1, as shown in Table 5. The frequencies and percentages of cognitive, social, and teaching presence in the discussions were similar between iterations. Further analysis was done to assess differences between iterations within each category.

Table 5.
Overall Differences between Cognitive, Social, and Teaching Presence

Element	*n*	Iteration 1		*n*	Iteration 2	
		Frequency	%		Frequency	%
Cognitive	10	48	62	9	49	55
Social	10	21	27	9	31	35
Teaching	10	8	11	9	9	10
Total	10	77	100	9	89	100

Cognitive presence. An examination of the instances of cognitive presence between iterations revealed a significantly higher rate of individual exploration in Iteration 1 in the discussions (p = .03) as shown in Table 6. In addition, this led to a significantly higher presence of total individual cognition for the students in Iteration 1 (p = .03). Subsequently, the students in Iteration 2 produced a significantly higher rate of total group cognition (p = .03) with higher frequencies achieved in all three categories. Overall, the students in Iteration 2 displayed a more even distribution between individual and group cognition.

A qualitative examination of the discussion posts revealed further differences among the groups. Posts to part A of the discussion, in which students presented their business models and asked for feedback were similar across iterations. However, responses to parts B (providing peer feedback) and C (reflection and improvement of the business model) were shorter and coded at a lower cognitive presence in Iteration 1. For example, a typical response by the students in Iteration 1 to part B of the assignment was “It all looks pretty solid to me #4 good work.”

While one of the shorter responses by a student in Iteration 2 was:

> Hey #7, love the idea!
> Many people in your target market are going to be interested in the business you have to offer and are going to love the service you provide them. Many large parties don’t like to be separated when having a good time in a hotel and hate to pay more just to get extra rooms. I like the idea of the unlimited access to entertainment in the lobby as well and the personal caterer.
> Your strategic alliance will be key to your business to get up and get going. Your business is going to need key investors that are willing to work with you and believe in the work you’re going to get done. Liability may be an issue as well because when you have a lot of people in one room it can tend to get a bit reckless and things may break in the room that were not your businesses fault. So make sure you have liability covered when people check in for your services.

In many cases, the responses of students from Iteration 1 to parts B and C of the assignment were 1 to 2 sentences that were coded with low cognitive presence or with none at all. However, the students in Iteration 2 more often responded in multiple paragraphs that were coded at various group levels.

Table 6
Differences in Categories within Cognitive Presence

Element	Iteration 1		Iteration 2		*p*
	Frequency	%	Frequency	%	
Triggering Event	0	0	0	0	0
Individual					
Exploration	25	52	14	29	.03*
Integration	7	15	5	10	.55
Resolution	3	6	5	10	.71
Sub-total	35	73	24	49	.03*
Group					
Exploration	6	10	9	18	.58
Integration	7	17	14	29	.15
Resolution	0	0	2	4	.50
Sub-total	13	27	25	51	.03*
Total	48	100	49	100	-

Note: * $p<.05$.

Social presence. There were no significant differences between iterations in terms of social presence (Table 7), although the students in Iteration 2 had a higher frequency of social presence. In each iteration, the social presence was largely focused in three areas: compliments, directing a comment to someone specific, and emotional expression. A qualitative difference among the iterations was that the students in Iteration 1 often expressed concern over doing the assignment correctly, while many of the students in Iteration 2 complimented fellow students' ideas as seen in the quote below.

> Iteration 1:
> #3, everything looks great only thing I would say is check with the professor to make sure its [sic] ok to number them. I was a bit lost of this because I remember she said no bullets. Are numbers considered bullets? I don't even want to chance it myself.

Iteration 2:
I value any input from my group members. Let me know what I need to work on.

Table 7
Differences in Categories within Social Presence

Element	Iteration 1		Iteration 2		*p*
	Frequency	%	Frequency	%	
Emotional Expression	3	14	4	13	.72
Open Communication	17	81	25	81	.18
Group Cohesion	1	5	2	6	1
Total	21	100	31	10	-

Teaching presence. There were no significant differences between iterations in terms of teaching presence as seen in Table 8. In each iteration, the teaching presence was focused on facilitating discourse. This was primarily accomplished by prompting others to generate new ideas as a reflection on the business model. Posts were similar across iterations.

Table 8
Differences in Categories within Teaching Presence

Element	Iteration 1		Iteration 2		*p*
	Frequency	%	Frequency	%	
Instructional Design & Organization	0	0	0	0	-
Facilitating Discourse	8	100	9	100	1
Direct Instruction	0	0	0	0	-
Total	8	100	9	100	-

Discussion

The current study examined the use of discussion protocols in large classes to engage students and facilitate discussions. In the following section, the theoretical and practical implications of the research findings are discussed.

Community of Inquiry

One unexplored area of research is the use of protocol-based discussions in large enrollment courses to improve students' perception of the CoI. After making several enhancements (e.g., simplifying the instructions, adding example posts, providing a rubric) in the second design iteration of protocol-based discussions, the research team noted changes associated with the cognitive presence, social presence, and teaching presence.

Cognitive presence. After the enhancements were made to the discussion-based protocols in Iteration 2, students perceived a significantly higher level of cognitive presence. This may be the result of the new perspectives gained from increased feedback received from peers. In closely examining one group's discussion posts from each iteration, the researchers also noted a statistically significantly higher level of group cognition in discussion posts in the second iteration. This increased group interaction may have resulted from the protocol-based discussions becoming more structured with clearer directions and expectations set for what students were supposed to do. This result corroborates an earlier study that found that group cognition increased with a protocol-based discussion over non-protocol-based discussion (Zydney et al., 2012). It may be that increasing the structure of protocol-based discussions further improves upon group cognition; however, additional research would be needed to confirm this idea.

Social presence. Similar to what was seen with cognitive presence, the enhancements to the discussion protocols in Iteration 2 significantly improved students' perception of social presence in the discussions. Although students' ratings of social presence improved in the second iteration, other data sources showed similarities between the two iterations. For example, the analysis of the open-ended survey results revealed a similar theme of community and collaboration across both iterations, with students noting feelings of community and belonging. Moreover, the analysis of the small group discussions revealed no difference pertaining to social presence between the two iterations. This lack of statistical significance may be the result of the fact that both design iterations included smaller sub-group discussions that enabled students to interact with one another more. (Kim, 2013).

Teaching presence. Students' perception of teaching presence also significantly improved after enhancements were made to the protocol-based discussions in Iteration 2. This difference may be the result of the increased peer feedback noted by students who received the second iteration of the discussion protocols. Although students' ratings of teaching presence improved in Iteration 2, examination of the group discussions revealed no statistical difference between levels of teaching presence in students' posts, regardless of which iteration they received. This was illustrated by similar numbers of posts from students helping to facilitate discussion by encouraging feedback from one another. This lack of statistical significance in levels of teaching presence in the discussion posts may be the result of the fact that, regardless of the enhancements made, the protocol itself is designed to enable students to facilitate themselves in the discussion. This result was also noted in an earlier study that found that the "protocol helped the students facilitate themselves and empowered them to design the discussion" (Zydney et al., 2012, p. 85).

Practical Implications

The findings of this study generate important implications for facilitating meaningful online discussions in large classes. Most importantly, the findings of this study suggest that implementing protocol-based discussions is a plausible teaching strategy in classes with large enrollments. Students could achieve effective communication and interactions within small groups of approximately 10 members in large classes, and the optimal group size is still an interesting area for future research. Students in this study enjoyed the small communities, felt a sense of connection to classmates, and benefited from peer learning that would not be achievable otherwise in large-size classes. Protocols help keep the large community on the same page; everyone knows what is expected, including the instructor.

It is essential to closely attend to the "protocol" nature of the discussion prompt. For larger classes, especially, the prompt needs to be explicit, with due dates being clear, and examples given. As the research team found, the second design iteration resulted in higher perceptions of community and more even distribution of the presences. Being explicit is extremely important in very large classes to prevent mass confusion. It is important to explain the purpose of the protocol, what to contribute, how to provide feedback to peers, and the timeline. In addition, the very nature of the discussion set-up is important. Including the original post and peer feedback in one discussion forum enables smooth communication between peers; this communication was less frequent in the first iteration when parts A and B were separated.

It appears that, in large undergraduate classes, students prefer scaffolding with authentic examples and simplified, but clear instruction. Although both cohorts felt that they received clear instruction for the discussion activity, Iteration 2 was particularly successful due to added features such as examples, grading rubrics, and due date reminders. It was surprising that students in Iteration 2 were more satisfied with the quality of the feedback, given that more detailed direction on how to give peer feedback was provided to students in Iteration 1. The design enhancements in the second iteration resulted in fewer concerns over the "correctness" of the assignment and reassured the students. There was also less concern about when the discussion parts were due. This finding suggests a "less is more" instructional design approach; instead of providing a large amount of clarified instruction, it may be advantageous to offer the most concise instruction possible along with an example.

It was also surprising that even after adding due date reminders in the course calendar, there was still feedback from students in Iteration 2 about needing more reminders of the multiple due dates for the protocol-based discussion. This points to the need for more student support in the use of the learning management system (LMS). Students need to not only be able to submit assignments and check grades, but also to skillfully leverage the features of the LMS (such as the course calendar) to support their own learning. Therefore, student training emerges as an issue for further pursuit.

Limitations and Future Research

There are several limitations to this study. One major limitation is the convenience sample selection. Even though this study had a large sample size over two iterations, the data only included undergraduate students in one business class at one southeastern university in the United States. In future studies, the research team will include classes of various disciplines and school environments to assess whether the current findings would generalize to varied contexts or samples. Moreover, the researchers only randomly selected one group from each iteration for discussion analyses, which limits the sample size of these quantitative analyses. More studentgroups should be analyzed to see if the results are consistent in future studies.

In addition, future research will use controlled experimental studies to measure the effect of protocol-based discussion on students' overall learning. Our study results showed a statistically significant group cognition but the effect was confounded by other factors, such as students' prior knowledge and the increase in assigned grade to protocol-based discussions in Iteration 2. The research team plans to implement pre- and post-examinations for future studies to control for individual differences between classes. Also, the effectiveness of modeling in protocol-based discussion is one that warrants further research.

In summary, this study provides some evidence that the protocol-based discussion activity can engage students and enhance their cognitive presence in large classes. More research is needed in large online classes to further examine the effects of the discussion strategies on students' learning. The researchers hope that this research will be useful to educational practitioners and researchers as they continue to investigate and build a knowledge base of teaching strategies for large online classes, especially engagement strategies to build personal learning communities without having to place heavy logistical burdens on instructors.

References

Anderson, T., Liam, R., Garrison, D. R., & Archer, W. (2001). Assessing teaching presence in a computer conferencing context. *Journal of the Asynchronous Learning Network*, *5*(2). Retrieved from http://www.sloan-c.org/publications/jaln/v5n2/index.asp

Anderson, T., Rourke, L., Garrison, D. R., & Archer, W. (2001). Assessing teaching presence in a computer conferencing environment. *Journal of Asynchronous Learning Networks*, 5 (2).

Arbaugh, J. B., Cleveland-Innes, M., Diaz, S. R., Garrison, D. R., Ice, P., Richardson, J. C., & Swan, K. P. (2008). Developing a community of inquiry instrument: Testing a measure of the community of inquiry framework Using a multi-institutional sample. *Internet and Higher Education*, *11*(3-4), 133–136. doi: 10.1016/j.iheduc.2008.06.003

Brannon, R., & Essex, C. (2001). Synchronous and asynchronous communication tools in distance education. *TechTrends*, *45*(1), 36-42.

Darabi, A., Liang, X., Suryavanshi, R., & Yurekli, H. (2013). Effectiveness of online discussion strategies: A meta-analysis. *American Journal of Distance Education, 27*(4), 228-241. Doi: 10.1080/08923647.2013.837651

deNoyelles, A., Zydney, J.M. & Chen, B. (2014). Best practices for creating a community through online discussions. *Journal of Online Learning and Teaching, 10*(1). Retrieved from http://jolt.merlot.org/vol10no1/denoyelles_0314.pdf

Eyitayo, O. (2005). Experimenting e-Learning with a large class. *International Journal of Education and Development Using ICT, 1*(3). Retrieved from http://ijedict.dec.uwi.edu/viewarticle.php?id=109&layout=html

Gao, F., Zhang, T., & Franklin, T. (2013). Designing asynchronous online discussion environments: Recent progress and possible future directions. *British Journal of Educational Technology, 44*(3), 469-483. doi: 10.1111/j.1467-8535.2012.01330.x

Garrison, D. R., Anderson, T., & Archer, W. (2010). The first decade of the community of inquiry framework: A retrospective. *The Internet and Higher Education*, *13*(1-2), 5–9. doi: 10.1016/j.iheduc.2009.10.003

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. doi: 10.1016/j.iheduc.2007.04.001

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*(2-3), 87–105. Doi: 10.1016/S1096-7516(00)00016-6

Garrison, D. R., Anderson, T., & Archer, W. (2001). Critical thinking, cognitive presence, and computer conferencing in distance education. *American Journal of Distance Education*, 15(1).

Hew, K.F., & Cheung, W.S. (2011). Higher-level knowledge construction in asynchronous online discussions: An analysis of group size, duration of online discussion, and student facilitation techniques. *Instructional Science, 39*, 303-319. Retrieved from: http://www.jstor.org/stable/23882804

Hew, K.F., Cheung, W.S., & Ng, C.S. (2010). Student contribution in asynchronous online discussion: A review of the research and empirical exploration. *Instructional Science, 38*, 571-606. doi: [add here]

Hewitt, J., & Brett, C. (2007). The relationship between class size and online activity patterns in asynchronous computer conferencing environments. *Computers & Education*, *49*(4), 1258–1271. doi: 10.1016/j.compedu.2006.02.001

Hsieh, H., & Shannon, S.E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, *15*(9), 1277-1288.

Kim, J. (2013). Influence of group size on students' participation in online discussion forums. *Computers & Education*, *62*, 123–129. doi: 10.1016/j.compedu.2012.10.025

Makitalo, K., Weinberger, A., Hakkinen, P., Jarvela, S., & Fischer, F. (2005). Epistemic cooperation scripts in online learning environments: fostering learning by reducing uncertainty in discourse? *Computers in Human Behavior, 21*, 603-622.

Maringe, F., & Sing, N. (2014). Teaching large classes in an increasingly internationalising higher education environment: Pedagogical, quality and equity issues. *Higher Education*, 67(6), 761–782. doi:1007/s10734-013-9710-0

McDonald, J.P., Zydney, J.M., Dichter, A., & McDonald, B. (2012). *Going online with protocols: New tools for teaching and learning*. NY: Teachers College Press.

Merriam, S. B. (2001). *Qualitative research and case study applications in education*. (2nd ed.). San Francisco: Jossey-Bass Publishers.

Orellana, A. (2006). Class size and interaction in online courses. *The Quarterly Review of Distance Education*, *7*(3), 229–248. Retrieved from http://wps.prenhall.com/wps/media/objects/4512/4621309/Survey_Online_Class_Size.pdf

Rocca, K.A. (2010). Student participation in the college classroom: An extended multidisciplinary literature review. *Communication Education, 59*(2), 185-213. doi: 10.1080/03634520903505936

Rourke, L., Anderson, T. Garrison, D. R., & Archer, W. (1999). Assessing social presence in asynchronous, text-based computer conferencing. *Journal of Distance Education*, 14(3), 51-70.

Tomei, L. (2004). The impact of online teaching on faculty load: Computing the ideal class size for online courses. *International Journal of Instructional Technology & Distance Learning*, *1*(1). Retrieved from http://itdl.org/journal/jan_04/article04.htm

Yang, Y.-T. C. (2007). A catalyst for teaching critical thinking in a large university class in Taiwan: asynchronous online discussions with the facilitation of teaching assistants. *Educational Technology Research and Development*, *56*(3), 241–264. doi: 10.1007/s11423-007-9054-5

Zydney, J. M., deNoyelles, A., & Seo, K. K.-J. (2012). Creating a community of inquiry in online environments: An exploratory study on the effect of a protocol on interactions within asynchronous discussions. *Computers & Education*, *58*(1), 77–87. doi:10.1016/j.compedu.2011.07.009

Acknowledgements. The research team is grateful to Ms. Kathie Holland, the course instructor, for giving us permission to conduct our research and continuously working with us on improving the online discussion activities.

Appendix A: Iteration 2 Protocol

For this group discussion posting, I would like you to use a specific structure called a "Tuning" protocol. The Tuning process allows for reflection and meaningful feedback about your Business Model, and will likely result in your creating a better business model, and earning a better grade, too. This structure assumes that you want to improve your business model, and that your Bazinga Circle members will deliver thoughtful and substantive feedback. Think of it like tuning up a car or an orchestra. Your communal goal is to help each other make A's on this assignment, so do this early in the week so you can use their feedback to improve your success on your assignment.

There are three parts to this discussion structure: Parts A, B, and C.

Part A: Due Week 8

Begin your post with the words "**PART A"**. Begin with a one- or two-sentence description of your business to remind everyone what your concept is. Share the elements of your emerging Business Model. Then ask your Circle members to review your submission and give you feedback about something specific in your Business Model you believe could be improved. (For example, perhaps you want their feedback on your Revenue Streams or on Validation and Pivoting.)

For Example: *(example provided here by the instructor)*

Part B: Due Week 10

Choose one group member's Business Model and provide feedback to him or her. Each member should give and get feedback from at least one member, so reply to someone who hasn't received any feedback yet. Your communal goal is to help each other make A's on this assignment.

Begin your post with the words "**Part B**". Then provide feedback that will help your group member to improve their submission.

For Example: *(example provided here by the instructor)*

Part C (*to be submitted with the Business Model assignment*): Reflect on what happened as a result of this structured discussion. What ideas did your group members provide that helped you to improve your Business Model? Who had the best one; why was it better than the others?

Appendix B: Survey Instrument

Teaching Presence

Design & Organization

1. The instructor clearly communicated important course topics.
2. The instructor clearly communicated important course goals.
3. The instructor provided clear instructions on how to participate in course learning activities.
4. The instructor clearly communicated important due dates/time frames for learning activities.

Facilitation

5. The instructor was helpful in identifying areas of agreement and disagreement on course topics that helped me to learn.
6. The instructor was helpful in guiding the class towards understanding course topics in a way that helped me clarify my thinking.
7. The instructor helped to keep course participants engaged and participating in productive dialogue.
8. The instructor helped keep the course participants on task in a way that helped me to learn.
9. The instructor encouraged course participants to explore new concepts in this course.
10. Instructor actions reinforced the development of a sense of community among course participants.

Direct Instruction

11. The instructor helped to focus discussion on relevant issues in a way that helped me to learn.
12. The instructor provided feedback that helped me understand my strengths and weaknesses.
13. The instructor provided feedback in a timely fashion.

Social Presence

Affective expression

14. Getting to know other course participants gave me a sense of belonging in the course.
15. I was able to form distinct impressions of some course participants.
16. Online or web-based communication is an excellent medium for social interaction.

Open communication

17. I felt comfortable conversing through the online medium.
18. I felt comfortable participating in the course discussions.
19. I felt comfortable interacting with other course participants.

Group cohesion

20. I felt comfortable disagreeing with other course participants while still maintaining a sense of trust.
21. I felt that my point of view was acknowledged by other course participants.
22. Online discussions help me to develop a sense of collaboration.

Cognitive Presence

Triggering event

23. Problems posed increased my interest in course issues.
24. Course activities piqued my curiosity.
25. I felt motivated to explore content related questions.

Exploration

26. I utilized a variety of information sources to explore problems posed in this course.
27. Brainstorming and finding relevant information helped me resolve content related questions.
28. Online discussions were valuable in helping me appreciate different perspectives.

Integration

29. Combining new information helped me answer questions raised in course activities.
30. Learning activities helped me construct explanations/solutions.
31. Reflection on course content and discussions helped me understand fundamental concepts in this class.

Resolution

32. I can describe ways to test and apply the knowledge created in this course.
33. I have developed solutions to course problems that can be applied in practice.
34. I can apply the knowledge created in this course to my work or other non-class related activities.

<u>5 point Likert-type scale</u>

1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree

Exploring Small Group Analysis of Instructional Design Cases in Online Learning Environments

Jesus Trespalacios
Boise State University

Abstract

The case-based approach is a constructivist instructional strategy that helps students apply their emerging knowledge by studying design problems in authentic real-world situations. One important instructional strategy in case-based instruction is to analyze cases in small groups before discussing them with the whole class. This study investigates the use of small-group structure to analyze case studies in online learning environments, as well as students' perceptions of the use of VoiceThread presentations to improve their learning of instructional design. The results show that a small group strategy has great potential to help students analyze case studies and consequently enhance learning. The implications of these findings for instructional designers and online instructors are discussed.

Keywords: Instructional Design; Case Studies; Online Learning; VoiceThread

Trespalacios, J. (2017). Exploring small group analysis of instructional design cases in online learning environments, *Online Learning 21*(1) 189-200. doi: 10.24059/ olj.v21i1.928

Introduction

Case-based approaches are constructivist instructional strategies that help students apply their emerging knowledge by studying design problems in authentic real-world situations (Jonassen, 1999, 2011; Jonassen & Hernandez-Serrano, 2002; Stepich & Ertmer, 2009). Research has shown that this approach is an effective strategy used to teach medicine, business, law, psychology, and teacher preparation (Lee et al., 2009; Pena-Shaff & Altman, 2009; Saleewong, Suwannatthachote & Kuhakran, 2012). For example, Honan and Rule (2002), Barnes, Christensen, and Hansen (1994) and Argyris (1980) agreed that real problems, analysis, and active student involvement are the central elements of case method teaching and learning. In the teaching of instructional design (ID), Carr-Chellman (1999) described this strategy as relevant because this field focuses essentially on solving ill-structured problems that possess incomplete information and multiple solutions. Additionally, Julian, Kinzie, and Larsen (2000) stated,

> In case analysis, instructional design students draw connections between their emerging knowledge of ID and the complex demands of actual practice. Cases can supplement student design projects, allowing further opportunity to reflect on relevant theory and methods as students explore a greater number of design issues in a broader array of environments (p. 165).

Wasserman (1994) described four necessary components in case-based instruction: (1) a case report, (2) study questions, (3) small group work, and (4) a whole group discussion. Considering cases as problems to solve, Jonassen (2011) also recommended the following four steps to support students' problem-based learning (PBL): (1) small group discussions to reason through the problem, (2) individual analysis of the case to understand the problem and find possible solutions, (3) students share what they have learned with the group and revisit the problem, and (4) "[a]t the end of the learning period (usually one week), students summarize and integrate their learning" (p. 154).

Instructional Design Expertise

One of the goals in instructional design courses is to provide students with opportunities to develop problem-solving skills to deal with instructional design situations where they need to identify issues and suggest instructional solutions (Ertmer & Stepich, 2005). However, developing this expertise in novice instructional designers is not an easy task because of the ill-structured nature of instructional design problems (Jonassen, 2011). Investigating the impact of guidance on the development of expertise by novice instructional designers, Ertmer et al. (2009) found that novices were able to perform more like instructional design experts after using the following analysis guidelines: (a) use your own words, (b) focus on the *big picture* rather than surface details, (c) make assumptions about missing information, (d) focus on root causes rather than quick fixes, (e) consider the core issues (those that are most central to your understanding of the situation), (f) consider the critical issues (those that are likely to have the greatest impact on a successful resolution), (g) if you identify multiple issues, think about how those issues fit together, and (h) think about where the issues you identify fit within the instructional design model. In fact, Ertmer et al. (2009) suggested, "the guidance encouraged novices to synthesize rather than summarize information, focus on principles rather than on surface features, identify relationships among identified issues, and make assumptions (i.e., to be reflective) based on what was stated in the case" (p. 121).

Small Group Activities in Online Environments

Research has shown that the process of peers working together in small groups appears to produce positive academic outcomes (Blumenfeld, Marx, Soloway & Krajcik, 1996; Wentzel & Watkins, 2011). More specifically, in a seminal work on the use of case studies to enhance instructional design education, Ertmer and Russell (1995) discussed the relevance of small group work:

> Following the case presentation, students work individually or in groups to analyze the data, evaluate the nature of the problem(s), decide upon applicable principles, and recommend a solution or course of action. Small group work, in or out of class, gives students the opportunity to discuss cases and questions with each other prior to the whole

> class discussion. These sessions give students their first chance to examine the issues presented in the case study; ideas are tried out in the safest of contexts. Study groups engage students in thoughtful consideration of the case issues and primes them for the more demanding whole-class discussion that follows (p. 24).

Thus, the creation of small groups to discuss cases is a relevant instructional activity that allows students to interact and identify key points before participating in the class discussion (Flynn & Klein, 2001).

After analyzing group-solving styles in two asynchronous online courses, Lowes (2014) recommended the following strategies to design collaborative group projects: (1) require a unique contribution from each group member; (2) provide clear instructions about collaborative activities; and (3) make available spaces for collaboration among the group members. One alternative to provide a space for students is to integrate computer-mediated communication (CMC) tools that allow students to communicate asynchronously and provide this important interaction among students (Benbunan-Fich & Hiltz, 1999; Rourke & Anderson, 2002). Recent studies have shown the relevance of video communication using tools such as VoiceThread to support communication and social presence among students in distance environments (Borup, West & Graham, 2013; Ching & Hsu, 2013). Thus, the purpose of this study was to explore the effectiveness of small group analysis of instructional design cases and students' perceptions of this activity in online learning. The research questions that guided this exploratory study were as follows: In an online learning environment,

> RQ1: How effective is the small group analysis of cases in instructional design when compared with experts' analysis?
> RQ2: What are students' perceptions of case-based VoiceThread presentations with regard to improving their learning of instructional design?

Methods

Twenty-one students enrolled in an online course in ID participated in this study. This three-credit course is required for the master's degree program in Educational Technology. Based on participants' introductions at the beginning of the course, students had a broad range of backgrounds, knowledge, and experiences. At the time of taking this course, five of the participants were technology coordinators or coaches for different schools, fifteen were teachers (elementary and secondary), and one worked for a consulting firm as an instructional designer. Thirteen participants were female (62%) and eight were male (38%). Eighteen participants lived in the United Stated and three lived overseas.

Course Setting

One week before the start of the course the participants received detailed information about the course objectives and activities and became familiar with the learning management system (LMS) Moodle in which the course was implemented. The content of the course was divided into weeks, starting on Mondays and ending at midnight on Sundays. This 15-week course contained different activities such as creating an ID job description, leading and participating in discussion forums, producing an instructional design project, and creating

VoiceThead presentations. During the first week, students were asked to briefly introduce themselves using VoiceThread.

The main readings of the course are from two textbooks. *Streamlined ID* (Larson & Lockee, 2014) is the ID textbook that the class follows to discuss introductory concepts in the field. The second textbook is the *ID casebook* written by Ertmer, Quinn and Glazewski (2014a), which is divided into three different sections. The first section contains 7 cases situated in K-12 environments, the second section contains 11 post-secondary cases, and the final section includes 12 cases situated in a corporate or manufacturing environment. Cases are approximately 4-7 pages long including text and pictorial material, and each one of them contains questions for preliminary analysis and implications for ID practice that could help instructors to organize their case-based instruction. As defined by the authors,

> The cases in this book are designed to be dilemma oriented: each case ends before the solution is clear. Students are expected to evaluate available evidence, to make reasonable assumptions as necessary, to judge alternative interpretations and actions, and, in doing so, to experience the uncertainty that commonly accompanies design decisions. (Ertmer et al., 2014a, p. xiv)

Small Group Activity

Following Wasserman (1994) and Jonassen's (2011) recommendations, students were randomly assigned to one of the five groups during the second week of the course. Each group was required to analyze three ID case studies (one for each level: K-12, higher education, and business) and lead a whole-class discussion. Members of the small groups were required to create a VoiceThread presentation where they analyzed the cases discussing the main issues and possible solutions. Requiring analysis of one main issue per slide, VoiceThread presentations contained 9 to 12 slides in total with audio comments ranging from 2 to 4 minutes per slide. All group members were required to participate in the presentation. Group members were encouraged to work on a Google presentation to create the slides, ensuring that the same format (background, font, and layout) would be used on each slide. Finally, five weeklong discussions were designed to discuss the cases with the whole class. VoiceThread presentations were shared at the beginning of each week to support the whole-class discussions that were led by the members of the small group.

Cases from the ID casebook (Ertmer, et al., 2014a) were chosen based on the relationship they had with the content of the ID textbook (Larson & Lockee, 2014). For instance, initial cases were related to task analysis or needs assessment because the analysis component of the instructional design process was discussed in the first chapters of the ID textbook (see Table 1). This decision was made based on the recommendation that "instruction created to help novices think like experts must be matched to the learners' existing knowledge and understanding, to make the tools of expertise accessible to them." (Hardré, Ge, & Thomas, 2006, p. 65). Additionally, the checklist developed by Ertmer et al. (2009) was provided to support the small group discussion and the individual analysis of the cases as experts.

Table 1
Discussions Content and Organization

Groups	Weeks in the semester	Cases assigned from the ID casebook	ID Content assigned from textbook
Group 1 (n=4)	Week 4	Cases 7, 18 and 28.	Analyzing needs and learners (chapters 1-3)
Group 2 (n=4)	Week 5	Cases 6, 10 and 19.	Analyzing context and content (chapters 4-5)
Group 3 (n=4)	Week 9	Cases 3, 8, and 29.	Aligning instruction and assessing learning (chapters 6-7)
Group 4 (n=4)	Week 10	Cases 2, 9 and 20.	Selecting strategies and technologies (chapters 8-9)
Group 5 (n=5)	Week 13	Cases 1, 13 and 21.	Producing and implementing instruction (chapters 10-11)

Data Collection and Analysis

To answer the first research question, (RQ1: How effective is the small group analysis of cases in instructional design when compared with experts' analysis?) VoiceThread presentations from the five small groups were analyzed using content analysis. As stated by Hsieh and Shannon (2005), "qualitative content analysis is defined as a research method for the subjective interpretation of the content of text data through the systematic classification process of coding and identifying themes or patterns" (p. 1278). To ensure consistency and comparability among the messages provided on VoiceThread, the information presented on each slide was treated as the unit of analysis. Using this strategy, 49 entries of information were identified for analysis. To establish validity in the content analysis process (Potter & Levine-Donnerstein, 1999; Rourke & Anderson, 2004), the instructor's resource manual for the ID casebook (Ertmer, Quinn & Glazewski, 2014b) was used to design the coding scheme. Thus, the data analysis process began with an identification of the main problems/issues that students found in their analysis of the instructional design cases across the written text and the audio presentation in each slide of the VoiceThread presentations. Then, these issues were compared with the issues identified by experts presented in the instructors' ID Casebook manual for each case. Besides analyzing the presence of experts' issues in the small group analysis of ID cases, a rubric with three levels was created to grade the level of similarity between the issues identified for the small groups and the experts. Three points were given if the issue described by the group matches completely to one of the issues discussed in the instructor's manual, two points if the issue was strongly related, one point if the issue is slightly related, and zero points if the issue was not described in the manual. It is important to note that because cases were chosen based on the relationship they have with the content of the ID textbook, initial small groups had less content knowledge and experiences as a reference than later groups. Thus, there is no intention to compare the performance between groups or evaluate who did better in the analysis of the ID cases.

Table 2

Examples of the Coding Scheme for the Type of Issues Identified

ID Case Studies	Issues Identified by Experts	Issues Identified by Students	Points
Case 7. Implementing New Instructional Approaches in a K-12 Setting	- Needs assessment - Change management - Instructional strategies	- Needs analysis: Not all stakeholders were interviewed. Not all needs were verified and supported by data.	3
		- Context analysis: Community's resistance to education was acknowledged but investigated or addressed	2
		- Theoretical context: Learning context does not adequately consider performance context but Ruth Ann is resistant to change her teaching style	2
Case 18. Designing Curriculum for Southeast Asian Trainers	- Instructional strategies - Learner/cultural analysis - Assessment	- Theoretical Context: Singaporean trainers prefer instructivist approach while the US trainers prefer constructivist/connectivist	2
		- Cultural Context: The Singaporean culture was not fully researched before interviews with the trainers therefore the US instructional designers were unable to collect stakeholders (Singaporean trainer) expectations	3
		- Learner Analysis: An inadequate learner analysis resulted in unmotivated Singaporean learners	3
Case 28. Managing Training in a Manufacturing Setting	- Managing company-wide training - Learner analysis - Needs assessment - Diversity/language needs	- Learner Analysis: Language barriers impact how training is completed and how target audience achieve necessary certifications.	3
		- Needs Analysis: Too many employees receiving training causing a shortage of people performing their assigned jobs.	3
		- Learner & Performance Context: There are no guidelines for implementing peer to peer training. Technicians do not have equal opportunities for training.	0
		- Context Analysis: The theoretical context between peer trainers varies on how they assess trainee learning.	0

To answer the second question (RQ2: What are students' perceptions of case-based VoiceThread presentations with regard to improving their learning of instructional design?) one open-ended question and three five-point Likert scale questions (1=Strongly disagree; 5=Strongly agree) related to this activity were added to the course evaluation survey:

1. How have case-based analysis and discussions played a role in your overall learning experience? (open ended question)
2. Creating a VoiceThread presentation with my group improved my understanding of the case(s) assigned.

3. VoiceThread presentations supported my understanding of all three cases assigned each week.
4. VoiceThread presentations helped me analyze the cases discussed in the Moodle forums.

Answers to the open-ended questions were analyzed using inductive coding (Miles, Huberman, & Saldaña, 2014). As defined by Thomas (2006), "inductive analysis refers to approaches that primarily use detailed readings of raw data to derive concepts, themes, or a model through interpretations made from the raw data by an evaluator or researcher." (p. 238).

Results and Discussion

Using the coding scheme presented in Appendix A, 49 entries (issues) presented by the small groups were analyzed. Table 3 presents the descriptive data showing the detailed breakdown of the scores for each of the five groups on each one of the categories of the ID cases. The average score for each presentation was approximately 2 points out of 3. The results of the content analysis showed that, on average, the issues identified for the small group members in each of the three case studies were "strongly similar" to the main issues identified by the authors of the ID casebook. The results support the notion that creating a small group discussion and requiring students to develop a VoiceThread presentation following scaffolding guidelines to analyze ID case studies helped students find relevant issues about the cases. As discussed by Kim and Hannafin (2008), peer collaboration in case-based activity helps individuals to generate and share ideas, and practice articulating those ideas.

Table 3
Scores for Each ID Case Category

Groups	K-12	Higher Education	Business	Average points
Group 1 (n=4)	2.3	2.6	1.5	2.1
Group 2 (n=4)	1.3	2.0	2.6	2.0
Group 3 (n=4)	1.3	2.3	1.3	1.7
Group 4 (n=4)	2.0	2.5	2.3	2.3
Group 5 (n=5)	1.3	1.3	2.3	1.6
Total	1.64	2.08	2.10	1.95

0=Issue identified is different; 3=Issue identified is the same

Answers from the Likert-scale questions about the use of VoiceThread presentations (Table 4) showed the relevance of watching small group presentations before the whole-group discussion. Students agreed that the presentations helped them to better understand the three cases assigned each week and to analyze the case assigned in the Moodle discussion forums more effectively. Additionally, students agreed that developing a VoiceThread presentation with the group improved their understanding of the assigned cases. These results confirmed the findings in the literature that learning and knowledge building in PBL environments is a collaborative experience (Hmelo-Silver & DeSimon, 2013).

Table 4
Students' Perceptions of Small-Group VoiceThread Presentations

Question	*M* (n=18)
VoiceThread presentations supported my understanding of the cases assigned each week.	4.22
VoiceThread presentations helped me analyze the case assigned in the discussion forums more effectively.	4.11
Creating a VoiceThread presentation with my group improved my understanding of the case(s) assigned.	4.17

1=Strongly disagree; 5=Strongly agree

Finally, 13 students' responses to the open-ended question on how case-based analysis and discussions played a role in their overall learning experience were inductively analyzed to determine common patterns or central themes. Most of the respondents indicated a positive view of the relevance of this activity in their instructional design learning as discussed by Carr-Chellman (1999). The most prevalent theme was the helpfulness of the instructional activity. The following students' opinions are some examples that illustrate this aspect. One student said, "The case-based analysis and discussions helped me understand the concepts better. I was able to listen to and read everyone else's interpretations and it helped me see different points of view." Another student said, "Definitely. They really helped me feel like I could step into the role of being an ID. The readings and discussions were extremely beneficial!" In addition, a student reported, "The case-based analysis and discussions were great! I thoroughly enjoyed the forum discussions. I felt that the VT [VoiceThread] presentations helped when I watched them before and after reading the cases. Overall, I like VT, but I felt that I learned a lot more from the forum discussions."

The use of case studies as real examples was another common theme raised by students that demonstrated the strength of this activity. As one student expressed, "Having real world examples to dissect was very helpful. Also, it was interesting to see how different students interpreted the problem and solution to various cases. It was like having a large think tank." Another student pointed out the variety of cases, stating that "They provided real-world examples in multiple areas; i.e. K-12, Higher Ed, and Industry." Finally, a few respondents reported that other activities were more relevant to their learning experience. One student thought the VoiceThread presentations were "Not as helpful as the actual project." Another student reported, "Overall, I like VT [VoiceThread], but I felt that I learned a lot more from the forum discussions."

Recommendations

Based on the results of this study and the experiences designing small group discussions on case-based instruction, several recommendations can be offered. First, implementing a small group discussion of a specific case study prior to the whole-class discussion can be an effective instructional strategy in online learning environments. In this study, members of the small groups were also the leaders of the whole-class online discussion forums. Since students in the small group analyzed the cases together previously, they could offer stronger feedback to their classmates during discussions. Second, as concluded by Lowes (2014), requiring unique contributions for each group member is key to the relevance of the small group work. Asking that each student present some analysis of the cases in VoiceThread was necessary to collect the points of view of different students as well as promote active participation of all students in the groups.

In addition, the integration of a Google presentation and VoiceThread as available spaces for collaboration among the group members facilitated the asynchronous communication among students. The small groups' VoiceThread presentations were posted for the whole class to view, increasing the potential learning benefits that come from student content creation and sharing (Bennett, Bishop, Delgarno, Waycott, & Kennedy, 2012). Finally, although it was not implemented in this study, a synchronous meeting of the instructor with members of the small group prior to the creation of the VoiceThread presentation is recommended. In this study, the examination of the presentations shows an acceptable level of analysis; however, low levels of critical thinking on some of the issues presented by the students were also present. An initial synchronous conversation with the small group about the case studies following the guidelines provided by Ertmer et al. (2009) could help students to provide stronger arguments about the issues and possible solutions to the different case studies.

Conclusions

Case-based instruction is an important strategy that has been widely utilized in areas such as law, medicine, nursing, and teacher education. The use of this strategy in distance education is important for developing students' critical thinking and problem solving skills in addition to improving communication and collaboration skills (Pena-Shaff & Altman, 2009; Rourke & Anderson, 2002). This study aimed to contribute to the online learning research and practice through exploring the design of PBL environments using case-based scenarios, as well as learners' perceptions of small group presentations to promote learning. This investigation also explored the formation of small groups to analyze instructional design cases and develop VoiceThread presentations that summarized the issues and possible solutions to three assigned cases.

Results supported previous findings that small group activities centered around case studies can enhance student learning. Specifically, our research confirmed that the small group activity involving the presentation of the case studies' analyses before the whole group discussion is a relevant strategy in distance learning environments. For educators in the field of instructional design, this research contributes to the literature by presenting an example of how small group discussions using VoiceThread provide PBL experiences in an online environment.

Finally, a limitation of this study is that data was not collected on participants' interaction in these small groups. Since the level of students' interaction in small groups is related to increased understanding (Webb, 1989), future studies with a similar design are encouraged to observe the interactions among the members of the groups and confirm the collaborative experience in the PBL online environment. In addition, results from this study need to be interpreted with caution due to the small number of participants and the specific learning context (i.e., adult learners in an online learning environment). Additional research with different types of students (K-12 and/or undergraduate) and different content knowledge is recommended to confirm these results.

References

Argyris, C. (1980). Some limitations of the case method: Experiences in a management development program. *Academy of Management Review, 5*(2), 291–298.

Barnes, L. B., Christensen, C. R., & Hansen, A. J. (1994). *Teaching and the case method: Text, cases, and readings* (3rd ed.) Boston, MA: Harvard Business School Press.

Benbunan-Fich, R., & Hiltz, S. R. (1999). Educational applications of CMCS: Solving case studies through asynchronous learning networks. *Journal of Computer-Mediated Communication*, *4*(3). Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1083-6101.1999.tb00098.x/full

Bennett, S., Bishop, A., Delgarno, B., Waycott, J., & Kennedy, G. (2012). Implementing web 2.0 technologies in higher education: A collective case study. *Computers and Education, 59,* 524-534.

Blumenfeld, P. C., Marx, R. W., Soloway, E., & Krajcik, J. (1996). Learning with peers: From small group cooperation to collaborative communities. *Educational Researcher,* 25(8), 37-40

Borup, J., West, R. E., & Graham, C. R. (2013). The influence of asynchronous video communication on learner social presence: A narrative analysis of four cases. *Distance Education. 34,* 48-63.

Carr-Chellman, A. A. (1999). I have a problem! The use of cases in educating instructional designers. *TechTrends, 43*(6), 15-19.

Ching, Y.-H., & Hsu, Y.-C. (2013). Collaborative learning using VoiceThread in an online graduate course. *Knowledge Management & ELearning*, *5*(3), 298–314.

Ertmer, P. A., Quinn, J. & Glazewski, K. D. (2014a). *The ID casebook: Case studies in instructional design* (4th ed.). Upper Saddle River, NJ: Pearson.

Ertmer, P. A., Quinn, J. & Glazewski, K. D. (2014b). *Instructor's resource manual for the ID casebook: Case studies in instructional design* (4th ed.). Upper Saddle River, NJ: Pearson.

Ertmer, P. A. & Russell, J. D. (1995). Using case studies to enhance instructional design education. *Educational Technology, 35*(4), 23-31.

Ertmer, P. A., & Stepich, D. A. (2005). Instructional design expertise: How will we know it when we see it? *Educational Technology, 45*(6), 38–43.

Ertmer, P. A., Stepich, D. A., Flanagan, S., Kocaman-Karoglu A., Reiner, C., Reyes, L., Santone, A. L. & Ushigusa, S. (2009). Impact of guidance on the problem-solving efforts of instructional design novices. *Performance Improvement Quarterly, 21*(4), 117–132.

Flynn, A. E., & Klein, J. D. (2001). The influence of discussion groups in a case-based learning environment. *Educational Technology Research and Development, 49*(3), 71-86.

Hardré, P. L., Ge, X., & Thomas, M. K. (2006). An investigation of development toward instructional design expertise. *Performance Improvement Quarterly, 19*(4), 63-90.

Hmelo-Silver, C. E., & DeSimone, C. (2013). Problem-based learning: An instructional model of collaborative learning. In C. E. Hmelo-Silver, C. A. Chinn, C. K. Chan, & A. O'Donnell (Eds.), *The international handbook of collaborative learning* (pp. 370-385). New York, NY: Routledge.

Honan, J. P., & Rule, C. S. (2002) *Using cases in higher education: A Guide for faculty and administrators*. San Francisco, CA: Jossey-Bass.

Hsieh, H-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research, 15*(9), 1277-1288.

Jonassen, D. H. (2011). *Learning to solve problems: A handbook for designing problem-solving learning environments*. New York, NY: Routledge.

Jonassen, D. H. (1999). Designing constructivist learning environments. In C. M. Reigeluth (Ed.), *Instructional-design theories and models* (2nd ed.) (pp. 215-239). Mahwah, NJ: Lawrence Erlbaum.

Jonassen, D. H. & Hernandez-Serrano, J. (2002). Case-based reasoning and instructional design: Using theories to support problem solving. *Educational Technology Research and Development, 50*(2), 65-77.

Kim, H. & Hannafin, M.J. (2008). Grounded design of web-enhanced case-based activity. *Educational Technology Research and Development, 56*(2), 161-79.

Larson, M. B., & Lockee, B. B. (2014). *Streamlined ID: A practical guide to instructional design*. New York, NY: Routledge.

Lee, S.-H., Lee, J., Liu, X., Bonk, C. J., & Magjuka, R. J. (2009). A review of case-based learning practices in an online MBA program: A program-level case study. *Educational Technology & Society, 12*(3), 178–190.

Lowes, S. (2014). How much "group" is there in online group work? *Online Learning Journal, 18*(1). Retrieved from http://olj.onlinelearningconsortium.org/index.php/olj/article/view/373/82

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis. A methods source*. (3rd Ed.). Thousand Oaks, CA: SAGE.

Pena-Shaff, J. B., & Altman, W. S. (2009). Case-based instruction using asynchronous online discussions: A synthesis. *Journal on Excellence in College Teaching, 20*(3), 97-121.

Potter, W. J., & Levine-Donnerstein, D. (1999). Rethinking validity and reliability in content analysis. *Journal of Applied Communication Research, 27*, 258-284.

Rourke, L., & Anderson, T. (2004). Validity in quantitative content analysis. *Educational Technology Research and Development, 52*(1), 5–18.

Rourke, L., & Anderson, T. (2002). Using web-based, group communication systems to support case study learning at a distance. *International Review of Research in Open and Distance Learning, 3*(2), 1–13.

Saleewong, D., Suwannatthachote, P. & Kuhakran, S. (2012). Case-based learning on Web in higher education: A Review of empirical research. *Creative Education, 3*, 31-34.

Stepich, D. A., & Ertmer, P. A. (2009). Teaching instructional design expertise: Strategies to support students' problem-finding skills. *Technology, Instruction, Cognition, and Learning, 7,* 147-170.

Thomas, D. R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation, 27*(2), 237-246.

Wasserman, S. (1994). *Introduction to case method teaching: A guide to the galaxy*. New York, NY: Teachers College Press.

Webb, N. M. (1989). Peer interaction and learning in small groups. *International Journal of Educational Research*, 13(1), 21-39.

Wentzel, K. R., & Watkins, D. E. (2011). Instruction based on peer interactions. In R. E. Mayer & P. A. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 322-343). New York, NY: Routledge.

Utilization of an Educational Web-Based Mobile App for Acquisition and Transfer of Critical Anatomical Knowledge, Thereby Increasing Classroom and Laboratory Preparedness in Veterinary Students

Kevin Hannon

Department of Basic Medical Sciences College of Veterinary Medicine, Purdue

Abstract

Contact time with students is becoming more valuable and must be utilized efficiently. Unfortunately, many students attend anatomy lectures and labs ill-prepared, and this limits efficiency. To address this issue we have created an interactive mobile app designed to facilitate the acquisition and transfer of critical anatomical knowledge in veterinary students, thereby increasing classroom and laboratory preparedness. We have found that in contrast to a traditional reading assignment, utilization of such an app to introduce students to a subject area significantly enhanced the initial learning of anatomy and the transfer of that learned material to a related, but novel area. We propose that students using the apps were subsequently better prepared for lecture and lab, than students using the more traditional method of reading a textbook. Exposure of students to a topic prior to lecture and laboratory, using methods that students embrace, can only lead to a more efficient and better educational experience.

Keywords: Mobile educational app; anatomy; class preparedness

Hannon, K. (2017). Utilization of an educational web-based mobile app for acquisition and transfer of critical anatomical knowledge, thereby increasing classroom and laboratory preparedness in veterinary students, *Online Learning 21*(1), 201-208. doi: 10.24059/olj.v21i1.882

Introduction

Anatomical instruction is faced with the same challenges of other curriculum: a reduction in contact time while maintaining an equivalent volume of content to be mastered (Sugand, Abrahams, & Khurana, 2010). Therefore, contact time with students is becoming more valuable and must be utilized efficiently. Unfortunately, many students attend anatomy lectures and labs ill-prepared (“What are we dissecting today?”). This lack of preparedness results in a suboptimal use of available contact periods. Current students resist reading traditional texts as pre-lecture or pre-lab assignments, as they find them time consuming and filled with unfamiliar terminology. Therefore, it was our objective to create a web-based learning mobile app with the purpose of preparing students for anatomy lecture and lab. While not used specifically for class preparedness, utilization of web-based computer-aided instructional resources by anatomy students will result in significantly higher scores on examinations (McNulty, Sonntag, & Sinacore, 2009). Therefore, it is a logical extension to hypothesize that computer-aided instructional resources such as mobile apps will improve anatomy student preparedness. The concept of web-based material for pre-lecture/lab preparation has been introduced previously for a physics curriculum (Novak, 1997). The students' outside-of-class preparation fundamentally affected what happened during the subsequent in-class time together. This type of pre-class preparation also brought a diverse group of students to a similar level before lecture/lab (Novak, 1997; Marrs, & Novak, 2004). A number of studies have outlined medical students' preference for mobile apps to deliver content as opposed to other more traditional resources (Gutmann, Kühbeck, Berberat, Fischer, Engelhardt, & Sarikas, 2015; Sandholzer, Rurik, Deutsch, & Frese, 2014). However fewer have documented the educational effectiveness of mobile web apps. It was the objective of our study to create a web-based anatomy mobile app, and analyze the effectiveness of such an app on the acquisition and transfer of critical anatomical knowledge in veterinary students, thereby increasing classroom and laboratory preparedness.

Method

The veterinary anatomy app used for this study was web-based and created using a commercial mobile app-creation platform (www.activelessonhq.com). Apps created using this platform are device neutral and therefore could be accessed by students on their computers and/or mobile devices. The utilization of mobile devices is an efficient and popular method to disseminate subject matter to students (Trelease, 2008). The app was designed to be interactive, utilize active retrieval in the form of self-guided quizzes (Karpicke & Blunt, 2011; Karpicke & Roediger, 2008), and be simple and intuitive to use. The self-guided quizzes utilized a graphics interface that allowed students to interact with anatomical images on the computer/mobile device monitor. In addition, the answers to the questions with these quizzes could be obtained in two ways: 1) by typing text into an input box, submitting an answer, and getting feedback as to whether the typed answer was correct; and 2) by clicking on a "reveal" button, where the answer was revealed without any text input (Figure 1). To test the efficacy of this app for enhancement of learning of veterinary anatomy, and to examine the importance of the text-input box versus the answer-reveal button, 84 first-year veterinary anatomy students were randomly separated into 3 groups (n=28 for each group). The age range of this group was 20-37 years old, and there were 67 women and 17 men. Each group was given an equivalent amount of time (2 hrs.) to complete the following tasks: A) Group 1- read a textbook assignment; B) Group 2- Use the anatomy app with text input box, and then read the textbook assignment; and C) Group 3- Use anatomy app with the answer-reveal button, and then read the textbook assignment. The reading assignment and the anatomy app were focused on muscles of the canine scapular region. The reading assignment was in a veterinary anatomy text that had been required for the course for the past 30 years. The app contained interactive images if muscles in situ, isolated muscles and muscle attachments, all information that was included in the reading assignment. The students had no previous exposure to this anatomical area. Upon completion of their assignment, all students were given an identical written exam. The first exam question was a cross-section through the canine scapula. The image in this question was found in the textbook reading assignment, but not in the app. Students were asked to identify and write-in the correct muscles (Figure 2). The second exam question was a lateral schematic view of the muscles of the canine scapular region. This image was found in the anatomy app, but not in the reading. Students were again asked to identify and write-in the correct muscles (Figure 3). The third exam question was a lateral schematic view of a dachshund. Students were asked to sketch in specific muscles in the dachshund image (Figure 4). This image was not found in the textbook reading assignment or in the anatomy app. This question tested the student's ability to transfer information they had obtained from the reading assignment and app from a dog with a more "standard" body shape (boxer), to a dog of a more "atypical" body conformation (dachshund). Transfer of anatomical knowledge is a common hurdle that confronts all students of veterinary anatomy, as they dissect and study not only dogs of different body types and breed, but other species (i.e. cat, horse, cow, and pig).

All exams were blind-graded with respect to the assignment group or student identification. Exam questions 1 and 2 were only scored correct if the answers were spelled correctly. Exam question 3 was examined for the correct placement of each attachment for each muscle, and the proper sketching of the muscle between the attachment points. Partial credit was given on question 3. ANOVA and Tukey's honest significance test were used for data analyses. This study was exempted by the Purdue Institutional Review Board.

Results

As noted above, question 1 asked students to identify canine scapular muscles from a cross-sectional perspective (Figure 2), a task which was covered in the textbook reading, but not in the app. Using the mobile app with either the text-input or reveal-answer function significantly improved scores on Question 1 compared to the read-only group ($F=6.24$, $p<0.01$) (Table 1). There was no difference between the text-input or reveal-answer groups.

Also as described above, question 2 asked students to identify and write-in the correct names of canine scapular muscles from a lateral perspective (Figure 3), an exercise found the app but not in the reading. Using the mobile app with either the text-input or reveal-answer function significantly improved scores on Question 2 compared to the read-only group ($F=26.1$, $p< 0.001$) (Table 1). There was no difference between the text-input or reveal-answer groups.

In question 3 students were asked to sketch muscles into a lateral view of a dachshund (Figure 4), an image not found in the textbook reading assignment or in the app. Using the mobile app with either the text-input or reveal-answer function significantly improved scores on Question 3 compared to the read-only group ($F=40.3$, $p<0.001$) (Table 1, next page).

Table 1. Exam Question Results.

	Question 1			**Question 2**			**Question 3**		
	Read Only	App + Text	App + Reveal	Read Only	App + Text	App + Reveal	Read Only	App + Text	App + Reveal
Range	0 - 6	0 - 6	0 - 6	0 - 5	0 - 7	0 - 7	0 – 3.5	0.5 – 5	1 - 5
Average	1.7[a]	3.4[b]	2.9[b]	1.4[a]	4.8[b]	4.6[b]	1.8[a]	3.9[b]	4.1[b]
SD	1.54	1.69	1.98	1.35	2.33	2.02	1.06	1.21	0.92
SE	0.29	0.31	0.37	0.26	0.44	0.38	0.20	0.22	0.17
p	-	**	*	-	***	***	-	***	***

Note. Possible ranges were: 0-7 for Questions 1 and 2, and 0-5 for Question 3. Superscripts indicate homogeneous subsets based on Tukey's HSD; *p* indicates the significance level of the comparison between each intervention group and the Read Only group based on Tukey's HSD: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. n=28/group

In summary, using the mobile app significantly improved scores on Questions 1, 2 and 3 when compared to the group that only read the text assignment. Scores within the app groups were significantly higher even on question 1, which contained an image found in the reading and not on the app. Exam scores obtained when using the app with the text input function were not significantly different from when using the app with the reveal answer function. These results demonstrate that the app significantly enhanced the initial learning of anatomy (Questions 1 and 2) and the transfer of that learned material to a related, but novel area (Question 3).

Analysis of the student opinion of the app was analyzed by asking "I felt that the app helped me understand the reading assignment". The students selected a numerical answer based on a scale spanning from 1 = yes, to 5 = no. Students reported feeling that both apps helped in understanding the reading assignment (App+Text mean = 1.2, SE = 0.13; App+Reveal mean = 1.4, SE = 0.13). These results demonstrate that students will accept, utilize and appreciate a mobile learning app when assigned as a precursor to a reading assignment.

Limitations

One of the limitations of our study design was that we did not perform a pretest of student knowledge of front limb muscle anatomy before they performed the exercises described in this manuscript. We did perform a general anatomy (none of the questions referred to front limb anatomy) pretest on this group of students and found no differences between student groups (data not shown). So the potential exists, albeit highly unlikely, that there was a significant knowledge difference between groups pertaining to front limb muscle anatomy that existed before the experiment. The design of this experiment could have been strengthened by having a pretest and examining and analyzing individual student improvement.

Having mentioned that lack of a pretest might have been a design weakness, it is also important to note that one advantage of our design of not having a pretest is that the students are not clued into "what is important to know" pertaining to the upcoming experiment.

Discussion

It was our objective to create an interactive mobile app designed to facilitate the acquisition and transfer of critical anatomical knowledge in veterinary students, thereby increasing classroom and laboratory preparedness. We found that utilization of such an app to introduce students to a subject area significantly enhanced the initial learning of anatomy and the transfer of that learned material to a related, but novel area. We propose that the students using the apps were subsequently better prepared for lecture and lab than the students just reading the textbook assignment. The increased knowledge these students demonstrated on the written exams would clearly help them follow discussion in lecture and facilitate reading a dissection guide and perform dissections. To our surprise, utilization of an app containing a more active form of learning (text input), was not a significantly more effective tool for obtaining anatomical knowledge than an app with a more passive method of learning (answers revealed without input). We hypothesized that using an app with a text input box would be an important consideration when introducing new terminology or when "spelling counts". Regardless, our results support those of McNulty et. al that utilization of web-based computer-aided instructional resources by anatomy students will result in significantly higher scores on examinations (McNulty, Sonntag, & Sinacore, 2009). In addition as reviewed by Guze, (2015) apps can contribute to educational goals of using technology in medical education include facilitating basic knowledge acquisition. Another positive about mobile applications is student acceptance. Mobile apps are preferred digital learning resources by medical students Gutmann, Kühbeck, Berberat, Fischer, Engelhardt, & Sarikas, 2015; Sandholzer, Rurik, Deutsch, & Frese, 2014).

While our study focused on the initial acquisition of anatomical knowledge, our app could still be easily used as a study review tool. The reasons why an app would enhance the initial learning of anatomy could be contributed to many factors, including a simple increase in exposure time to the topic, increased student acceptance when compared to a traditional reading assignment, or by simply being another study option. Studies have shown that successful anatomy students use a multitude of study methods while struggling students relied on single methods (Ward & Walker 2008). In addition, the apps utilized retrieval practice which has been shown to optimize learning (Karpicke & Blunt, 2011; Karpicke & Roediger, 2008). Regardless, exposure of students to a topic prior to lecture and laboratory, using methods that students embrace, can only lead to a better and more efficient educational environment.

References

Gutmann, J., Kühbeck, F., Berberat, P. O., Fischer, M. R., Engelhardt S, & Sarikas A. (2015). Use of learning media by undergraduate medical students in pharmacology: a prospective cohort study. *PLoS One* 10, e0122624.

Guze, P. A. (2015). Using technology to meet the challenges of medical education. *Trans Am Clin Climatol Assoc*, 126, 260-270.

Karpicke, J. D., & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science*, 331, 772-775.

Karpicke, J. D., & Roediger, H. L., 3rd. (2008). The critical importance of retrieval for learning. *Science*, 319, 966-968.

Marrs, K. A., & Novak, G. (2004). Just-in-Time teaching in biology: Creating an active learner classroom using the internet. *Cell Biology Education*, 3, 49-61.

McNulty, J. A., Sonntag, B., & Sinacore, J. M. (2009). Evaluation of computer-aided instruction in a gross anatomy course: a six-year study. *Anat Sci Ed*, 2, 2-8.

Novak, E. T. (1997). World wide web technology as a new teaching and learning environment. *Int J Mod Phys C*, 8, 19-39.

Sugand, K., Abrahams, P., & Khurana, A. (2010). The anatomy of anatomy: a review for its modernization. *Anat Sci Ed*, 3, 83-93.

Sandholzer, M., Rurik, I., Deutsch, T., & Frese T. (2014). Medical students' expectations towards an implementation of a family medicine textbook as a comprehensive app in Germany. *J Med Syst*, 38, 125-134.

Trelease, R. B. (2008). Diffusion of innovations: smartphones and wireless anatomy learning resources. *Anat Sci Ed*, 1, 233-239.

Ward P. J., & Walker J. J. (2008). The influence of study methods and knowledge processing on academic success and long-term recall of anatomy learning by first-year veterinary students. *Anat Sci Ed*, 1:68-74.

Figure Legends

Figure 1. Screen-shot of an example of a quiz question contained within the anatomy app. An answer can be generated by either: 1) typing an answer into the text-input box and receiving automatically graded input on your typed answer by tapping the “Check” button; or 2) by tapping on the “Reveal” button.

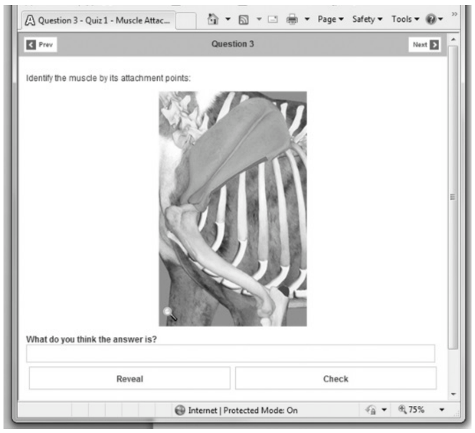


Figure 2. Quiz question 1. The image was found in the traditional reading assignment.

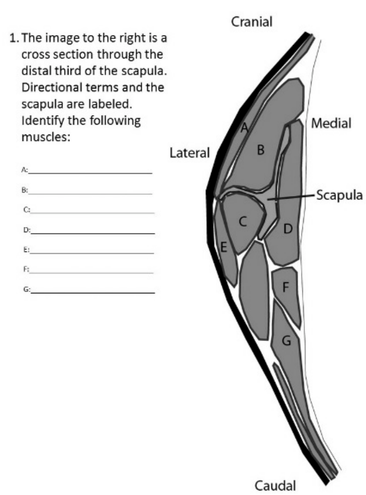
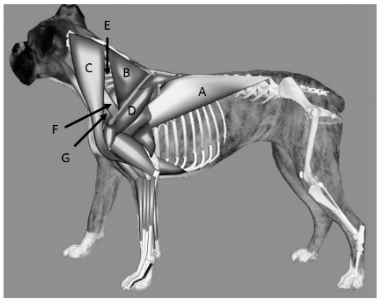


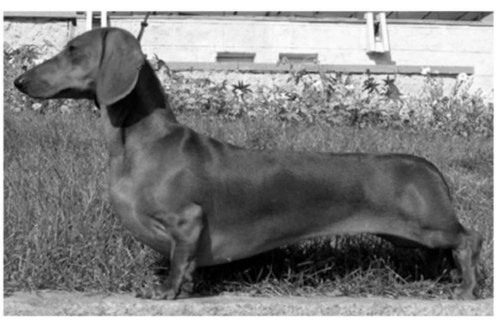
Figure 3. Quiz question 2. The image was found only in the app.



2. Identify the following muscles in the image above:

A:____________________
B:____________________
C:____________________
D:____________________
E:____________________
F:____________________
G:____________________

Figure 4. Quiz question 3. The students were not exposed to this image in the reading or the app.



3. Sketch in and label the following muscles on the image above:

a. Rhomboideus
b. Trapezius
c. Omotransversarius
d. Deep pectoral
e. Latissimus dorsi

A Critical Review of the Use of Wenger's Community of Practice (CoP) Theoretical Framework in Online and Blended Learning Research, 2000-2014

Sedef Uzuner Smith
University of Houston Downtown

Suzanne Hayes and Peter Shea
State University of New York, Albany

Abstract

After presenting a brief overview of the key elements that underpin Etienne Wenger's communities of practice (CoP) theoretical framework, one of the most widely cited and influential conceptions of social learning, this paper reviews extant empirical work grounded in this framework to investigate online/blended learning in higher education and in professional development. The review is based on integrative research approaches, using quantitative and qualitative analysis, and includes CoP oriented research articles published between 2000 and 2014. Findings are presented under three questions: Which research studies within the online/blended learning literature made central use of the CoP framework? Among those studies identified, which ones established strong linkages between the CoP framework and their findings? Within this last group of identified studies, what do the patterns in their use of the CoP framework suggest as opportunities for future research in online teaching and learning?

Keywords: community of practice; Wenger; online and blended learning.

Smith, S. U., Hayes, S., & Shea, P (2017). A critical review of the use of Wenger's Community of Practice (CoP) theoretical framework in online and blended learning research, 2000-2014, *Online Learning 21*(1), 209-237. doi: 10.24059/olj.v21i1.963

Introduction

A great deal of empirical research investigating the use of online and blended approaches in higher education and professional development has drawn primarily on social constructivist theories of learning (Vygotsky, 1978). In many instances, this research was directly inspired by Lave and Wenger's (1991) and Wenger's (1998) theoretical claims, and in others it was motivated by the assumptions put forth by other influential social learning theorists. In this review, we look at published research studies where Wenger's communities of practice (CoP) theoretical framework provided a conceptual direction for the investigation of online and blended learning environments in higher education and in professional development. Our purpose is to examine critically the ways in which these studies used the CoP framework to research online/blended learning.

The impetus for this review came from our interest in examining the theory-research links in published studies of online/blended learning. Most major research methods textbooks and articles remind us that theory influences the types of questions (or hypotheses) researchers generate, and consequently it influences the answers obtained from those questions. The following quote from Kilbourn (2006, p. 545) attests to this view:

> A fundamental assumption for any academic research is that the phenomena (data) that we wish to understand are filtered through a point of view (a theoretical perspective) – that is to say, it is assumed that there is no such thing as a value-free or unbiased or correct interpretation of an event. Interpretations are always filtered through one or more lenses or theoretical perspectives that we have for "seeing"; reality is not something that we find under a rock.

There is also the contrasting view that theory is not always needed in research. Yet, what appears to be the general consensus on this point is the idea that "Research that is not theoretically informed, not grounded in the existing body of knowledge, or of the 'shotgun' variety that fails to raise and investigate conceptually grounded questions, is likely to generate findings of a narrow and ungeneralizable value" (Yiannakis, 1992, p. 8). It is this idea that initially gave rise to this review, and we determined our purpose to be that of critically examining how the CoP framework is used in published research studies on online/blended learning in higher education and professional development.

Before embarking on this task, we searched for any existing publications that might have already attempted what we sought to do. Our search yielded no such publication, and we found only one article (Consalvo, Schallert, & Elias, 2015) that came closest to the focus of our review. This article critically examined the use of Lave and Wenger's (1991) construct of legitimate peripheral participation in literacy research. Although its focus and content are quite different from the review we present here, the Consalvo et al. (2015) article provided insights that helped us think through ways to conduct this review.

In conducting this review, we were guided by two goals. One was to critique the ways in which the CoP framework has been used in studies focusing on online/blended learning environments in higher education and professional development and to identify new possibilities for future research. Another was to aid current and future researchers in examining their own application of the CoP framework in detail.

Having described our purpose, we will now provide a brief summary of Wenger's CoP framework. For those readers seeking a detailed presentation of Wenger's ideas, we provide references to Wenger's own writings rather than secondary sources. Lastly, it is important to note that we did not intend to offer a critical analysis of Wenger's CoP framework. Rather, our goal is to provide a critical analysis of how this framework has been used in published research on online/blended learning. For those readers interested in a critical analysis of Wenger's ideas, we recommend the collection of essays in Hughes, Jewson, and Unwin (2007) as a good starting point.

Summary of the CoP Framework

When speaking of Wenger's notion of CoP, it is important to note that it has continued to grow in complexity and focus. The initial concept of CoP originated in Wenger's partnership with Jean Lave in their 1991 publication, "*Situated learning: Legitimate peripheral publication.*" In this work, Lave and Wenger used an anthropological perspective to argue that learning is not just receiving or absorbing information. Rather, in their view, learning is "increasing participation in communities of practice" (Lave & Wenger, 1991, p. 49). In his groundbreaking 1998 book, "*Communities of practice: Learning, meaning, and identity,"* focusing on workplace learning, Wenger expanded upon this idea of CoP, articulating how social resources shape people's learning trajectories and their professional identity. Following this publication, Wenger developed the concept of CoP further by presenting it as an approach to knowing and learning that is applicable to various contexts, including business, organizational design, government, education, and civic life. Undoubtedly, Wenger's notion of CoP is one of the most widely cited and influential conceptions of social learning to date.

More recently, Wenger, Trayner, and de Laat (2011) defined CoP as a "learning partnership among people who find it useful to learn from and with each other about a particular domain. They use each other's experience of practice as a learning resource" (p. 9). Taking this definition as our starting point, below we briefly explore the important concepts that underpin the principles of CoP.

The Domain

For Wenger (2004), the domain of a CoP constitutes "the area of knowledge that brings the community together, gives it its identity, and defines the key issues that members need to address" (para. 13). The domain, therefore, is what gives a group its identity and distinguishes it from a club of friends or a network of connections between people.

The Community

For Wenger (2004), the community constitutes "the group of people for whom the domain is relevant, the quality of the relationships among members, and the definition of the boundary between the inside and the outside" (para. 14). For a group of people to constitute a CoP, its members must come together around ideas or topics of interest (the domain) and interact with each other to learn together.

The Practice

Wenger (2004) defines practice as "the body of knowledge, methods, tools, stories, cases, documents, which members share and develop together" to address recurring problems in their specific contexts (para. 15). To our knowledge, the most recent attempt to define this construct from a Wengerian perspective comes from Consalvo et al. (2015). These authors defined practice as "a way of acting in the world" and as "a field of endeavor and expertise" (p. 3). Combined, these definitions suggest that practice implies knowledge of and engagement with a domain.

Participation and Reification

Wenger (1998) contended that individuals' engagement in a CoP always entails a process of negotiation of meaning which takes place in the convergence of two processes: participation and reification. Participation involves acting and interacting, and reification involves producing artifacts (such as tools, words, symbols, rules, documents, concepts, theories, and so on) around which the negotiation of meaning is organized. Participation and reification are complementary processes in that each has the capacity to make up for the limitations of the other. For instance, when reading about an idea does not make it clear to an individual, peers who have a better grasp of it may become a source for the individual's understanding through conversation, a form of participation. In the same way, giving shape to an idea through writing (a form of reification) may enhance one's meaning making in ways that discussing it with other people could not. Wenger, White, and Smith (2009) noted that learning in a CoP "requires both participation and reification to be present and in interplay" (p. 57).

Joint Enterprise, Mutual Engagement, and Shared Repertoire

Wenger (2010) emphasized that over time, through participation and reification, participants of a CoP develop and negotiate "a set of criteria and expectations by which they recognize membership" (p. 180). These criteria include:

- *joint enterprise* - a collective understanding of what the community is about, its purpose
- *mutual engagement* - interacting and establishing norms, expectations, and relationships; and
- *shared repertoire* - using the communal resources, such as language, artifacts, tools, concepts, methods, standards.

Wenger (1998, p. 137) posited that it is through joint enterprise, mutual engagement, and shared repertoire that a community establishes guidelines as to "what it is to be a competent participant, an outsider, or somewhere in between" and further adds that establishing such guidelines is crucial for learning to take place in a CoP.

Engagement, Imagination, and Alignment

According to Wenger, as people participate in a CoP, they express their belonging through three modes of identification:

- *engagement* – doing things together, talking, producing artifacts;
- *imagination* – reflecting, constructing an image of the practice and its members and seeing self as one of them;
- *alignment* – following directions, aligning self with expectations/standards, coordinating actions towards a common goal.

Wenger posited that these three modes of identification are not mutually exclusive and their presence is crucial to the transformation of a CoP into a site of learning. He noted, "The creation of learning communities […] depends on dynamic combination of engagement, imagination, and alignment […]" (Wenger, 1998, p. 228).

Boundaries

People often belong to more than one CoP with each having boundaries that separate them from one another. In Wenger's view, boundaries connote difference: "They arise from different enterprises; different ways of engaging with one another; different histories, repertoires, ways of communicating, and capabilities" (Wenger, 2000, p. 125). In other words, being members of multiple CoPs means crossing boundaries.

Brokering

Crossing boundaries between different communities provides opportunities for brokering, a concept Wenger (1998) defined as the process of "transfer[ring] some element of one practice into another" (p. 109). He further added that good brokers are those that cause learning as they engage in import-export.

Legitimate Peripheral Participation

When individuals cross boundaries as outsiders or newcomers, they are offered possibilities for participation called peripheries. A newcomer's participation in a CoP often starts on the periphery – "a region that is neither fully inside nor fully outside" (Wenger, 1998, p. 117) and leads towards the center through growing involvement. This process of moving from the periphery to center is characterized by the concept of legitimate peripheral participation – a concept that was first developed by Lave and Wenger (1991). In Wenger's writings, the notion of legitimate peripheral participation is mentioned but it does not take center stage. Rather, it serves as important background condition under which newcomers become included in a CoP. Wenger's contribution to the development of this notion lies in his articulation of the special measures (e.g., observation, special assistance, close supervision, etc.) that may be taken to open up a practice to newcomers. He also noted, "No matter how the peripherality of initial participation is achieved, it must engage newcomers and provide a sense of how the community operates" (Wenger, 1998, p. 100).

Identity

Identity construction as a result of participating in and learning from the practices of a community is another topic that is initially explored in Lave and Wenger (1991) and further elucidated in Wenger's (1998) later work. Wenger reminded us that as people participate in a CoP, they acquire new knowledge and simultaneously their sense of who they are, their identities, change. As he stated:

> Because learning transforms who we are and what we can do, it is an experience of identity. It is not just an accumulation of skills and information, but a process of becoming – to become a certain person or, conversely, to avoid becoming a certain person (Wenger, 1998, p. 215).

Knowledge

Participants in CoP generate knowledge as they interact with each other, share information, experience, insight and advice and help each other solve problems. Over time, this combination of action and discourse eventually represents communal approaches to understanding and solving problems, and the process of reification transforms these shared knowledge into the tools and artifacts that embody a CoP's regime of competence. The community's knowledge is dynamic, not static. It is also explicit and tacit, as well as social and individual (Wenger, McDermott & Snyder, 2002).

Learning Architectures

In his discussion of learning as participation and becoming, Wenger (1998) introduced four dualities to capture the general elements for designing learning in CoPs. These dualities are: (1) participation and reification; (2) designed/emergent; (3) local/global; and (4) identification/negotiability.

The first duality reminds us of the need to hold doing/talking (participation) and producing objects (reification) in the correct proportion to each other in social learning systems. The second duality expresses the need to include improvisation and innovation (emergent) into the prescriptions of practice (designed), such as policies and plans. The third duality highlights the need to involve "those who organize learning and those who realize it" in the design of learning (Wenger, 1998, p. 234). The fourth duality expresses the need to distribute power to shape both the community and the individual. Along with these dualities, Wenger emphasized that a robust design for learning should involve:

- interactive technologies, communication facilities, joint tasks, availability of help, and peripherality (indication of engagement);
- transparency, explanations, reflection, and pushing boundaries (indication of imagination); and
- common focus, direction, plans, standards, policies, and distribution of authority (indication of alignment).

Value Creation

Wenger's later writings (Wenger, Trayner, & de Laat, 2011) presented the concept of value creation as a way to describe and assess the nature of social learning in a CoP and what, if any, value is created as a result of CoP members' activities and in their interactions with others in informal networks. The primary recipients of this value are participants of a CoP, but value may also accrue to other stakeholders, such as the organizations in which CoP operate and their sponsors who invest resources.

Wenger et al. (2011) defined five different cycles of value creation generated within CoP: immediate value; potential value, applied value; realized value; and reframed value. Immediate value includes learning that is put to use immediately to solve a problem. Potential value includes benefits related to the shared skills and knowledge that can be realized at some time in the future. Applied value results from the application of shared skills and knowledge to new contexts. Realized value includes CoP participant and stakeholder reflections on how the skills and knowledge gained as a result of their participation in a CoP made a difference in their ability to achieve important goals. Lastly, reframed value involves the identification and definition of new criteria for success.

Methodology

Analytic Framework for the Review

This review takes the form of an integrative research review, a type of literature review that comes closest to fulfilling the methodological requirements of traditional research. As stated by Szmigiel and Lee (2014), an integrative review consists of "five stages comparable to those in empirical research: research question formulation, data collection, data evaluation, data analysis, and interpretation and reporting" (p. 37). The process we followed for this review mirrored these stages. We began by developing the overarching questions that provided the boundaries for the review. Next, we searched for and selected the studies relating to our inquiry. We treated each study as a data source and used both qualitative and quantitative approaches to achieve a systematic data analysis procedure.

In the next sections, we provide our guiding questions, search strategy, and analytical approach followed by our findings and interpretations.

Guiding Questions

The overarching questions that guided this review were:

1. Which research studies within the online/blended learning literature made central use of the CoP framework?
2. Among those studies identified, which ones established strong linkages between the CoP framework and their findings?
3. Within the final group of studies identified, what do the patterns in their use of the CoP framework suggest as opportunities for future research in online teaching and learning?

Search Strategy

To identify the studies to be included in this review, we conducted a comprehensive search using six aggregator research database services: EBSCO Academic Search Premier, Gale One Search, ProQuest, EdIT (*Education and Information Technology*), Science Direct, and Sage. We used the following three topics to guide our search: 1) community of practice, 2) Wenger, and 3) online and blended learning. We also developed synonyms and phrases for each topic including their singular, plural and abbreviated forms, and then combined them using the Boolean operator AND. (See Appendix A for complete list of search terms). A limitation of this search was that only studies in the above-mentioned databases were identified. To account for the possibility of exclusion of relevant articles outside of these databases, we searched Google Scholar and used citation chaining.

We targeted research articles from peer-reviewed journals during this search and excluded non-research articles, conference papers, dissertations, books, and book chapters. We also excluded articles published in languages other than English. While we set no specific time range for our inclusion criteria, the latest publications identified for the review were from 2014 given that we conducted our search during the summer of 2015.

Initially, our search yielded 82 research studies. The majority of these focused on online/blended courses or programs in higher education, and some focused on professional development that uses online/blended delivery. From this latter group, we selected only those papers that had a community focus and represented formal learning experiences bounded by time limits as well as predetermined, communal goals and outcomes. We excluded papers in which online/blended professional development environments represented “affinity spaces” (Gee, 2005) where participants come and go as they please to connect with each other, to share personal/vocational interests/passions, and to learn something connected to those interests/passions. Upon completion of this selection process, we ended up with 60 research articles. This became our sample.

Analytical Approach

To answer the first two questions guiding this review, we used content analysis and coded all 60 articles by using coding schemes that we developed for the purposes of this review. During the content analysis, we evaluated our coding decisions for inter-rater reliability using Holsti's (1969) coefficient of reliability.

To determine which articles made central use of the CoP framework (Guiding question 1), we first developed the ‘Theoretical Foundation’ coding scheme (Table 1) based on Bates & Taylor’s (2013) argument that, “the quality of theory application depends not so much on where (i.e., literature review, method, discussion) theory is used, but how thoroughly theory is applied to the study” (p. 63). Accordingly, if the CoP theory formed the conceptual framework for a study, either solely or jointly with other theories, we assigned that study “1,” and if the CoP theory was referenced or mentioned but did not provide the conceptual direction for the study, we assigned that study “0.”

Table 1

Theoretical foundation coding scheme

Criteria	Code	Description
Extent of theoretical foundation	1	CoP theory formed the conceptual framework for the study, either solely or jointly with other theories.
	0	CoP theory is referenced or mentioned but did not provide the conceptual direction for the study.

Next, we developed the ‘Theoretical Linkage’ coding scheme (Table 2) to determine which studies established strong linkages between the CoP framework and their findings (Guiding question 2). We developed this scheme based on the argument that theory “serves as the structure and support for the rationale for the study, the problem statement, the purpose, the significance and the research questions,” but most importantly, it provides a grounding base for methods and a conceptual anchor for analysis and findings (Grant & Osanloo, 2012, p. 12). Against this backdrop, we assigned “2” to a study whose analysis/findings clearly connected to CoP theory; assigned “1” to a study whose analysis/findings somewhat or partially connected to CoP theory; and assigned “0” to a study whose analysis/findings were not connected to CoP theory.

Table 2

Theoretical linkage coding scheme

Criteria	Code	Description
Extent of connection between CoP theory and analysis/findings	2	Analysis/findings are clearly connected to CoP theory
	1	Analysis/findings are somewhat or partially connected to CoP theory
	0	Analysis/findings are not connected to CoP theory

To address the third guiding question, we used the data analysis technique of the constant comparison method. This method involved rereading each of the studies identified as having strong/clear theoretical linkages and exploring patterns in their use of the CoP framework. While doing this, we constantly compared patterns that emerged from one study to those that emerged from another. As patterns became apparent, we noted that some were related. We sorted and reclassified them and arrived at six core patterns, which we reported in our findings.

Findings

Guiding question 1. The coding we conducted with our sample of 60 research articles, using guiding question 1, yielded 41 studies that made substantial use of Wenger's CoP framework, either solely or jointly with other theories. In the remaining 19 studies, the CoP framework was referenced or mentioned but did not provide the conceptual direction for the investigation. Initial inter-rater reliability for this analysis was .95 and was negotiated to 1.00.

As can be seen in figure 1, among those 41 studies, the first study on online/blended learning making central use of the CoP theoretical framework dated back to the year 2000. Between 2000 and 2010, 26 papers of this nature were published, paralleling the overall growth in enrollments and scholarship in online/blended learning (Allen & Seaman, 2013). In the subsequent four years, 15 additional papers were published, followed by a peak of seven papers in 2011. It appears that within the domain of online/blended learning, the number of publications using CoP theory has begun to decline.

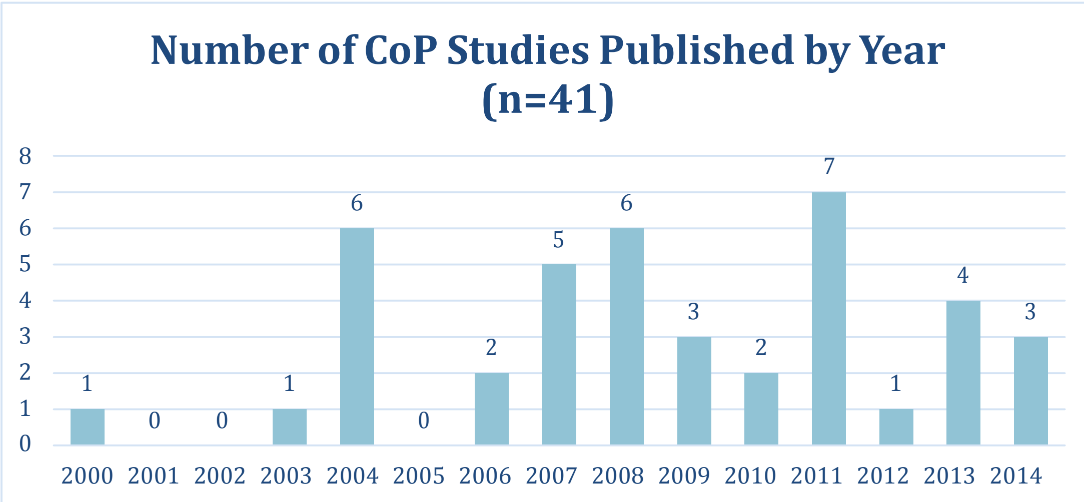


Figure 1. Publication year distribution of CoP-based research studies in online/blended learning

The majority of the studies we located were qualitative (53.7%) and some used mixed methods (36.6%). Only 9.8% of all studies used a quantitative approach. Investigation of undergraduate and graduate level learning (78%) far outnumbered investigation of professional development (22%). Similarly, education courses were most often the site for CoP-based research (65.9%), followed by those in health care and social services (12.2%). The remainder were in other fields (17.1%) or unidentified (4.9%). Studies examining online courses accounted for just 36 % versus 46.3% for blended courses. The remaining 17.7% used an online element but did not explicitly identify whether the course was fully online or blended. Lastly, 78% of the studies focused on the course level, with only 14.6% examining online/blended programs and 7.3% addressing both areas.

Guiding question 2. In the second round of content analysis, we revisited the 41 studies identified for the guiding question 1 to determine if the authors had established strong linkages between the CoP framework and their analysis and findings. We found 17 studies that met this criterion (see Appendix B). In these studies, Wenger's CoP theory was central to the authors' data analysis efforts while also serving as the conceptual lens for interpretation of their findings. The 24 excluded studies failed to ground their analysis and/or results in Wenger's theory because of its absence or cursory use. For this analysis, the initial inter-rater reliability was .88 and was negotiated to 1.00.

Guiding question 3. To address our third guiding question, which focused on identifying patterns in the use of the CoP framework among the final group of 17 studies, we used the data analysis technique of the constant comparison method. Below, we present the six primary patterns/themes that emerged from our analysis as potential avenues that future research may pursue.

Problematization versus theory verification. Overall, the majority of the studies generated from our search terms oriented toward theory verification – that is, they provided empirical confirmation of Wenger's theoretical assumptions. Theory verification was also evident in the final group of 17 studies that we identified for our guiding question 3. In this group, we found six studies showing this trend: Evans, Yeung, Markoulakis, and Guilcher (2014), Gray (2004), Guldberg and Pilkington (2006), Moule (2006), Rogers (2000), and Brosnan and Burgess (2003). The aims of the first five studies coincided: Each sought to understand the extent or the nature of CoP formulation in the online learning environment. A common thread running through these studies was that they looked for evidence of the three essential characteristics of a CoP, i.e., mutual engagement, joint enterprise and shared repertoire, in their data. In the sixth study, Brosnan and Burgess (2003), the validity of Wenger's learning architecture concept was tested against what was seen to work well in an online course. All of these studies were theoretically sound; they verified how Wenger's theoretical assumptions correspond to the ways teaching and learning function in online and blended environments. However, the emphasis placed on theory verification led some of these studies to repeat the same general conclusions. It appeared that while the theory verification approach resulted in the design aspect of these studies being well grounded in CoP theory, it made their findings repetitive of the assumptions and findings that are already present in the literature.

One way of going beyond theory verification and avoiding the production of repetitive findings in research is to use the problematization approach proposed by Sandberg and Alvesson (2010). Sandberg and Alvesson defined this approach as "think[ing] differently, instead of what is already known" and "being able to formulate informed and novel questions" (p. 32). We argue that for CoP grounded online/blended learning research, the strategy of problematization implies disrupting the research emphasis on the verification of the best-known and over-researched Wengerian concepts (such as mutual engagement, joint enterprise and shared repertoire) and opening up new and previously unexplored areas for investigation.

A problematizing attitude, in the way described above, appeared evident in eight of the 17 studies included in our final analysis: Adams (2007); Clarke (2009); Cowan and Menchaca (2014); Goggins, Laffey, and Gallagher (2011); Guldberg and Mackness (2009); Mackey and Evans (2011); Nelson and Temples (2011); and Stacey, Smith, and Barty (2004). For example, drawing on Wenger's theory of identity formation in practice, Adams' (2007) study documented the struggles that ensued when art and design graduates transitioned from being an artist to becoming a teacher. The analysis revolved around analyzing the expressive and confessional nature of these new arts-teachers' online forum posts that revealed the complexities of their experiences. With this focus, the study was able to provide a refreshing perspective on the use of online forums as venues for personal exploration of identity and agency, particularly for newcomers to the teaching profession.

Another COP grounded study that offered a novel perspective on the contributions of online learning in teacher education is Clarke (2008). This study began by critiquing the standards model of teacher education, arguing that it is a flawed model that reduces teaching to a set of competencies to be mastered. Following this argument, it presented Wenger's CoP framework as a "more useful approach with which to analyze the complexity of new teachers' experiences" (p. 522) and illustrated the ways in which online components of a teacher education course can foster aspects of a CoP – namely the community, the domain of the community, community's shared practice, and community members' boundary crossings. Rather than verifying Wenger's CoP theory, this study developed from that theory a model of student teachers' online learning.

In sum, we contend that even though verification of the various elements of Wenger's CoP construct has provided many important insights into online/blended learning processes, there is need for researchers to develop a more problematizing attitude towards their investigations, as was done in Adams (2007) and Clarke (2008), than is currently the case.

Formation of a CoP versus formation of different types of community. More than half of the 17 studies included in our final analysis sought to determine whether the study participants actually attained a CoP (Clarke, 2009; Correia & Davis, 2008; Ellaway, Dewhurst & McLeod 2004; Goggins et al., 2001; Gray, 2004; Guldberg & Mackness, 2009; Guldberg & Pilkington, 2006; Rogers, 2000). Among these, two studies (Correia & Davis, 2008; Guldberg & Pilkington, 2006) stood out because, before the actual analysis, they identified a need for acknowledging how different dynamics lead to different types of communities. To illustrate this point, both studies drew on Henri and Pudelko's (2008) classification of four levels of communities: communities of interest; goal oriented communities; a learner's community; and communities of practice. By considering the strength of a group's social bonds (i.e., its level of cohesion) and the extent of its intentionality (i.e., the demonstrated purposefulness of its efforts) as a starting point, Henri and Pudelko differentiated these communities in the following ways: Communities of interest have the lowest cohesion of collective endeavor because they generate knowledge solely for individual use. Goal-oriented communities are driven by external forces to carry out a particular task within a specified timeframe. A learner's community relies upon the instructor for guidance and results in the generation of both individual and shared products. Finally, communities of practice are organized around professionals who perform similar activities and use their strong social bonds and high levels of intentionality to extend and improve their practices by building a base of shared knowledge or knowledge system(s). What is most notable about Henri and Pudelko's framework illustrating distinctions among communities is that it recognized the idea that learners do not necessarily form a CoP when they are part of a learning environment. We contend that this is an important point that any research seeking to understand the extent of CoP formation in online/blended learning environments should consider.

The dimension of time in CoP formation. Time is an important element in Wenger's CoP framework. The growth of novices into experts as they become enculturated into a CoP's regime of competence necessitates the passage of time. Time is also essential to Wenger's concept of identity. In a recent interview, Wenger noted that "… [I]dentity itself is a time/space concept" … [in that] "you become a person out of a whole series of experiences over time" (Farnsworth, Kleanthous & Wenger-Trayer, 2016, p. 11-12). In Wenger's own publications or in secondary sources that describe or interpret his theories, there is no clear-cut answer to the question of how much time is needed to arrive at a functioning CoP, but a general contention is that "a shared repertoire cannot be rushed into existence" (Cousin & Deepwell, 2005, p 61).

Within the final group of 17 studies, in all but one study (Ellaway et al., 2004) the temporal element of CoP formation was manifest as the duration of the online/blended courses or professional development opportunities that formed the research context. In these studies, the duration of time varied from as brief as 6 weeks (Moule, 2006) to 10 to 16 weeks (Correia &

Davis, 2008; Cross & Pryor, 2008; Goggins et al., 2011; Mackey & Evans, 2011; Nelson & Temples, 2011; Stacy, Smith & Barty, 2004). There were three exceptions that examined longer time frames. One was Clarke (2009), which reported on a 36-week long course that led to post graduate certificate in geography education. The others were Adams' (2007) yearlong study of a cohort of former art students making the transition to student teachers and Cowan and Menchaca's (2014) longitudinal analysis that examined an educational technology graduate program over a ten-year period. Similar variation was present when looking across the studies focusing on professional development. These studies ranged from 3 weeks (Rogers, 2000), 7 weeks (Gray, 2004; Guldberg & Mackness, 2009), and 10 to 12-weeks (Brosnan & Burgess, 2003; Evans et al., 2014), to one year (Guldberg & Pilkington, 2006).

None of the aforementioned studies explored how time has contributed to the establishment of a functioning CoP. It appears that despite its importance in Wenger's CoP framework, time remains an unexplored variable in online/blended learning research grounded in this theory. We suggest that future research attempt to bring the issue of time to the forefront. One way to do this would be to identify how time contributes to, for example, the growth of novices into experts in online/blended course environments. Another way would be to examine how time impacts an online/blended course/program participant's development of an identity as someone who belongs to a CoP.

The need to unpack the epistemic and discursive practices typical of social practices. The idea that learning happens through people's engagement in social practices lies at the heart of Wenger's CoP theory. Nevertheless, as important as social practices are to embodying and sustaining learning and knowledge within a CoP, an articulation of the epistemic and discursive practices typical of the communities that make up a social practice is missing from the literature. Arguing that Wenger's notion of social practices is largely undifferentiated, Amin and Roberts (2008) noted, "It is time that a more heterogeneous lexicon for different types of situated practice was developed" (p. 365). Amin and Roberts' critique emphasized that further clarity is warranted to identify the distinctive properties of learning and knowing that are situated within different types of social practices. We concur with this assessment and believe that more attention is needed to highlight the specialized ways of knowing, thinking, and doing that people need to internalize in order to participate in a particular social practice.

Looking across the final group of 17 studies, we found that only two studies (Crossard & Pryor, 2008; Evans et al., 2014) detailed participants' epistemic engagement. In their exploration of the online components of a blended course where doctoral students began their trajectory from the periphery of educational research to more central roles, Crossouard and Pryor (2008) documented how the students engaged with the disciplinary norms to develop the ability to think and act like researchers. Similar illustrations were also present in the Evans et al. (2014) study, which focused on physical therapists' use of evidence-based approaches in their practice. Because these studies did not set out to explore how one goes about doing things in the practices that formed the context for their investigation, they did not provide a full account of the situated learning and knowing embedded within those practices. Nevertheless, they merit praise for acknowledging, and partially documenting, the epistemic structures that need to be orchestrated to facilitate individuals' entry into particular practices. We contend that this is an area that needs further consideration in future online/blended learning research grounded in Wenger's CoP framework.

As Larreamendy-Joerns and Leinhardt (2006) wrote, participation in a social practice entails understanding "not only its substantive structure (i.e., facts, concepts, theories), but also its syntax – that is, the questions that guide inquiry, the tools that allow inferences and interconnections, and the actions and principles (rules) that validate knowledge" (p. 590). Future research focusing on online/blended learning environments through the lens of CoP theory should consider the identification and articulation of these structures and syntax (i.e., the epistemic and discursive practices) that are valued in specific disciplines and professions. Doing so will provide the beginnings of an understanding of how to better focus learners' social interactions in online/blended learning environments to facilitate their professional socialization.

Using technology tools to support learning activities in CoP focused courses. When introducing technology into a CoP, Wenger, White, Smith and Rowe (2005) warned of the danger of "confusing the community with the technology" (p. 2). In other words, merely establishing an electronic site to host distributed members of an existing or aspiring CoP to engage with each other is no guarantee of its success. The same cautions and principles apply to online and blended learning environments. In these environments, web-based technologies such as asynchronous and/or synchronous discussions typically serve as a means of ensuring learner engagement with each other for the purposes of generating communal knowledge and resources that form their social practice. Nevertheless, just adding these interactive spaces to an online/blended learning environment does not guarantee that the resulting interactions support the kinds of meaning making necessary for the development of a CoP.

Within the final group of 17 studies, nearly all relied on asynchronous discussions to examine the ways in which the various components of Wenger's CoP framework play out in online/blended learning environments. Among these, three studies stood out (Clarke, 2009; Crossouard & Pryor, 2008; Evans et al., 2014). These studies distinguished themselves by their clarity in explicating how those discussions were used to support participants' interaction-based meaning making for the development of community specific practices. Clarke (2009) discussed using online discussion forums for reflective practice as well as for formal and informal sharing of resources. In Crossouard and Pryor (2008), the discussion forum enjoyed use as a space where students could problematize and reflect upon the process of conducting research and engage in peer assessment. Evans et al. (2014) mentioned the use of discussion areas for a series of authentic evidence-based practice learning activities. These three studies provide readers with a clear idea of how technology was used for the purposes of generating communal knowledge and resources. In the remaining studies, however, such clarity was absent.

We contend that in online/blended learning studies that are grounded in Wenger's CoP framework, if asynchronous/synchronous discussions are used as data sources to examine how learning unfolds in these spaces, then the learning activities that are used in those discussions need to be made explicit. Future research should be sensitive to this issue and avoid the tendency to present asynchronous/synchronous discussions as data sources without further explication.

On a related note, we also argue that the heavy reliance upon discussion areas as the most common site of exploration for CoP oriented research suggests that there is a need to focus on other collaboration tools to examine the uptake and enactment of social practices within a CoP. In our final sample of 17 studies that relied on web-based interactive technologies as their focus of analysis, we identified only one study (Goggins et al., 2011) that used a learning management system-based wiki rather than a discussion board. This study demonstrated how wikis offer a viable alternative space to support both participation and reification of shared and negotiated meaning in an online learning environment. We suggest future CoP oriented research in online/blended learning go beyond the analysis of discussions and consider the integration of alternative spaces for studying participation and reification in the generation of social practices and communal knowledge.

Communicating the practical implications of CoP theory with caution. Wenger's theoretical assumptions about CoP constitute a very rich and complex theory that is challenging to apprehend and apply. As such, it is not surprising that in the final group of 17 studies we only found three studies (Brosnan & Burgess, 2003; Rogers, 2000; and Ellaway et al., 2004) that provided practical implications of this theory. The Brosnan and Burgess study provided contextualized accounts of how the key elements of Wenger's learning architecture notion can be employed to evaluate and guide the design and support of a Web-based continuing professional development course. Rogers' study offered guidelines as well as examples of how the principles of Wenger's mutual engagement, joint enterprise, and shared repertoire concepts can be applied in online learning environments to foster cohesive communities. Ellaway et al. (2004) differed from the Bronson and Burgess (2003) and Rogers (2000) studies in that it did not offer straightforward practical implications. Nevertheless, the authors contribute a post-hoc evaluation model for assessing the success and value of a virtual learning environment in supporting the general characteristics of CoP as articulated in Wenger's construct of learning architecture. The 60-item survey presented in this study can be used by instructors, designers, and students in the context of a specific course. As Ellaway et al. (2004) put it, their survey offers "a perspective of how successfully the [virtual learning environment] is serving the communities of practice involved with the course in question, and […] provide[s] pointers to areas in which it could be improved to the benefit of that community" (p. 142).

The fact that we were only able to identify three studies of this nature out of the final 17 speaks to the need for more researchers conducting CoP oriented research on online/blended learning to clearly articulate and demonstrate the practical aspects of their findings. However, we say this with caution. Often there is a tendency for researchers to accept a theory uncritically, and thus the implications they draw out from it can be too neat or too facile. There can be no denying that some applications of the CoP theoretical framework that we see in online/blended learning research are cursory. With this in mind, we argue that future CoP oriented research should go beyond simplistic or cursory applications of Wenger's theoretical ideas by exploring and engaging with this theory in its complexity.

Discussion and Conclusions

In this literature review, our analysis showed how research publications prior to 2015 have used Wenger's CoP framework to investigate online/blended learning in higher education and in professional development. It also described avenues that both current and future researchers employing this framework can pursue in their investigation of online/blended learning environments. Below we summarize our findings and provide overall conclusions and recommendations.

An important finding was that authors of 41 studies made explicit reference to Wenger's CoP framework, stating that this theory provided conceptual direction for their investigation. However, 24 of these studies' use of the CoP framework was questionable when judged against criteria we generated from the guidelines provided by Grant and Osanloo (2012) and Bates and Taylor (2013) on the use of theory in research. Another important finding was that the majority of the remaining 17 studies concentrated on the verification of the best-known elements of the CoP theory: joint enterprise, mutual engagement, and shared repertoire. Few studies went further than this focus. This review therefore argues that online/blended learning research employing the CoP theory should enter a new phase of development. There is a need for studies that not only take up different aspects of Wenger's CoP theory but also go beyond the traditional practice of theory verification to provide more complex and more nuanced understandings of online/blended learning environments.

We have identified certain areas in this review as being worthy of further consideration for future CoP-oriented online/blended learning research. These include:

- moving toward more sophisticated ways of gauging the progress of CoP formation, as seen in Henri and Pudelko's (2008) model;
- considering how the passage of time impacts the establishment of a CoP and/or the process of professional identity development within an online course/program;
- making visible the nature of epistemic engagement, the modes of thinking and acting that matter in certain social practices to help individuals learn how to participate meaningfully within those contexts;
- articulating in detail the functions and uses of the technological tools that most effectively support and mediate a community's social and intellectual engagement; and
- exercising caution when demonstrating how specific elements from the CoP theory can be applied to inform the design and execution of online/blended learning.

Overall, we hope that these findings will help current and future researchers to think more critically about their own use of the CoP framework when researching online/blended learning environments. We believe this theoretical framework continues to have potential for shedding light on how individuals learn within these social and situated contexts. Yet, at this point, it seems fair to argue that we have not fully exploited what this theory has to offer. We hope that this review will serve as a trigger for new lines of inquiry that take full advantage of this theory to expand our understandings of the process of online/blended learning in higher education and professional development.

References

Papers with three asterisks were eliminated during the analysis conducted for RQ1, i.e., "theoretical foundations" coding.

Papers with two asterisks were eliminated during the analysis conducted for RQ2, i.e., the "theoretical linkage" coding

Papers with a single asterisk remained at the end of two rounds of coding. Their analysis/findings illustrated the strongest connections to Wenger's theory. These are analyzed in the review conducted for RQ3 and are detailed in Appendix B.

Allen, I. E., & Seaman, J. (2013). *Changing Course: Ten Years of Tracking Online Education in the United States*. Needham MA: Babson College Research Group.

***Arnold, N., Ducate, L., Lomicka, L., & Lord, G. (2005). Using Computer-mediated Communication to Establish Social and Supportive. *CALICO Journal*, *22*(3), 537–566.

***Assaf, L. C. (2005). Exploring Identities in a Reading Specialization Program. *Journal of Literacy Research*, *37*(2), 201–236. http://doi.org/10.1207/s15548430jlr3702_4

***Asoodar, M., Atai, M. R., Vaezi, S., & Marandi, S. S. (2014). Examining effectiveness of communities of practice in online English for academic purposes (EAP) assessment in virtual classes. *Computers & Education*, *70*, 291–300. http://doi.org/10.1016/j.compedu.2013.08.016

*Adams, J. N. P. (2007). Artists becoming teachers: Expressions of identity transformation in a virtual forum goldsmiths research online artists becoming teachers: Expressions of identity transformation in a virtual forum. *International Journal of Art & Design Education*, *23*(6), 264–273.

**Amrein-Beardsley, A., Foulger, T. S., & Toth, M. (2007). Examining the development of a hybrid degree program: Using student and instructor data to inform decision-making. *Journal of Research on Technology in Education*, *39*(4), 331.

Amin, A. & Roberts, J. (2008). Knowing in action: Beyond communities of practice. *Research Policy*, *37*, 353–369.

**Annese, S., & Traetta, M. (2009). Psychosocial dynamics of blended communities: participating and positioning in learning contexts. *Open and Interdisciplinary Journal of Technology, Culture and Education*, *4*(2), 90–106. http://www.ckbg.org/qwerty/index.php/qwerty/article/viewArticle/59

Bates, J. S., & Taylor, A. C. (2013). Taking stock of theory in grandparent studies. In Mark A. Fine and Frank D. Fincham (Eds). *Handbook of family theories* (pp. 51-70). New York, NY: Taylor & Francis.

***Bell-Robertson, C. (2014). “Staying On Our Feet” Novice Music Teachers’ Sharing of Emotions and Experiences Within an Online Community. *Journal of Research in Music Education*, *61*(4), 431–451. Retrieved from http://jrm.sagepub.com/content/61/4/431.short

***Baxter, J. (2012). Who am I and What Keeps Me Going? Profiling the Distance Learning Student in Higher Education. *International Review of Research in Open and Distance Learning*, *13* (4), 107-129.

**Baxter, J., & Haycock, J. (2014). Roles and student identities in online large course forums: Implications for practice. *The International Review of Research in Open and Distributed Learning*, *15*(1), 21–40. Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/1593

**Beckett, G. H., Amaro‐ Jiménez, C., & Beckett, K. S. (2010). Students’ use of asynchronous discussions for academic discourse socialization. *Distance Education*, *31*(3), 315–335. doi: 10.1080/01587919.2010.513956

*Brosnan, K., & Burgess, R. C. (2003). Web based continuing professional development – A learning architecture approach. *Journal of Workplace Learning*, *15*(1), 24–33. doi: 10.1108/13665620310458794

**Cesareni, D., Martini, F., & Mancini, I. (2011). Building a community among teachers, researchers and university students: A blended approach to training. *International Journal of Computer-Supported Collaborative Learning*, *6*(4), 625–646. doi: 10.1007/s11412-011-9126-8

**Chang, C. K., Chen, G. D., & Li, L. Y. (2008). Constructing a community of practice to improve coursework activity. *Computers & Education*, *50*(1), 235–247. doi: 10.1016/j.compedu.2006.05.003

*Clarke, L. (2009). The POD model: Using communities of practice theory to conceptualise student teachers’ professional learning online. *Computers & Education*, *52*(3), 521–529. doi: 10.1016/j.compedu.2008.10.006

**Conrad, D. L. (2008). From community to community of practice: Exploring the connection of online learners to informal learning in the workplace. *American Journal of Distance Education*, *22*(1), 3–23. doi: 0.1080/08923640701713414

*Correia, A., & Davis, N. (2008). Intersecting communities of practice in distance education: the program team and the online course community. *Distance Education*, *29*(3), 289–306. doi: 10.1080/01587910802395813

Consalvo, A. L., Schallert, D. L., & Elias, E. M. (2015). An examination of the construct of legitimate peripheral participation as a theoretical framework in literacy research. *Educational Research Review*, *16*, 1–18. doi:10.1016/j.edurev.2015.07.001

Cousin, G., & Deepwell, F. (2005). Designs for network learning: A communities of practice perspective. *Studies in Higher Education*, *30*(1), 57–66. doi: 10.1080/0307507052000307795

**Cowan, J. (2012). Strategies for developing a community of practice: Nine years of lessons learned in a hybrid technology education master's program. *TechTrends*, *56*(1), 12–18. doi: 10.1007/s11528-011-0549-x

*Cowan, J. E., & Menchaca, M. P. (2014). Investigating value creation in a community of practice with social network analysis in a hybrid online graduate education program. *Distance Education*, *35*(1), 43–74. doi: 10.1080/01587919.2014.893813

*Crossard, B. and Pyror, J. (2008). Becoming researchers: A sociocultural perspective on assessment, learning and the construction of identity in a professional doctorate. *Pedagogy, Culture & Society*, *16(3)*, 221–237.

***Dahlberg, G. M. & Bagga-Gupta. (2013). Communication in the virtual classroom in higher education: Language beyond the boundaries of time and space. *Learning, Culture and Social Interaction 2*, 127–142. http://doi.org/10.1016/j.lcsi.2013.04.003

***Dawson, K., Cavanaugh, C., & Sessums, C. (2011). Designing a professional practice doctoral degree in Educational Technology: Signature pedagogies, implications and recommendations. *International Journal of …*, *25*(3), 1–15. Retrieved from http://www.ijede.ca/index.php/jde/article/view/767

***Delahunty, J., Verenikina, I., & Jones, P. (2014). Socio-emotional connections: identity, belonging and learning in online interactions. A literature review. *Technology, Pedagogy and Education*, *23*(2), 243-265.

**Egan, J. P. (2013). Learning community or community of practice: Preliminary findings of a transfer of learning study. *Issues in Information Science and Information Technology*, *10*, 137-144.

*Ellaway, R., Dewhurst, D., & McLeod, H. (2004). Evaluating a virtual learning environment in the context of its community of practice. *Alt-J*, *12*(2), 125–145. doi: 10.1080/0968776042000216192

*Evans, C., Yeung, E., Markoulakis, R., & Guilcher, S. (2014). An online community of practice to support evidence-based physiotherapy practice in manual therapy. *The Journal of Continuing Education in the Health Professions*, *34*(4), 215–23. doi:10.1002/chp.21253

Farnsworth, V., Kleanthous, I., & Wenger-Trayner, E. (2015). Communities of practice as a social theory of learning: A conversation with Etienne Wenger. *British Journal of Educational Studies*, p. 1-22. Advanced online publication. doi:10.1080/00071005.2015.1133799

Gee, J. P. (2005). Affinity spaces: From age of mythology to today's schools. In D. Barton, & K. Tusting, *Beyond communities of practice: Language power and social context* (pp. 214-232). New York: Cambridge University Press.

*Goggins, S. P., Laffey, J., & Gallagher, M. (2011). Completely online group formation and development: Small groups as socio‐ technical systems. *Information Technology & People*, *24*(2), 104–133. doi: 10.1108/09593841111137322

Grant, C., & Osanloo, A. (2014). Understanding, selecting, and integrating a theoretical framework in dissertation research: creating the blueprint for your "house." *Administrative Issues Journal: Connecting Education, Practice, and Research*, *14*(2), 12-26. doi: 10.1046/j.0266-4909.2003.00051.x

*Gray, B. (2004). Informal learning in an online community of practice. *Journal of Distance Education*, *19*(1), 20–35. Retrieved from http://eric.ed.gov/?id=EJ807836

*Guldberg, K., & Mackness, J. (2009). Foundations of communities of practice: Enablers and barriers to participation. *Journal of Computer Assisted Learning*, *25*(6), 528–538. doi: 10.1111/j.1365-2729.2009.00327.x

*Guldberg, K., & Pilkington, R. (2006). A community of practice approach to the development of non-traditional learners through networked learning. *Journal of Computer Assisted Learning*, *22*(3), 159–171. doi: 10.1111/j.1365-2729.2006.00171.x

**Henderson, M. (2007). Sustaining online teacher professional development through community design. *Campus-Wide Information Systems*, *24*(3), 162–173. doi: 10.1108/10650740710762202

Henri, F. & Pudelko, (2003). Understanding and analyzing activity and learning in virtual communities. *Journal of Computer Assisted Learning*, *19,* 474-487.

**Hodges, T. E., & Cady, J. (2012). Blended-format professional development and the emergence of communities of practice. *Mathematics Education Research Journal*, *25*(2), 299–316. doi: 10.1007/s13394-012-0065-0

Holsti, O. (1969). *Content analysis for social sciences and humanities*. Reading, MA: Addison-Wesley.

**Hough, B., Smithey, M., & Evertson, C. (2004). Using computer-mediated communication to create virtual communities of practice for intern teachers. *Journal of Technology and Teacher Education, 12*(3), 361–386.

Hughes, J., Jewson, N., & Unwin, L. (2007). *Communities of practice: Critical perspectives*. New York: Routledge.

**Kayler, M., & Weller, K. (2007). Pedagogy, self-assessment, and online discussion groups. *Journal of Educational Technology & Society, 10*(1), 136-147.

Kilbourn, B. (2006). The qualitative doctoral dissertation proposal. *Teachers College Record, 108* (4), 529-576.

***Kellogg, S., Booth, S. & Oliver, K. (2014). A Social Network Perspective on Peer Supported Learning in MOOCs for Educators. *International Review of Research in Open and Distance Learning, 15*(5), 263-289.

***Lee, K., & Brett, C. (2015). Dialogic understanding of teachers' online transformative learning: A qualitative case study of teacher discussions in a graduate-level online course. *Teaching and Teacher Education, 46*, 72–83. http://doi.org/10.1016/j.tate.2014.11.001

***Lee, J., Carter-Wells, J., & Glaeser, B. (2006). Facilitating the Development of a Learning Community in an Online Graduate Program. *Quarterly Review of Distance Education, 7*(949), 13–33. *Retrieved from http://eric.ed.gov/?id=EJ875021*

***Lou, Y. (2004). Learning to solve complex problems through between‐ group collaboration in project‐ based online courses. *Distance Education, 25*(1). Retrieved from http://www.tandfonline.com/doi/abs/10.1080/0158791042000212459

Larreamendy-Joerns, J., & Leinhardt, G. (2006). Going the distance with online education. *Review of Educational Research, 76*(4), 567–605

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. London: Cambridge University Press. doi:10.2307/2804509

**Lim, J. H., Dannels, S. A., & Watkins, R. (2008). Qualitative investigation of doctoral students' learning experiences in online research methods courses. *Quarterly Review of Distance Education, 9*(3), 223–236.

**Lin, T., Hsu, Y., & Cheng, Y. (2011). Emerging innovative teacher education from situated cognition in a web-based environment. *Online Journal of Educational Technology-TOJET, 10*(2), 100-112. Retrieved from http://eric.ed.gov/?id=EJ932230

*Mackey, J., & Evans, T. (2011). Interconnecting networks of practice for professional learning. *The International Review of Research in Open and Distributed Learning*, *12*(3), 1-18.

***Mackness, J., Waite, M., Roberts, G. & Lovegrove, E. (2013). Learning in a Small, Task–Oriented, Connectivist MOOC: Pedagogical Issues and Implications for Higher Education, International Review of Research in Open and Distance Learning, 14(4), 140-159.

**Marken, J. A., & Dickinson, G. K. (2013). Perceptions of community of practice development in online graduate education. *Journal of Education for Library and Information Science*, *54*(4), 299–307.

*Moule, P. (2006). E-learning for healthcare students: Developing the communities of practice framework. *Journal of Advanced Nursing*, *54*(3), 370–80. doi: 10.1111/j.1365-2648.2006.03813.x

**Murchu, D. O., & Sorensen, E. K. (2004). Online master communities of practice: Collaborative learning in an intercultural perspective. *European Journal of Open and Distance Learning, 2004(1).* Retrieved from http://www.eurodl.org/materials/contrib/2004/Identifying_COPs.pdf

*Nelson, G., & Temples, A. L. (2011). Identity construction as nexus of multimembership: Attempts at reconciliation through an online intercultural communication course. *Journal of Language, Identity & Education*, *10*(2), 63–82. doi: 10.1080/15348458.2011.563636

***Obura, T., Brant, W. E., Miller, F., & Parboosingh, I. J. (2011). Participating in a Community of Learners enhances resident perceptions of learning in an e-mentoring program: proof of concept. *BMC Medical Education*, *11*(1), 3. http://doi.org/10.1186/1472-6920-11-3

***Oliveira, I., Tinoca, L., & Pereira, a. (2011). Online group work patterns: How to promote a successful collaboration. *Computers & Education*, *57*(1), 1348–1357. http://doi.org/10.1016/j.compedu.2011.01.017

***Olofsson, A. D. (2007). Participation in an Educational Online Learning Community From Individuality to Community – Net-based Learning in Transition. *Educational Technology & Society*, *10*(4), 28–38.

***Owston, R., Wideman, H., Murphy, J., & Lupshenyuk, D. (2008). Blended teacher professional development: A synthesis of three program evaluations. *The Internet and Higher Education*, *11*(3-4), 201–210. http://doi.org/10.1016/j.iheduc.2008.07.003

**Razak, A. N., Saeed, M., & Ahmad, Z. (2013). Adopting social networking sites (SNSs) as interactive communities among English foreign language (EFL) learners in writing: Opportunities and challenges. *English Language Teaching*, *6*(11), 187–199. doi: 10.5539/elt.v6n11p187

*Rogers, J. (2000). Communities of practice: A framework for fostering coherence in virtual learning communities. *Educational Technology & Society*, *3*(3), 384-392.

**Rohleder, P., Bozalek, V., Carolissen, R., Leibowitz, B., & Swartz, L. (2007). Students' evaluations of the use of e-learning in a collaborative project between two South African universities. *Higher Education*, *56*(1), 95–107. doi: 10.1007/s10734-007-9091-3

Sandberg, J., & Alvesson, M. (2011). Ways of constructing research questions: Gap-spotting or problematization? *Organization*, *18*(1), 23-44.

**Schrum, L., Burbank, M. D., & Capps, R. (2007). Preparing future teachers for diverse schools in an online learning community: Perceptions and practice. *The Internet and Higher Education*, *10*(3), 204–211. doi: 10.1016/j.iheduc.2007.06.002

*Stacey, E., Smith, P. J., & Barty, K. (2004). Adult learners in the workplace: Online learning and communities of practice. *Distance Education*, *25*(1), 107–123. doi: 10.1080/0158791042000212486

**Stewart, A. R., Harlow, D. B., & DeBacco, K. (2011). Students' experience of synchronous learning in distributed environments. *Distance Education*, *32*(3), 357–381. doi: 10.1080/01587919.2011.610289

Szmigiel, R., & Lee, D. (2014). Avatars in E-and U-learning. In F. B. M. Neto (Ed.), *Technology platform innovations and forthcoming trends in ubiquitous learning* (pp. 35-58). Hershey, PA: Information Science Reference.

**Thang, S. M., Hall, C., Murugaiah, P., & Azman, H. (2011). Creating and maintaining online communities of practice in Malaysian Smart Schools: challenging realities. *Educational Action Research*, *19*(1), 87–105. doi: 10.1080/09650792.2011.547724

**Vaughan, N., & Lawrence, K. (2013). Investigating the role of mobile devices in a blended pre-service teacher education program, *Canadian Journal of Higher Education*, *43*(3), 56–78.

Vygotsky, L. S. (1978). *Mind in society.* Cambridge, MA: Harvard University Press.

***Walker, S., & Creanor, L. (2005). Crossing complex boundaries: transnational online education in European trade unions. *Journal of Computer Assisted Learning*, *21*(5), 343–354. http://doi.org/10.1111/j.1365-2729.2005.00141.x

Wenger E. (1998). *Communities of practice: Learning, meaning and identity*. New York, NY: Cambridge University Press.

Wenger, E. (2000). Communities of practice and social learning systems. *Organization*, *7*(2), 225–246.

Wenger, E., McDermott, R., & Snyder, W. (2002). *Cultivating communities of practice: A guide to managing knowledge*. Cambridge, MA: Harvard University Press.

Wenger, E. (2004). Knowledge management as a doughnut. *Ivey Business Journal*. Retrieved from: http://iveybusinessjournal.com/publication/knowledge-management-as-a-doughnut/

Wenger, E., White, N., Smith, J.D., Rowe, K. (2005). Technology for communities. In L. Langelier (Ed.), *Working, learning and collaborating in a network: Guide to the implementation and leadership of intentional communities of practice* (pp. 71–94). Quebec City: CEFIRO. doi:10.1109/ICSMC.2004.1399962

Wenger, E., White, N., & Smith, J. D. (2009). *Digital habitats: Stewarding technology for communities*. Portland, OR: CPSquare.

Wenger, E. (2010). Communities of practice and social learning systems: The career of a concept. In C. Blackmore (Ed.), *Social Learning Systems and Communities of Practice* (pp. 179–198). London: Springer. doi:10.1007/978-1-84996-133-2

Wenger, E., Trayner, B., & de Laat, M. (2011). *Promoting and assessing value creation in communities and networks: A conceptual framework*. Heerlen, The Netherlands: Ruud de Moor Centrum, Open University of the Netherlands.

***Wikeley, F., & Muschamp, Y. (2004). Pedagogical implications of working with doctoral students at a distance. *Distance Education*, *25*(1), 125–142. http://doi.org/10.1080/0158791042000212495

** Yukawa, J. (2010). Communities of practice for blended learning: Toward an integrated model for LIS education. *Journal of Education for Library and Information …*, *51*(2), 54. Retrieved from http://www.jstor.org/stable/20720485

Appendix A

Search Topics Used to Conduct Literature Search

Topic 1	Topic 2	Topic 3
• community of practice • communities of practice • CoP	Wenger	• online course(s) • blended course(s) • online learning • blended learning • distance learning • e learning • elearning • learning management system • LMS • virtual learning environment • VLE

Appendix B

Characteristics of Final Group of 17 CoP-grounded Research Studies

	Authors	Discipline	Research Approach	Online (O) Blended (B)	Course (C) Program (P)	Focus	Context	CoP Concept(s) Used
1	Adams (2007)	Education - Arts	Qualitative	B	C	Explores the experiences of art and design graduates as they transition to becoming K-12 art teachers using Wenger's concepts of identity and borders.	An online forum used in LMS to provide communications and support to student teachers during their 9-month school placement.	Identity, boundaries
2	Brosnan & Burgess (2003)	Professional Development - Health and human services	Qualitative	O	C	Explores how Wenger's notion of a learning architecture can be applied to evaluate and guide the design of an online professional development course	A 12-week long online professional development course for professionals from health, education, pharmacy and social work backgrounds	Learning architecture
3	Clarke (2009)	Teacher Education - Geography	Qualitative	B	B	Explores students' perspectives of the online components of a teacher education course designed using CoP concepts.	Cohorts of student teachers use LMS-designed environment and forums to support reflection, sharing of classroom resources and mutual support during 36- week long course.	Domain, community, practice, and brokering
4	Correia & Davis (2008)	Teacher Education	Qualitative	B	C, P	Examines the dynamics of two complementary	Program staff meet F2F and online over several	Community, legitimate

						and interacting communities: One involving program instructors and instructional developers; the other involving classroom and remote students in a blended program.	years; Students in semester long graduate course, where some students participate via teleconferencing.	peripheral participation
5	Cowan & Menchaca (2014)	Education - Technology	Mixed	B	P	Examines Master’s program to assess value created over its ten-year history.	An established hybrid graduate program designed and taught using CoP principles.	Value creation Framework
6	Crossard & Pryor (2008)	Education – Ed.D program	Qualitative	B	C	Examines how tutor and peer use of convergent and divergent assessment affected shift in identity.	Part-time doctoral students in a 16-week blended educational research methods course.	Legitimate peripheral participation, identity
7	Ellaway, Dewhurst, & McLeod (2004)	Heath Care - Medicine	Quantitative	NA	C	Describes the development and implementation of a post-hoc evaluation instrument based on Wenger’s learning architecture to assess the effectiveness of custom-designed virtual learning environment to support a CoP.	Administered to students, faculty and program staff in undergraduate medical school program	Learning architecture
8	Evans, Yeung, Markoulakis, & Guilcher (2014)	Professional Development Health Care - Physical Therapy	Qualitative	O	C	Examines how the LMS and CoP concepts were used to engage students in authentic learning in evidence-based research practices course.	Physical therapists in a ten-week online professional development course	Community

9	Goggins, Laffey, & Gallagher (2011)	Information Systems – Software development	Mixed	O	C	Explores how small groups cooperate to develop characteristics of CoP in a tool-based online course.	Graduate students in a semester long course in Designing Performance Support Systems.	Community
10	Gray (2004)	Professional Development - Human Services	Qualitative	O	C	Examines use of CoP theory to shape informal workplace learning in an online course.	A 7-week long professional development course for local Adult Learning Council coordinators.	Community, identity
11	Guldberg & Mackness (2009)	Professional Development - Organization Effectiveness	Mixed	O	C	Explores barriers and enablers to participation in an international online workshop designed as a CoP.	A 7-week long professional development course sponsored for academics and managers hosted by Wenger's consulting firm.	Community, identity
12	Guldberg & Pilkington (2006)	Professional Development - Human Services	Mixed	O	C	Examines how collaborative activities based on Wenger's concepts contributes to the development of a CoP	Year-long professional development course for parents and caregivers of people with Autism spectrum disorder.	Community, identity, boundaries
13	Mackey & Evans (2011)	Teacher education – educational technology	Mixed	NA	C	Explores how participants in formal course-based learning relate interconnecting experiences between practices, communities, and opportunities.	Wenger's theories used to examine how teachers in a graduate course use learning and resources from course-based CoP to their school-based CoPs.	Identity; multi-membership in multiple communities; boundaries; brokers
14	Moule (2006)	Heath care – Nursing, radiology, radiotherapy	Qualitative	O	C	Examines how the essential characteristics of CoPs develop in higher education online	Assessed whether students in their final year of study in a 6-week long online	Mutual engagement, Joint enterprise,

						learning environments.	interprofessional healthcare course were able to develop a CoP	Shared repertoire
15	Nelson & Temples (2011)	Linguistics & Foreign language teaching	Qualitative	O	C, P	Examined experiences of two exchange students using Wenger's CoP concepts of identity and reconciliation as students participated in online course, university and host country.	Online 15-week intercultural communications course served as bridge to help students negotiate their identities at host university and country.	Identity, reconciliation, multi-membership multiple communities
16	Rogers (2000)	Professional Development - TESOL	Qualitative	O	C	Used Wenger's concepts in exploratory study as framework for analyzing participant discourse to determine whether discourse was characterized by three essential elements of CoP	3-week long web-based professional development workshop for ESOL teachers and administrators	Community; mutual engagement, joint enterprise shared repertoire
17	Stacey, Smith & Barty (2004)	Education - Educational Technology	Mixed	O	C	Examined tensions that result when students move between course learning community and workplace CoPs.	22 Masters students enrolled in 13-week course on the theory and practice of e-learning.	Community, reconciliation, multi-membership

Institutional Factors for Supporting Electronic Learning Communities

Jayme N. Linton
Lenoir-Rhyne University

Abstract

This study was designed to explore how the electronic learning community (eLC) process at an established state virtual high school (SVHS) supported new and veteran online high school teachers through the communities of practice (CoP) framework. Specifically, this study focused on the institutionally-driven nature of the eLC process, using Wenger's CoP framework to analyze institutional factors that influenced the eLC process. Case study methods, including observation, interviews, and document analysis, were used to provide a rich and dynamic analysis of the eLC process in light of what research says about preparation and support for quality online teaching. While the institutionally-driven nature of the eLC process posed some barriers to alignment with the domain and community elements of the CoP framework, case study participants expressed that the eLC process impacted their practice and connected them to colleagues with which they could collaborate and problem solve. The use of strategies such as valuing the work of eLCs, removing barriers, and connecting the eLC process to the organizational strategy served to facilitate alignment with the CoP framework and overcome some of the potential disadvantages of an institutionally-driven eLC process.

Keywords: Electronic learning community, online learning community, K-12, virtual school, community of practice

Linton, J. N. (2017). Institutional factors for supporting electronic learning communities, *Online Learning 21*(1), 238-256. doi: 10.24059/olj.v21i1.953

Introduction

While much of what constitutes effective teaching in traditional classrooms also translates to good teaching online (Garrison & Anderson, 2003; Journell et al., 2013), an additional set of skills and competencies is needed to ensure high levels of student engagement and student learning in virtual settings (Learn NC, 2008; NEA, 2006; Palloff & Pratt, 2011; Redmond, 2011). These skills and qualities for teaching online courses, such as developing presence online and facilitating asynchronous discussions, are absent from teacher education programs (Barbour, Siko, Gross, & Waddell, 2013). In some cases, this leads to K-12 administrators touting online learning as unsuccessful when, in reality, the lack of training and support may be what is setting up many online instructors and online learners to fail (Learn NC, 2008). Forced to fend for themselves, many online instructors have adopted a “sink or swim” mentality, taking responsibility for their own professional learning (Hawkins, Graham, & Barbour, 2012; Marek, 2009; Ray, 2009).

This study was designed to explore how the electronic learning community (eLC) process at an established state virtual high school supported new and veteran online high school teachers through the communities of practice (CoP) framework. The institutionally-driven eLC process was designed to support online teachers through ongoing collaboration and professional development. While a research base on preparation and support for online instructors is growing, a theoretical framework is often missing in these studies. his study sought to apply the CoP framework to an institutionalized eLC process for supporting online instructors.

Literature Review

Communities of Practice

Essential to any CoP is a common domain, consisting of concepts and issues related to a body of knowledge that is necessary for members to develop their competencies and skills. The domain connects the work of the community to a broader community of practitioners. According to Wenger, McDermott, and Snyder (2002), “A shared domain creates a sense of accountability to a body of knowledge and therefore to the development of a practice” (p. 30). The second necessary element of a CoP, according to Wenger (1998), is a community of people who “interact, learn together, build relationships, and in the process develop a sense of belonging and mutual engagement” (Wenger et al., 2002, p. 33). Community within a CoP involves sustained, community-building engagement among members and a shared repertoire of strategies, tools, routines, and language as they engage in joint enterprise with other members. In addition to a community of members committed to a common domain, Wenger’s (1998) CoP framework includes the element of practice, which is defined as “doing in a historical and social context that gives structure and meaning to what we do” (Wenger, 1998, p. 46). In other words, practice is the “doing” of a CoP. Wenger’s (1998) concept of practice can be examined through the ways in which community members explore and produce ideas together.

The relationship between a CoP and the organization in which members work can range from unrecognized by the organization to highly institutionalized (Wenger et al., 2002). While the health of CoPs “depends primarily on the voluntary engagement of their members and on the emergence of internal leadership” (Wenger et al., 2002, p. 11), organizations can foster the growth and development of CoPs by valuing their work, creating time and space for CoP tasks, encouraging participation, and removing barriers (Wenger et al., 2002). In fact, successful CoPs which are likely to inspire growth, leadership, and innovation exist at the intersection of “strategic relevance” to the organization and the passions of community members (Wenger et al., 2002, p. 31).

Wenger et al. (2002) suggested that “without intentional cultivation, the communities that do develop will depend on the spare time of members” (p. 13). As CoPs address complex issues, solve problems, and contribute to improved practice, they offer value to the organization, measurable by both tangible results (i.e., improved skills and faster access to information) and intangible results (i.e., relationships, confidence, and a sense of belonging) (Wenger et al., 2002).

Perhaps of primary importance, CoPs bring value to the organization by "connecting the personal development and professional identities of practitioners to the strategy of the organization" (Wenger et al., 2002, p. 17). Whether a CoP is initiated by members or an organization, the ultimate success of the community will depend on the energy which members of the community generate (Wenger et al., 2002).

Communities of practice can exist along a continuum from unrecognized by the institution to highly institutionalized. The eLC process selected for this case study was institutionalized in that it was conceived and structured by state virtual high school (SVHS) leadership and mandated for teachers. While the success of a CoP is dependent upon the engagement of members (Wenger et al., 2002), organizations can support and increase the effectiveness of the work of CoPs in five ways: value the work of eLCs, create time and space, encourage participation, remove barriers, and connect to the organizational strategy. These five organizational strategies were examined during this case study to explore ways in which the eLC process supported new and veteran online teachers.

Electronic Learning Communities

An electronic learning community is defined as an online space to which members are committed and involved professionally over an extended period of time, with opportunities for synchronous and asynchronous communication (Duncan-Howell, 2010). Electronic learning communities create a third space for participants, where learners and experts are equals in the knowledge building process (So, Loss, Lim, & Jacobson, 2009). Participants use this third space to discuss common interests. According to an online survey of 98 members of three online communities for teachers, participants joined those communities to learn from their peers, keep up-to-date with current trends, engage in discussions, share professional knowledge, obtain support from colleagues, and build a safety net of like-minded educators. Approximately 87% of those survey respondents felt their online communities were meaningful. Seventy-seven percent reported that they made changes to their teaching practices as a result of their participation in an electronic learning community (Duncan-Howell, 2010).

By moving an existing professional learning community (PLC) into an online environment, K-12 teachers can extend their collaboration outside of their work day and transcend geographical boundaries (Tsai, Laffey, & Hanuscin, 2010). The Internet also can facilitate relationships within local communities of learners by providing them with a set of learning and collaboration tools that can be tailored to meet the needs of the community (Clary & Wandersee, 2009; Schlager & Fusco, 2003). Online environments that support existing school-based learning communities allow community members to take on leadership roles within different contexts (Schlager & Fusco, 2003). Results from a recent mixed-methods study revealed that teacher teams functioned well when using an online space to strengthen their existing learning communities (Parr & Ward, 2006). These teachers, first and foremost, felt safe within their existing learning community first, which contributed to the success of their online community. Parr and Ward (2006) also found that the existence of a well-functioning PLC within a school increased the likelihood that teachers would find success in an online learning community. Similarly, K-12 teachers involved in online professional development reported that their participation in an online discussion board reinforced the learning that had taken place among colleagues within the same school (Signer, 2008). Data gathered by Holmes, Signer, and

MacLeod (2010) showed that social presence was the greatest factor influencing teachers' learning and satisfaction online. When an existing PLC moved into an online environment, the social presence of the group contributed to the group's online learning.

Components of Electronic Learning Communities

Building an electronic learning community requires that community members combine technology and procedures that facilitate collaborative learning (Yeh, 2010b). Technology integration can support and motivate teachers to focus on continuous growth and school improvement (Williams, Atkinson, Cate, & O'Hair, 2008). Factors that influence the success of electronic learning communities include motivation to participate, a sense of group trust, cooperation, sociability, and usability. Similarity among group members also contributed to the belongingness felt by members of an electronic learning community (Yeh, 2010b). Rovai (2001) conducted a mixed-methods study of adult learners interested in distance learning. He described four dimensions that build a sense of community online: spirit, trust, interaction, and learning. Yeh (2010a) identified four types of electronic learning communities by analyzing discussion board messages: active collaboration, passive collaboration, individualized participation, and indifference. The active collaboration communities, which consisted of high levels of member participation and collaboration, performed best in assigned tasks online.

In a study of 32 pre-service teachers participating in a blended learning environment, composed of online and face-to-face learning experiences, Yeh (2010b) identified four stages for building an online learning community. Teachers moved through the stages of motivation and acquaintance, socialization and belongingness, information exchange and consensus, and tacit understanding and development. Electronic learning community members working at the highest stage of Yeh's model communicated well with one another and achieved goals effectively. A similar model developed by Waltonen-Moore, Stuart, Newton, Oswald, and Varonis (2006) included five stages of online group development: introduction, identification, interaction, involvement, and inquiry. The final stage, inquiry, occurred when teachers put what they learned into practice. Online communities who reached the inquiry stage behaved in ways similar to face to face conversations, with a lot of give and take among community members.

Benefits of Electronic Learning Communities

Taking advantage of electronic learning communities can provide numerous benefits to K-12 teachers. However, before teachers can benefit from an online learning community, they must first perceive a need and recognize that an online community can be a solution to address that need (Parr & Ward, 2006). Teachers involved in electronic learning communities have increased access to resources and flexibility with regard to the time and place in which they work (Lock, 2006). Learning in an online community has been described as immediate, relevant, authentic, and linked to real life as teachers directed their own discovery and construction of knowledge (Duncan-Howell, 2010). As opposed to expert-directed professional development, online communities can build teachers' capacity by giving them ownership of their own learning (Lock, 2006). For example, teachers participating in the online course studied by Holmes et al. (2010) reported that their online learning community provided them with a variety of instructional strategies. The access to resources afforded them by the electronic learning community impacted their teaching practice. In addition, online learning communities offered teachers a common language for communicating about teaching and learning (Chen, Chen, &

Tsai, 2009). Chen et al. (2009) further found that the use of technology as a tool to develop PLCs contributed directly to instructional practices. Their data also showed that technology made teacher collaboration faster and simpler.

Challenges of Electronic Learning Communities

While providing numerous benefits for K-12 teachers, online learning environments can pose several challenges as well. Duncan-Howell (2010) conducted an online survey of 98 teacher members of an electronic learning community. Participating teachers self-reported that time management and sidetracked conversations were barriers to effective learning online. Teachers involved in online professional development identified personal technological preferences, such as familiarity or comfort with specific types of software or web programs, as the basis for most problems within the online learning environment (Clary & Wandersee, 2009). Chen et al. (2009) also identified technical expertise as a factor in building a successful online learning experience for teachers. Similarly, Holmes et al. (2010) found that teachers with prior online learning experience were more satisfied with online professional development courses.

In her literature review of online teacher communities in K-12 education, Lock (2006) summarized the reasons why many electronic learning communities have failed, including problems with technology, lack of time, learner readiness, mismatch to the school culture, and quality of the community. The success of an electronic learning community is partly dependent on the technology available to facilitate teacher learning online. Technology tools used to support the online community should be flexible and meet the needs of community members. Lock (2006) also pointed out that online communities failed when teachers were not ready to participate. They must be self-motivated and independent learners and have a level of confidence with technology use. Effective online communities require teachers to transition from an isolated, autonomous working environment to one that is collaborative, but school culture can hinder the effectiveness of electronic learning communities. If a school's culture does not foster collaboration and collective learning, it can be difficult for teachers to break free from the traditional school culture of independence and autonomy. In addition, the electronic learning community should be integrated into teachers' professional development practices rather than being perceived as an add-on. "The power and direction of the community must come from community members. It cannot be imposed on them," (Lock, 2006, p. 673).

A large case study was conducted to explore several facets of the eLC process through the lens of the CoP framework. Specifically, the findings presented here focused on the institutionally-driven nature of the eLC process, using Wenger's (1998) CoP framework to analyze institutional factors that influenced the eLC process. Other aspects of Wenger's CoP framework are beyond the intended scope of this article. Case study methods, including observation, interviews, and document analysis, were used to provide a rich and dynamic analysis of the eLC process in light of what research says about preparation and support for quality online teaching. This case study was designed to examine data gathered through the eLC process at SVHS in order to address the following research question: In what ways do institutionally-driven electronic learning communities operate like communities of practice from the perspective of experienced online teachers, novice online teachers, and learning community leaders?

Methods

Setting and Sample

The second largest state virtual school in the country, referred to in this study as the State Virtual High School (SVHS), was commissioned in 2005 to provide e-learning opportunities to high school students from across a state in the southeastern United States. Courses for students were first offered in the summer of 2007. During its first year, 17,325 students enrolled in courses through SVHS. During the 2014-15 school year, SVHS enrolled 71,932 students in grades six through 12, with a total of 111,634 course enrollments (Watson et al., 2015). During the 2012-13 school year, SVHS contracted with nearly 700 online teachers who provided instruction to students from all 115 school districts in the state as well as 44 charter schools.

Virtual teachers at SVHS must first complete an 18-week induction program, which is designed to orient teachers to SVHS expectations as well as strategies and tools for effective online teaching. The induction program consists of a nine-week orientation to online teaching and a nine-week practicum which allows SVHS teaching candidates to apply online teaching skills by co-teaching online courses with veteran SVHS teachers. The first nine weeks orient candidates to online teaching competencies and to the expectations for online teaching at SVHS. Topics include instructional design models, expectations for communication with SVHS students and parents, web tools, copyright, and reporting student progress, among others. These expectations are presented synchronously and asynchronously by the orientation leader and veteran SVHS teachers. All candidates must attend synchronous sessions, participate in asynchronous discussion, and complete assignments each week. These synchronous and asynchronous experiences require SVHS teaching candidates to reflect on and apply course content.

Following the nine-week orientation, teaching candidates are matched with veteran SVHS teachers for a nine-week practicum experience. Each candidate takes on gradual responsibility, with mentorship and supervision by the veteran teacher, for online teaching in a course within his or her content area. Responsibilities taken on by SVHS candidates include maintaining regular communication with students and their parents, posting announcements, grading and providing feedback on student work, and reporting student progress to local school districts. Upon successful completion of the orientation, which is measured by weekly active participation in synchronous meetings and asynchronous discussions and multiple opportunities to apply and demonstrate online teaching skills through assignment submission, SVHS teachers become members of course-specific electronic learning communities. Within eLCs, teachers work collaboratively both synchronously and asynchronously to carry out their work as virtual teachers. This work includes course revision, sharing effective teaching practices, and goal-setting and reflection. During the virtual practicum, each candidate joins the eLC to which his or her mentor teacher belongs.

Electronic learning communities (eLCs) function as part of the overall continuous professional learning program for SVHS teachers. All SVHS teachers are contractually obligated to participate in the eLC process, which is designed to facilitate professional learning, collaboration, and growth among SVHS teachers. Other professional learning opportunities are offered to SVHS teachers on an as-needed basis, including a recent self-paced online training on a new learning management system. According to the SVHS chief academic officer, the purpose of eLCs is to provide a collaborative process for teachers to enhance their practice and improve student learning.

Electronic learning communities are organized by program of study and content area. Programs offered by teachers at SVHS include the following: traditional program of study, credit recovery, occupational course of study, and STEM (science, technology, engineering, and math). Within this structure, all SVHS teachers belong to a course-specific eLC. For example, all SVHS teachers who teach sections of Biology belong to the Biology eLC. Further, with traditional courses that have high enrollment, such as psychology, teachers of the general and honors sections function as separate eLCs. That is, all teachers of the general psychology course belong to the psychology eLC, while all teachers of the honors psychology course belong to the honors psychology eLC. This qualitative case study focused on the eLC process for online English teachers during the spring of 2014.

The total sample for this study included seven female SVHS employees, six of whom taught high school English, and the chief academic officer for SVHS. Two participants were new online teachers, having taught for SVHS for one year or less; two were classified as veteran online teachers, having taught for SVHS for three years or more; and two were eLC facilitators who were responsible for leading the eLC process for their respective courses. Institutional Review Board approval was obtained from the researcher's institution and SVHS.

Data Collection and Analysis

The institutionally-driven eLC process at SVHS was selected for this case study. Case study data were gathered over twelve weeks via interview, observation, and the collection of documents and shared websites. Three eLCs within the English department were selected by the SVHS research coordinator, who selects participants for and coordinates all external research, for observations and document analysis. Observations were conducted during seven monthly synchronous online eLC meetings that were hosted by each eLC facilitator within the learning management system. Shared documents, such as collaboratively constructed formative assessments, archives of announcements, and reflections to monthly eLC questions distributed by the chief academic officer, used by eLC participants for ongoing asynchronous eLC participation were collected and analyzed, along with eLCCommunication via email. Further, seven eLC participants, all female, were interviewed, including the chief academic officer, two instructional leaders, two veteran teachers, and two new teachers at SVHS. The following interview questions were used to guide one semi-structured interview with each of the seven participants.

1. Tell me about your experience as an online teacher (or instructional leader or chief academic officer) for SVHS and a member (or facilitator) of an eLC.
2. How is the eLC process structured at SVHS?
3. Have you participated in a face-to-face learning community? If so, how is participation in the eLC similar to or different from participation in a face-to-face learning community?
4. What are the areas of focus of your eLC?
5. What types of support are available to you through membership in the eLC?
6. What is expected of you as an eLC member (or facilitator)?

7. How do you participate in the eLC? In what ways are you involved?
8. What expectations do eLC members have of each other?
9. How do you communicate with other members of your eLC?
10. What kinds of relationships have you developed with other members? How have those relationships changed over time?
11. How would you say your own practice has changed, or not, as a result of being a member of an eLC?

Interview participants varied widely in their teaching experience, ranging from seven to 32 years of traditional face-to-face teaching and from one semester to eight years of online teaching. Pseudonyms are used to protect the identities of participants. See Table 1 for participant demographic information.

Table 1
Electronic Learning Community Member Demographics

	Mean	Standard Deviation	Range
Age in Years	42	8.50	27
Years of Face-to-Face Teaching Experience	16	9.18	27
Years of Online Teaching Experience	4.48	2.79	7.33
Total Years of Teaching Experience	17	7.81	24

While drawing the sample from a diverse group of eLCs from different disciplines was preferred, the research coordinator for SVHS selected the eLCs and participants for the study due to the amount of SVHS teacher participation in other research studies. With multiple studies occurring at SVHS during the spring 2014 semester, the SVHS research coordinator selected the six teacher participants and English eLCs involved in this case study. This poses a limitation to this study, since these participants may not be representative of the rest of the population of SVHS teachers and leaders in other disciplines or other eLCs. Further, while this case study provides a glimpse into the nature of the eLC process, twelve weeks is not enough time to truly determine alignment with the CoP framework. In order to better understand the institutional factors that supported the eLC process for online teachers, more time for data collection would be needed.

All data were analyzed using NVivo, a qualitative data analysis software to explore relationships and identify trends (Richards, 1999), through Wenger's (1998) CoP framework and were coded according to five organizational strategies identified by Wenger et al. (2002) as techniques for supporting communities of practice. Data analysis software was used to organize and code each piece of data collected. For example, phrases from interviews were tagged according to the five organizational strategies for supporting CoPs, as were responses to eLC reflection questions in shared documents, comments made during synchronous eLC meetings, and emails. Prior to data collection, the researcher generated possible examples of data to represent each code. This coding structure was used to analyze shared documents, field notes collected during synchronous observations, transcripts from interviews, and emails among eLC members.

Findings

Introduction

According to five of the seven interview participants, the eLC process has improved since its inception in 2010. Amy described that "over the years the eLCs have become much more valuable and meaningful (Amy, personal communication, February 26, 2014)." Similarly, Simone referred to the eLC process during an interview as "a gradual improvement. It gets better every year (Simone, personal communication, March 9, 2014)." Tina, a course lead, described how this improvement has occurred, saying, "They were always asking, 'Is this helpful? How can we make this better?' And they listened. And they still do that. Like I said, we're ever changing the courses, we're always trying to improve. They're very responsive (Tina, personal communication, February 26, 2014)."

Donna agreed, "We've gotten better at it year by year as we really look at what works and what doesn't work (Donna, personal communication, February 26, 2014)." She stated that during the spring of 2014, she was, "the happiest I've ever been, last semester and this semester, with what the eLCs look like." She went on to say, "I think we're getting there. I think we have a ways to go, but I think we are doing it the best this semester than we have ever done it, and that is simply because of trial-and-error (Donna, personal communication, February 26, 2014)."

Findings below, organized and coded according to strategies recommended by Wenger et al. (2002), revealed efforts made by SVHS to foster the growth and development of the eLC process.

Value the Work of eLCs

Repeatedly, in weekly reflections, emails, and synchronous meetings, eLC members were thanked for their work. During a live English I eLC meeting, Donna, the chief academic officer, expressed her gratitude, "You guys have been fantastic at how you have approached it in your eLCs. We cannot thank you enough for analyzing student work in such a proactive way (Donna, eLC meeting, March 27, 2014)." During another synchronous English meeting, Donna told teachers, "I appreciate what you guys do so much (Donna, department meeting, March 17, 2014)." This sentiment was also regularly expressed within instructions for weekly reflections, which often concluded with a statement such as, "Thank you for your work this week (Donna, eLC newsletter, March 2014)!" Donna expressed that it was important to her to ensure that teachers felt their time was honored. "I never want the teachers to feel like their time has been wasted (Donna, personal communication, February 26, 2014)." A veteran teacher and course lead, Tina expressed that she "always appreciated being treated like a professional by [SVHS]. I always feel like they appreciate me (Tina, personal communication, February 26, 2014)."

In addition to expressions of gratitude, SVHS demonstrated that it valued eLC work in other ways. For instance, synchronous meetings often included celebrations, both personal and professional, and recognition of effective practices used by SVHS teachers. Further, SVHS regularly gathered feedback from eLC members and used that feedback to make adjustments to the eLC process. During a synchronous meeting, all honors teachers were asked provide feedback on what was working well with the eLC process as well as suggestions for improvement. As previously shared by Tina, SVHS frequently asked, "'Is this helpful? How can we make this better?' And they listened. And they still do (Tina, personal communication, February 26, 2014)."

Create Time and Space

The eLC structure, with weekly asynchronous work via reflection questions and monthly synchronous meetings, provided consistent, focused time and space for eLC work to happen. While Wendy felt that she did not typically like that type of structure all time, she believed that "it works in this situation, because it keeps us on track and it keeps us talking about what we're supposed to be talking about (Wendy, personal communication, March 12, 2014)." Instructional leader Simone also appreciated the structure of the eLC process, which allowed teachers to review shared practice from previous months. According to Simone, compared to face-to-face learning community work, the eLC structure was "more accessible. It's more permanent (Simone, personal communication, March 9, 2014)."

In addition to the structure of time and space for ongoing eLC work, SVHS worked to provide important content in an easily digestible format. Specifically, the honors portfolio process was broken down into small chunks of information and specific steps for eLCs to complete. In an email to the English I eLC, Amy explained that Donna had "broken it down really carefully for us and we are to move through it one week at a time (Amy, eLC meeting, February 24, 2014)." Later, during a synchronous meeting, she stated that Donna had "laid out a document to walk teams through the portfolio week-by-week (Amy, eLC meeting, February 24, 2014)." Also, Advanced Placement® (AP®) data were presented in an organized, condensed format for AP® teachers to use during their eLC work. The presentation of weekly and monthly information related to the eLC focus in a consistent and structured format provided evidence that SVHS created time and space for ongoing eLC work.

Encourage Participation

Since regular and active participation in the eLC process was an expectation of all SVHS teachers enforced through the teacher evaluation process, participation was not voluntary. However, the levels at which members participated and contributed to both the eLC process and the organization varied. Further, eLC members were encouraged to participate at different levels of engagement. In interviews, course leads and instructional leaders described being approached by SVHS leaders and asked to assume leadership roles within the eLC process. These leaders were then provided professional development and support as they moved into leadership roles. The selection and preparation of eLC members to take on core leadership roles within the eLC process encouraged different levels of participation. In addition to course lead and instructional leader roles, eLC members were also provided additional opportunities to participate via mini-contracts for course revision. Donna explained that mini-contracts would be used to complete course revisions at the end of the honors portfolio process. Teachers expressed appreciation that mini-contracts would be used for course revision to honor teachers' time in the eLC process.

According to Wenger et al. (2002), maintaining small community sizes is one technique for encouraging all members to participate actively. Each eLC participating in this case study had either two, three, or four community members. Within these small communities, each member was encouraged and expected to contribute to synchronous and asynchronous eLC work. Other data coded as "encourage participation" included reminders and prompts to participate from eLC leaders. For instance, Amy emailed level I English teachers, "If you haven't completed week 1 yet please go ahead and do so - it was due yesterday (Amy, email, March 3, 2014)." Amy described her efforts to encourage participation within synchronous meetings, particularly with one eLC that was a "really quiet group" as "pulling teeth (Amy, personal communication, February 26, 2014)." On the contrary, the other eLC in which Amy participated was very talkative and participatory. Donna also encouraged participation during synchronous meetings, as evidenced during a previously described synchronous meeting when she asked participants to be active in the conversation and avoid multi-tasking. She communicated to meeting participants that she expected them to be active in the chat.

Remove Barriers

Within the SVHS eLC process, some data revealed that distance presented a barrier to community development, while other data showed that distance was not a barrier. According to instructional leader Amy, some eLCs used online tools such as Skype and Google Hangouts to foster community development. In this way, eLC members were able to see and hear each other and, in addition, they got to know each other's families. Other communities were described by Amy as being more "business-like (Amy, personal communication, February 26, 2014)." New teacher Cheryl described that she was able to meet some of her fellow eLC members, who lived close to her, face-to-face for breakfast. She added that she was not able to meet other eLC members due to the distance between them.

Some evidence of distance posing a barrier within the eLC process was found in asynchronous work. Although eLCs participated in monthly synchronous meetings, the majority of eLC work occurred asynchronously. Members were required to post responses to weekly reflection questions then return later to the document to read one another's responses. Instructions for eLC work typically included a statement such as this one, "Wait a few days and come back to this document and read through your team's responses (Weekly reflection, January 27, 2014)." In a face-to-face community, this discussion would take place synchronously, with community members engaging in natural dialogue, commenting on one another's thoughts, and contributing in real-time. The delayed response via asynchronous discussions added wait time to eLC conversations.

Contrasting evidence was found, supporting the notion that distance did not present a barrier to participation in the eLC process. In fact, eLC members felt that the electronic nature of the eLC process facilitated participation. Tina described that "sharing is so easy on the computer with Google Docs to actually copy / paste precisely what you did (Tina, personal communication, February 26, 2014)." SVHS worked to remove barriers created by distance through the provision of frequent opportunities for interaction among eLC members. In contrast to Wenger et al. (2002), eLC members were not likely to forget that the community existed, as participation was mandatory and both structures and resources were provided to facilitate active participation among all members. Participation within the joint enterprise of the eLC was an ongoing part of work as an online teacher for SVHS, and communication within the eLC process was frequent, occurring via email, phone, synchronous meetings, and shared documents.

Specific strategies were employed by SVHS to remove potential barriers to participation in the eLC process. For example, all synchronous meetings were archived and posted to the shared community space. This allowed all members to participate in synchronous meetings even when scheduling conflicts occurred. Also, as described by instructional leader Amy, weekly asynchronous work was broken down and organized in a way that made it manageable and easy to follow. The English III eLC leader emailed eLC members, "Because this is so nicely broken down for us, I think it will be manageable, especially working together (Amy, email, March 3, 2014)."

Connect to the Organizational Strategy

One way in which SVHS connected the eLC process to the organizational strategy was by determining the focus of monthly eLC work. Donna described that the eLC focus was driven by monthly results from the eLC process, with the goal of designing eLC experiences that "can be applied right now to improve student learning (Donna, personal communication, February 26, 2014)." The SVHS leadership team used the eLC process to engage teachers in work around critical issues related to the organization, which during the spring of 2014 consisted primarily of focusing on the honors portfolio process and AP® data.

For the first time in the history of the eLC process, AP® eLCs were provided with a different topic for their work than non- AP® eLCs during the spring of 2014. This work focused on AP® data and the culture of AP® courses at SVHS. Simone, the instructional leader for AP® English Language, explained that in her time at SVHS, she had "never known AP® to have its own PD, so this has been really nice (Simone, personal communication, March 9, 2014)." In an interview, Donna explained that SVHS felt that "the culture of AP® is not where it needs to be. It's not reflective of the rest of [SVHS] (Donna, personal communication, February 26, 2014)." The differentiated focus for honors, AP®, and other eLCs during the spring of 2014 provided evidence of the eLC process being used to target specific components of the SVHS organizational strategy.

All SVHS teachers were required to actively participate in the eLC process, which was included in the teacher evaluation structure at SVHS and part of the SVHS teacher contract. Connecting eLC participation to teacher evaluation ensured that SVHS could extend its organizational priorities and expectations to every SVHS teacher in a systematic way. As described by Donna, SVHS "incorporated it that fall into the teacher contract so they would know we were serious about it, that we expected eLCs to be a requirement weekly (Donna, personal communication, February 26, 2014)." She later added, "I think that the eLC process puts feet to our expectations, and I think it also shows the reinforcement and the support we're going to provide (Donna, personal communication, February 26, 2014)," further revealing the strong connection between the eLC process and the SVHS organizational strategy.

Discussion

To answer this study's research question, data were gathered related to five strategies identified by Wenger et al. (2002) as ways for an organization to support and increase the effectiveness of the work of CoPs: value the work of eLCs, create time and space, encourage participation, remove barriers, and connect to the organizational strategy. Findings revealed that these strategies did, in fact, support and increase the effectiveness of the eLC process in many ways. However, findings also revealed that the institutionally-driven nature of the eLC process also posed a barrier to the alignment of the eLC process with the CoP framework.

Wenger et al. (2002) emphasized the notion of reciprocity within the CoP framework. Individual CoP members should benefit from membership in the community, while the community also benefits from the contributions of individual members. On a larger scale, the organization should benefit from CoPs, while communities should benefit from the process as well. In this case study, the eLC process brought value to all three levels – individual, community, and organization. As described previously, teachers expressed ways their teaching improved due to participation in the eLC process. Case study participants also described finding value in relationships built through the eLC process. Just as teachers benefited from participation in the eLC process, the communities themselves were enhanced through the eLC process. The contributions of each eLC member led to the improvement of courses, and ongoing interactions among community members strengthened the community and practice of the eLCs. Further, the eLC process brought value to the organization.

Wenger et al. (2002) described that successful CoPs existed at the intersection of "strategic relevance" to the organization and the passions of CoP members (p. 31). Truman (2004) described ways in which an institutionalized support system for online faculty can be successful for instructors and the organization. In the case of the honors portfolio process, in particular, the eLC process did exist at the intersection of the needs of SVHS and the ongoing interest of eLC members in improving their practice and their courses to meet the needs of students. At a broader level, the intersection of organizational relevance and eLC members' passions existed in a focus on students. Chief academic officer Donna described the purpose of the eLC process as "our ability to be collaborative among teachers to improve student learning (Donna, personal communication, February 26, 2014)." Case study participants, including teachers and eLC leaders, also described student learning as a focus of the eLC process. Findings revealed that students were the major focus of the eLC process, with the word student(s) being the most frequently used word across all data sources as revealed by frequency counts conducted in NVivo. The eLC process, then, provided a way for teachers to collaborate and improve their teaching and their courses in the interest of increasing student learning. Organizational competence, particularly with teaching practice and student learning, was increased through the eLC process.

One way in which the institutionally-driven nature of the eLC process posed a barrier to alignment with the CoP process was related to mandatory participation. Wenger et al. (2002) argued that participation within a CoP could be mandatory, but the level of participation must be voluntary. Further, Parr and Ward (2006) found that teachers must first perceive a need, and then recognize that an electronic learning community can be a solution to that need. In the case of the SVHS eLC process, curriculum leaders at SVHS perceived a need and believed that the eLC process would meet that need. Chief academic officer Donna described in an interview that curriculum and instruction leaders were discussing ways to get teachers to collaborate with one another to address low pass rates in their courses. According to Donna, "we decided, well we know that PLCs have a great purpose behind them, so let's do an eLC (Donna, personal communication, February 26, 2014)." She further explained, "So that was how the eLCs came to be, out of a desperate need to have teachers collaborate and talk with each other" (Donna, personal communication, February 26, 2014). While the leadership of SVHS recognized this need and believed strongly that eLCs were the solution, it is difficult to say whether teachers themselves perceived this need and believed in the eLC process as a way to meet that need. Previous research has shown that online instructors desire community and opportunities for collaboration with other online instructors, suggesting that SVHS teachers may have recognized the same need (Terosky & Heasley, 2015).

Lock (2006) found that eLCs can build capacity by giving teachers ownership of their learning. Lock also argued that the power and direction of eLCs comes from members. The institutionally-driven nature of the eLC process leads to questions regarding who owns and directs the learning. Aside from the month when eLCs were allowed to choose their own focus, the topics and issues addressed through the eLC process were selected by the curriculum and instruction leadership team. Although chief academic officer Donna explained that the leadership team selected topics based on the current needs of SVHS teachers and students, the argument can be made that the institutionally-driven nature of the eLC process prevented teachers from owning and directing their learning. What follows are recommendations for organizations and researchers based on findings from this case study.

Recommendations for Organizations Implementing eLCs

Provide opportunities for connecting the eLC process with professional organizations and professional learning opportunities outside of the organization. This could include providing funding for attendance at regional and national conferences, purchasing subscriptions to print or online publications, and providing access to online resources made available via professional organizations.

Implement a mentor program to support new online teachers. The mentor program can serve as a supplement to the eLC process by matching new online teachers with veteran online teachers. Mentees should be given regular opportunities to interact with their mentors, synchronously and asynchronously. The mentor program can provide new online teachers a safe space to ask questions, seek information, and gain confidence in their own practice.

Build a culture of celebration. To overcome barriers due to separations in distance and time, community-building must be an intentional component of the eLC process. Professional learning opportunities could be provided to help eLC facilitators develop skills and processes for community-building within the eLC process. Further, during organization-wide synchronous and asynchronous interactions, organization leaders can model community-building efforts.

Use a framework to guide and evaluate the eLC process. Organization leaders can collaborate to select and adapt a framework, such as the communities of practice framework, in order to ensure a consistent and systematic approach to the eLC process across the organization. Each organization should adapt the selected framework to meet the specific professional learning needs of its teachers. The revised framework can then be used to monitor and evaluate eLC implementation.

Recommendations for Researchers

Findings and limitations from this case study led to several recommendations for researchers interested in exploring electronic learning communities. These recommendations include: 1) ground research in a theoretical framework; 2) determine effective teaching practices in online environments; 3) determine effective practices for preparing and supporting online instructors; 4) use design research to improve educational practice; and 5) explore the impact of eLC participation on online teaching practice. Each of these recommendations is described in more detail below.

First, the use of the CoP framework served to ground this study in research and provided a useful structure for organizing and analyzing data. Researchers are encouraged to use a theoretical framework as the foundation of future studies exploring eLCs, utilizing the framework to design research questions, structure the data collection process, create processes for data analysis, and interpret findings.

Furthermore, in order to support K-12 online instructors in using best practices and supporting student learning, researchers need to determine effective practices for training and supporting K-12 online instructors. Research into effective online instructional practices is expanding at a rapid rate, but research into effective models for K-12 online instructor preparation and support is still greatly needed. Teacher education programs are failing to prepare pre-service teachers for their potential future work as online teachers (Barbour et al., 2013; Journell et al., 2013; NEA, 2006).

Perhaps more important than the content of future research is the design, implementation, and reporting of this research (Barbour, 2010). According to a review of the research on K-12 online learning conducted by Barbour (2010), most current literature in the field of K-12 online education has been based on personal experiences rather than systematic research. One of the leading researchers in the field, Michael Barbour (2010) recommended that researchers use a design research approach to conduct research in K-12 online education settings. The purpose of design research is to improve educational practice through a cycle of analysis, design, development, and implementation conducted collaboratively among researchers and practitioners in authentic settings. In contrast to traditional research, the goal of design research is not to generalize findings to other settings but to collaborate with members of the research site to solve their problems (Barbour, 2010). Design research is a methodology that is systematic yet flexible enough to be practical for dynamic K-12 online environments.

Finally, one important issue that was not explored in this case study was the impact of online teachers' participation in the eLC process on their teaching. Future research could explore the impact of eLC participation on online teaching practice, providing insight for eLC leaders and teachers. A study such as this would require the researchers to dig deeper than the eLC process, gaining access to courses and students. To best explore the impact of eLC participation on online teachers' practice, a longitudinal study would be useful. Research could observe eLC participation and online teaching practice over time, exploring the relationship between the two.

Conclusion

Case study participants described previous professional development experiences, including face-to-face learning communities, as segmented, unproductive, and disconnected from students. On the contrary, the eLC process was described by these teachers as authentic, genuine, accessible, and student-centered. For these teachers, the eLC process served as ongoing, productive professional learning that was focused on improving teaching to increase student learning. While the institutionally-driven nature of the eLC process posed some barriers to alignment with the domain and community elements of the CoP framework, case study participants expressed that the eLC process impacted their practice and connected them to colleagues with which they could collaborate and problem solve. Case study participants repeatedly mentioned that improving teaching practices and increasing student learning outcomes were desired results of participating in the eLC process. Requiring mandatory participation in the eLC process served to support SVHS teachers and students while also supporting the overall organization. Had SVHS not mandated participation in the eLC process and provided a consistent structure for ongoing eLC work, it could be argued that the beliefs and practices of these online teachers, particularly the new online teachers, would have been very different.

Further, the institutionally-driven nature also allowed the eLC process to concurrently offer value to the organization itself. The argument can be made that the supports provided the eLC process by SVHS, which were recommended by Wenger et al. (2002), allowed the institutionally-driven nature of the eLC process to act in more positive than negative ways in regards to the CoP framework. That is, the use of strategies such as valuing the work of eLCs, removing barriers, and connecting the eLC process to the organizational strategy served to facilitate alignment with the CoP framework and overcome some of the potential disadvantages of an institutionally-driven eLC process. Therefore, it is recommended that organizations follow guidelines from Wenger et al. (2002) to provide support for eLCs, including valuing the work of eLCs, creating time and space, removing barriers, encouraging participation, and connecting the eLCs to the organization's strategy.

Historically, SVHS curriculum and instruction leaders selected topics for the monthly eLC focus. However, during the spring of 2014, eLCs were allowed to select their own topics for one month of eLC work. This proved to be a decision that tightened the alignment of the eLC process with the CoP framework by allowing eLCs to direct and own their learning. Findings revealed that when eLCs were allowed to choose their own topics, each eLC engaged in work that was aligned to the overall focus of the eLC process and the organizational strategy. If SVHS were to continue to provide regular opportunities for eLCs to direct their own work, it would follow that this pattern would continue, maintaining connections between the eLC process and the organizational strategy while allowing each eLC to be a true CoP, in which the work of each community exists at the intersection of the goals of the organization and the passions of its members.

References

Barbour, M. K. (2010). Researching K-12 online learning: What do we know and what should we examine? *Distance Learning, 7*(2), 7-12.

Barbour, M. K., Siko, J., Gross, E., & Waddell, K. (2013). Virtually unprepared: Examining the preparation of K-12 online teachers. In R. Hartshorne, T. L. Heafner, & T. M. Petty, (Eds.), *Teacher education programs and online learning tools: Innovations in teacher preparation* (pp. 120-143). Hershey, PA: IGI Global.

Chen, Y., Chen, N.-S., & Tsai, C.-C. (2009). The use of online synchronous discussion for web-based professional development for teachers. *Computers & Education, 53*(4), 1155-1166.

Clary, R. M., & Wandersee, J. H. (2009). Can teachers learn in an online environment? *KappaDelta Pi Record, 46*(1), 34-38.

Duncan-Howell, J. (2010). Teachers making connections: Online communities as a source of professional learning. *British Journal of Educational Technology, 41*(2), 324-340.

Garrison, D. R., & Anderson, T. (2003). *E–Learning in the 21st century: A framework for research and practice*. London: Routledge/Falmer.

Hawkins, A.; Graham, C. R., & Barbour, M. K. (2012). Everybody is their own island: Teacher disconnection in a virtual school. *The International Review of Research in Open and Distance Learning, 13*(2), 123-144.

Holmes, A., Signer, B., & MacLeod, A. (2010). Professional development at a distance: A mixed-methods study exploring inservice teachers' views on presence online. *Journal of Digital Learning in Teacher Education, 27*(2), 76-85.

Journell, W., Beeson, M. W., Crave, J. J., Gomez, M., Linton, J. N., & Taylor, M. O. (2013). Training teachers for virtual classrooms: A description of an experimental course in online pedagogy. In R. Hartshorne, T. L. Heafner, & T. M. Petty, (Eds.), *Teacher education programs and online learning tools: Innovations in teacher preparation* (pp. 120-143). Hershey, PA: IGI Global.

Learn NC. (2008). Training teachers to teach online: The importance of certification for online instructors [White paper]. Retrieved May 15, 2013, from: http://www.learnnc.org/lp/media/uploads/2008/11/colt_white_paper.pdf

Lock, J. (2006). A new image: Online communities to facilitate teacher professional development. *Journal of Technology and Teacher Education, 14*(4), 663-678.

Marek, K. (2009). Learning to teach online: Creating a culture of support for faculty. *Journal of Education for Library and Information Science, 50*(4), 275-292.

National Education Association (NEA). (2006). *Guide to teaching online courses.* Retrieved from http://www.nea.org/assets/docs/onlineteachguide.pdf

Palloff, R. M., & Pratt, K. (2011). *The excellent online instructor: Strategies for professional development*. San Francisco: Jossey-Bass.

Parr, J., & Ward, L. (2006). Building on foundations: Creating an online community. *Journal of Technology and Teacher Education, 14*(4), 775-793.

Ray, J. (2009). Faculty perspective: Training and course development for the online classroom. *Journal of Online Learning and Teaching, 5*(2), 263-276.

Redmond, P. (2011). From face-to-face teaching to online teaching: Pedagogical transitions. Paper presented at Ascilite 2011, Hobart, Tasmania, Australia. Retrieved from http://www.ascilite.org.au/conferences/hobart11/downloads/papers/Redmond-full.pdf

Richards, L. (1999). *Using NVivo in qualitative research*. Thousand Oaks: Sage.

Rovai, A. (2001). Building community at a distance: A case study. *Educational Technology Research and Development, 49*(4), 33-48.

Schlager, M. S., & Fusco, J. (2003). Teacher professional development, technology, and communities of practice: Are we putting the cart before the horse? *The Information Society*, *19*(3), 203-220.

Signer, B. (2008). Online professional development: Combining best practices from teacher, technology and distance education. *Journal of In-service Education, 34*(2), 205-218.

So, H.-J., Lossman, H., Lim, W.-Y., & Jacobson, M. J. (2009). Designing an online video based platform for teacher learning in Singapore. *Australasian Journal of Educational Technology,25*(3), 440-457.

Terosky, A., & Heasley, C. (2015). Supporting online faculty through a sense of community and collegiality. *Online Learning Journal, 19*(3), 147-161.

Truman, B. E. (2004). UCF's exemplary faculty support: An institutionalized ecosystem. *Journal of Asynchronous Learning, 8*(3), 89-96.

Tsai, I.-C., Laffey, J. M., & Hanuscin, D. (2010). Effectiveness of an online community of practice for learning to teach elementary science. *Journal of Educational Computing Research, 43*(2), 225-258.

Waltonen-Moore, S., Stuart, D., Newton, E., Oswald, R., & Varonis, E. (2006). From virtual strangers to a cohesive online learning community: The evolution of online group development in a professional development course. *Journal of Technology and Teacher Education*, *14*(2), 287-311.

Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge: Cambridge University Press.

Wenger, E., McDermott, R., & Snyder, W. M. (2002). *Cultivating communities of practice: A guide to managing knowledge.* Boston, MA: Harvard Business School Publishing.

Williams, L. A., Atkinson, L. C., Cate, J. M., & O'Hair, M. J. (2008). Mutual support between learning community development and technology integration: Impact on school practices and student achievement. *Theory into Practice*, *47*(4), 294-302.

Yeh, Y.-C. (2010a). Analyzing online behaviors, roles, & learning communities via online discussions. *Educational Technology and Society*, *13*(1), 140-151.

Yeh, Y.-C. (2010b). Integrating collaborative PBL with blended learning to explore preservice teachers' development of online learning communities. *Teaching and Teacher Education, 26*(8), 1630-1640.

Adapting for Scalability: Automating the Video Assessment of Instructional Learning

Amy M. Roberts
University of Nebraska

Jennifer LoCasale-Crouch, Bridget K. Hamre, Jordan M. Buckrop
University of Virginia

Abstract

Although scalable programs, such as online courses, have the potential to reach broad audiences, they may pose challenges to evaluating learners' knowledge and skills. Automated scoring offers a possible solution. In the current paper, we describe the process of creating and testing an automated means of scoring a validated measure of teachers' observational skills, known as the Video Assessment of Instructional Learning (VAIL). Findings show that automated VAIL scores were consistently correlated with scores assigned by the hand scoring system. In addition, the automated VAIL replicated intervention effects found in the hand scoring system. The automated scoring technique appears to offer an efficient and reliable assessment. This study may offer additional insight into how to utilize similar techniques in other large-scale programs and interventions.

Keywords: automated assessment, scalability, teacher education

Roberts, A. M., LoCasale-Crouch, J., Hamre, B. K., & Buckrop, J. M. (2017). Adapting for scalability: Automating the video assessment of instructional learning, *Online Learning 21*(1), 257-272. doi: 10.24059/olj.v21i1.961

Introduction

Implementing large-scale evidence-based programs offers a promising means of reaching broad audiences (Franks & Schroder, 2013). Massively Open Online Courses demonstrate one-way educational content is being disseminated widely (Vale & Littlejohn, 2014). The use of online courses has increased among various groups of professionals, including teachers, to advance learners' technical knowledge and skills (Gill, 2011; U.S. Department of Education, National Center for Education Statistics, 2016). Although recent shifts towards large-scale programs are promising, they pose challenges to assessment, particularly the assessment of learners' skills. Skill assessments are an integral part of most educational programs and are often used to understand individuals' growth trajectories (Biggs & Tang, 2011). Automated coding systems offer a possible means of assessing learners' skills in larger scale programs (Williamson, Xi, & Bryer, 2012).

Given the increased offering of online professional development to teachers (Gill, 2011; Means, Toyama, Murphy, Bakia, & Jones, 2009), the present study sought to adapt a validated measure of teachers' observational skills of effective teacher-child interactions, the Video Assessment of Instructional Learning, or VAIL (Jamil, Sabol, Hamre, & Pianta, 2015), to be automatically scored, rather than manually hand scored, and thus applicable for large-scale interventions. More specifically, the goals of the study were to determine whether the automated VAIL scoring system related to the previously validated hand scoring system and if it was sensitive to intervention effects in a previous professional development program. To achieve these aims, we first correlated automated scores with hand scores and then compared intervention and control group means to determine if the automated VAIL replicated results previously found using the hand scored VAIL. These results are presented along with a discussion of how automated measures may be useful in other large-scale programs.

Literature Review

As the availability of online coursework grows, course designers and instructors are faced with the challenge of determining how to accurately and efficiently assess students (Palloff & Pratt, 2008). Assessments provide valuable information regarding motivation and progress, and can be used to provide feedback to both learners and instructors. Assessing learners' knowledge and skills from the beginning to the end of a course can also determine the efficacy of a program and may suggest modifications that need to be made (Boston, 2002). In general, assessment questions may be open-ended (short-answer or essays) or closed-ended (true/false or multiple choice.) Open-ended questions provide more thorough information regarding learners' mastery because they require learners to generate responses, rather than simply identifying correct answers from a prescribed list of options (Foddy, 1993). Although, open-ended questions may provide more useful information, they are more arduous to score which may be difficult in courses with large numbers of enrollees (Landauer, Laham, & Foltz, 2003).

Assessing open-ended items typically relies on the knowledge and expertise of human raters, such as instructors or teaching assistants, to manually score and make personal judgments for each response. This technique is not always feasible in large-scale programs because it is time consuming, and thus, costly. Interventions that require humans to score large quantities of responses may take inordinate amounts of time and deplete resources (Landauer et al., 2003; Williamson et al., 2012). Additionally, placing such potentially burdensome demands on human raters may increase the likelihood of fatigue and error (Ramineni & Williamson, 2013). In response to the practical limitations associated with manual hand scoring, automated scoring techniques may offer a possible solution.

Automated Scoring Techniques

Automated scoring systems, in which responses are scored by machines, have been used to assess short-answer responses, essays, and spoken responses. For example, the Educational Testing Service (ETS) has utilized automated essay scoring (AES) for high-stakes assessments, such as the GMAT or GRE, for over 10 years (Williamson et al., 2012). Various AES systems exist, which all require large numbers of essay samples to base their scoring and feedback on. These systems tend to provide both holistic and specific feedback, although the exact content varies by system. Overall, AES systems have been found to be valid and reliable (Dikli, 2006).

Automated scoring techniques are not without controversy, however. In particular, AES has been criticized for oversimplifying the assessment of writing to focus on rote elements, such as word count or complexity of word choice, rather than less easily quantifiable aspects, such as thoughtfulness of response or writing to a specific audience (Condon, 2013; Perelman, 2014). The application of automated scoring techniques for short answer responses is less highly contested, especially when such techniques are not used to holistically rate the quality of writing with high-stakes implications. However, it has been suggested that automated short answer scoring is more arduous to create than AES, because AES tends to focus on grammar and mechanics while automated short answer scoring tends to focus on content (Brew & Leacock, 2013). Brew and Leacock (2013) further suggested that automated short answer scoring systems are underutilized because "it is currently impossible to buy an off-the-shelf short answer scoring engine that will work for all items" (p. 151).

Despite the aforementioned challenges, automated short answer systems are best suited to measure explicit concepts, facts, or skills (Brew & Leacock, 2013). Perhaps the most well-known automated system for short answer responses is known as "C-rater" which was developed by ETS (Leacock & Chodorow, 2003). The validity of C-rater has been evaluated by comparing automated scores with hand scores. Leacock and Chodorow (2003) found that automated scores matched scores assigned by human raters 84% of the time. Similar findings have been replicated elsewhere (Burstein, Chodorow, & Leacock, 2003), suggesting that automated short answer systems can offer a valid means of assessment. Despite the fact that the technology for creating and using similar automated systems has existed for at least a decade, few have been developed and disseminated across fields. Building off previous work, this study focused on the VAIL assessment, described in more detail below. The VAIL relies on short answer responses, rather than essays, and appears to be particularly conducive to adaptation into an automated system.

Video Assessment of Instructional Learning (VAIL)

The VAIL is grounded in social learning theory, the notion that learning occurs largely through observation, as well as evidence suggesting that observational skills are valuable in developing expertise (Bandura, 1986; Jamil et al., 2015; Miller, 2011). The VAIL assesses observational skill by first asking teachers to watch a short video clip of an actual classroom and then identify and describe the effective teaching behaviors they observed. The process of "seeing" is an integral part of intentional teaching. One must be able to objectively identify and assess the effectiveness of specific practices in the classroom, and subsequently reflect on and modify personal teaching practices as necessary (Hamre, Downer, Jamil, & Pianta, 2012).

The content focus of the VAIL is teacher-child interactions, a topic particularly relevant to education. Teacher-child interactions have been consistently implicated as a means of promoting positive development in children (Burchinal et al., 2008; Thomason & LaParo, 2009; Yoshikawa et al., 2013). Subsequently, professional development has increasingly focused on training teachers to engage in positive interactions with children (Bierman, Nix, Greenberg, Blair, & Domitrovich, 2008; Domitrovich, Gest, Gill, Jones, & DeRousie, 2009; Pianta, Mashburn, Downer, Hamre, & Justice, 2008).

The VAIL has been validated (Jamil et al., 2015) and utilized in studies of both in-service (Hamre, Pianta, et al., 2012) and pre-service (Wiens, Hessberg, LoCasale-Crouch, & DeCoster, 2013) educators, offering a potentially valuable tool for teacher training and professional development. In a sample of pre-service teachers, demographic and programmatic characteristics did not consistently predict VAIL scores, suggesting that variation in VAIL scores were due largely to individual differences rather than group membership (Wiens et al., 2013). Furthermore, Jamil et al. (2015) found that the VAIL related to teachers' observed practices, suggesting that teachers who were more adept at identifying effective teacher-child interactions were also more likely to implement high quality teaching practices.

Furthermore, the National Center for Research on Early Childhood Education (NCRECE) Professional Development study, a randomized controlled evaluation of two forms of professional development, coursework and coaching, utilized the hand-scored VAIL as a measure of teachers' observational skills. At the end of the intervention, teachers enrolled in coursework were found to have significantly improved observational skills than teachers not enrolled in the course, and these improvements translated into meaningful changes in teachers' practice (Hamre, Pianta, et al., 2012; Downer et al., in press). Put differently, the VAIL was sensitive to intervention effects in the NCRECE course; this assessment tool seemed to detect important material teachers learned in the course that ultimately led to demonstrated improvements in practice. On the contrary, teachers' observational skills did not significantly change for teachers enrolled in the *coaching* intervention (Downer et al., in press). Although the reason for this is unclear, it may have been the result of the differences in content and delivery. For instance, it is possible that the VAIL may be more proximal to the content of the course, which focused on observing other teachers' practices, as opposed to the coaching intervention, which focused mostly on observing teachers' own personal practices.

In summary, although the VAIL measure is especially pertinent to current trends in education, the previously utilized hand scoring techniques may limit the scalability of the measure. As a result, the development of an automated means of scoring the VAIL was warranted. Automated scoring systems are indefatigable, systematic, and reliable, and offer a sustainable alternative to hand scoring. Nevertheless, it is necessary to ascertain that automated scoring systems are valid and useful, which were the aims of the present study.

Current Study

The present study explored the extent to which the VAIL, an assessment of teachers' observational skills of teacher-child interactions, could be adapted into an automated scoring system. Specifically, the goals were to determine whether the automated VAIL scoring system related to the previously validated hand scoring system; and was sensitive to intervention effects in the NCRECE professional development study. Building off of previous work utilizing hand-scores from the same study (Hamre, Pianta, et al., 2012; Downer et al., in press), we anticipated that the automated VAIL would be sensitive to intervention effects in the NCRECE course, but not the coaching intervention.

Method

Study Overview and Participants

This study utilized data from the NCRECE randomized, controlled evaluation of two forms of professional development designed to improve prekindergarten teachers' interactions with children. The NCRECE study was designed to evaluate scalable approaches to early childhood professional development (Pianta, Hamre, & Hadden, 2012). Teachers were placed randomly into treatment or control groups for the first phase, a 14 week (one semester) in-person course, and their group placements were then re-randomized for the second phase, the year-long MyTeachingPartner web-mediated coaching intervention.

The present study utilizes data on all teachers across all conditions who completed the post-intervention survey (n = 175). Seventy-two (41.1%) teachers received the course in phase I and 88 (50.3%) received coaching in phase II. Most teachers (95.9%) were female, and 48.5% of all teachers were African American, 34.5% White, 10.5% Hispanic, 4.1% multi-racial, and 2.3% Asian. Roughly half (50.9%) of teachers taught in Head Start centers and 37.1% worked in public schools. On average, teachers were 41.54 years old (SD = 10.41) with 10.93 years of experience teaching pre-kindergarten (SD = 7.64). In terms of their degree attainment, 34.5% of teachers held a bachelor's degree, 33.3% had less than a bachelor's degree, and 32.2% held an advanced degree.

It is important to note that the current sample represents a fraction (43.5%) of the entire NCRECE sample (n = 402). This reduced rate of completion is likely due to the fact that the post-intervention survey was optional. A series of t-tests were conducted to test whether those who completed the survey differed from those who chose not to complete the survey. Teachers did not significantly differ in terms of gender, age, and study condition; however, teachers who completed the survey were more likely to hold at least a bachelor's degree (M = .65, SD = .47) than those who did not complete the survey (M = .58, SD = .49; t (378) = -1.5, p <.01.) Teachers who completed the survey were also less likely to be Latino (M = .10, SD = .30) than those who did not complete the survey (M = .19, SD = .39; t (378) = 2.31, p <.001.)

The post intervention survey was sent via email and completed online. A portion of the survey included the Video Assessment of Interactions and Learning (VAIL) which is based on the Classroom Assessment Scoring System, or CLASS (Pianta, La Paro, & Hamre, 2008) a commonly used observational tool of teacher-child interactions. There are multiple forms of the VAIL focused on different elements of teaching; the present study focused on the VAIL designed to capture teachers' ability to detect aspects of Emotional Support, which includes creating a positive classroom climate, being sensitive to students' needs, and having regard for students' perspectives. To complete the VAIL, teachers viewed a two-and-a-half-minute video clip that depicted a teacher engaging a student in a conversation about her weekend. Teachers were instructed to watch the video as many times as they wished, but were encouraged not to spend more than 10 minutes on the video. Then teachers were asked to name up to five strategies the teacher used to support the student's social and emotional development. Finally, teachers were asked to, list a specific, behavioral example of each strategy from the clip.

Measures

Hand Scoring System. All VAIL responses were hand scored based on a previously established standardized rubric that aligns with the CLASS (Pianta, La Paro, et al., 2008). For each response, trained coders assessed four elements: (1) *strategy*, if the learner identified a behavioral indicator consistent with the CLASS; (2) *example*, if the learner provides a specific behavioral description from the video of the teacher demonstrating a strategy; (3) *breadth*, the specific CLASS indicator that the strategy is most consistent with; and (4) *match*, whether the strategy and example pair is representative of the same CLASS indicator.

In keeping with the behavioral indicators for the Emotional Support domain of CLASS (Pianta, La Paro, et al., 2008), VAIL responses could fall into twelve possible *breadth* categories, including: (1) *Relationships*: being in close physical proximity, engaging in shared activity, matching the child's affect, and engaging in social conversation with the student; (2) *Positive Affect*, smiling, laughing, showing enthusiasm; (3) *Positive Communication:* demonstrating verbal affection (4) *Respect*: maintaining eye contact, having a warm and calm voice, cooperating and sharing; (5) *Awareness:* anticipating problems, noticing difficulties; (6) *Responsiveness*: acknowledging emotions, providing comfort and individualized support; (7) *Addressing Problems:* helping in a timely and effective manner and resolving problems; (8) *Student Comfort*: the student seeks support, freely participates, and takes risks; (9) *Flexibility and Student Focus:* showing flexibility, incorporating the student's ideas, following the student's lead; (10) *Support for Autonomy and Leadership:* allowing choice, allowing students to lead lessons, giving students responsibility; (11) *Student Expression*: encouraging student talk and eliciting ideas; and (12) *Restriction of Movement:* allowing movement, not being rigid.

Collectively, this information was then used to calculate, per teacher, the total number of: correct *strategies*, correct *examples*, unique *breadth* scores, and strategy-example *matches*. Because teachers could enter up to five possible responses, the range for each of the four aforementioned categories was 0-5. The four categories rendered a Cronbach's alpha coefficient of .79. An excerpt from the coding manual is shown in Figure 1 along with further description of how the manual was used to assess the four components (*strategy, example, breadth,* and *match*) that were described above.

Automated Scoring System. The VAIL hand coding manual was used to build a "dictionary" for the automated system. The dictionary was organized in a format that could be utilized in automated scoring software, specifically the Linguistic Inquiry Word Count (LIWC) software (Pennebaker, Booth, & Francis, 2007). In an attempt to replicate the hand scoring system, key words and phrases for the automated dictionary were pulled directly from the hand coding manual. For instance, according to the VAIL manual the teacher demonstrates regard for the student's perspective by allowing choice, so the term "choic*" was included in the automated dictionary. As part of the LIWC software, any letters in the word that appeared *after* the asterisk were disregarded (therefore "choice" and "choices" would both be considered correct.) Additionally, notes were collected from trained VAIL coders regarding language used in teacher responses that varied from the coding manual, but was still considered correct (i.e., synonyms). Consistent with the previous example, words such as “choose”, “choosing” and “option” were included as synonyms of “choice”.

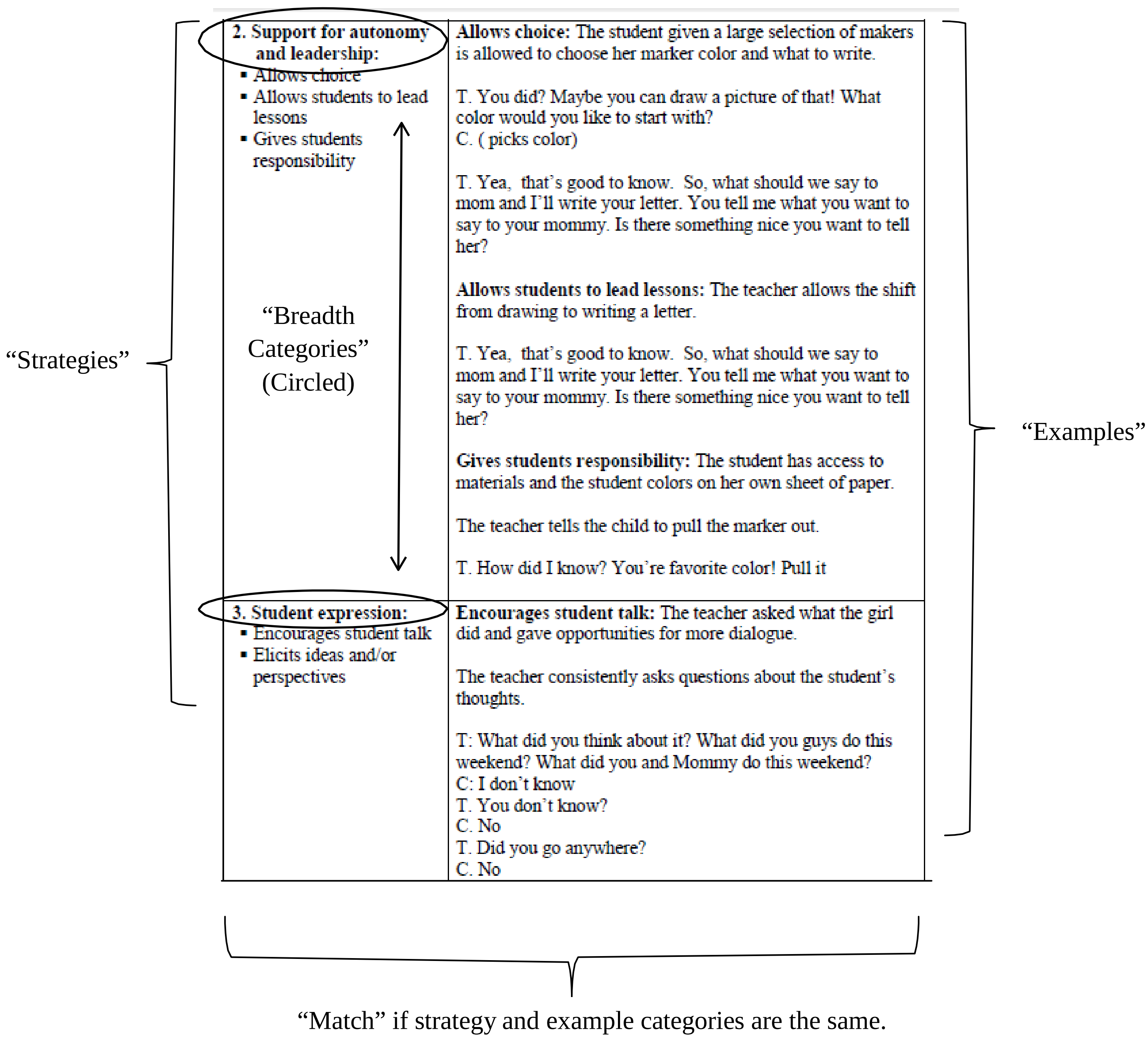


Figure 1. Excerpt from the VAIL Coding Manual

The LIWC software program allows the user to upload a dictionary organized broadly by a designated number of categories with various words comprising each category. As a result, our dictionary contained a total of 24 categories that represented the 12 possible breadth categories for strategies and, similarly, the 12 possible categories for examples. Thus, further building on the example outlined in Figure 1, in the automated dictionary, the *strategy* category for the "Support for Leadership and Autonomy" would be comprised of words such as "choice" and "responsibility". The corresponding *example* category would be comprised of phrases, such as "choose the marker" or "say to mommy." As previously mentioned, each teacher had up to five VAIL responses; to maintain consistency with the hand scoring technique, each response (one response per text file) was independently run through the software. The process of actually running all responses through the software was done simultaneously (all text files were run at the same time), allowing all responses to be scored in a matter of minutes.

LIWC provides output indicating the proportion of words that fall into certain categories, which, in this case, was used to assess whether or not the words used fell into any of the 24 previously mentioned categories. This information was then used to create: (1) *strategy* scores, represented as "correct" if a number other than zero appeared in any of the 12 strategy categories (suggesting that a word in the prescribed dictionary was used) (2) *example* scores, similarly, were "correct" if a number other than zero appeared in any of the 12 example categories (3) *breadth* scores, the specific category for which the strategy was located (4) *match* scores, if the strategy category matched the example category. Similar to the hand scoring technique, results ranged from 0-5. The four components rendered a Cronbach's alpha coefficient of .81. In addition, a mean score was also created by averaging across the four components in order to maintain consistency with previous work (Jamil et al., 2015).

Results

First, descriptive statistics (Table 1) were examined to compare the various components (i.e., strategy, example, breadth, and match scores) of the automated scoring system and the hand scoring system. The means, ranges, and overall distributions proved to be very similar. The automated scoring system did not consistently yield higher or lower results than the hand scoring system.

Table 1
Descriptive Statistics

	Mean	Std. Dev.	Range
Automated: Strategies	.86	1.26	0-5
Hand-scored: Strategies	1.05	1.35	0-5
Automated: Examples	3.60	1.52	0-5
Hand-scored: Examples	3.51	1.50	0-5
Automated: Breadth	.85	1.11	0-4
Hand-scored Breadth	.87	1.04	0-4
Hand-scored: Match	.65	1.05	0-5
Hand Code: Match	.78	1.18	0-5

Next, the four components of the automated system were correlated with the four components of the hand scoring system (Table 2). The bivariate correlations were consistently high across the four components, ranging from .72 to .87, suggesting that the automated scoring system was consistently replicating the hand scoring. High correlations were also observed among the strategy, breadth, and match components for both the automated and hand-scoring systems, which is not surprising given that breadth and match scores are only assigned when a strategy is “correct”, rendering high correlations among these elements. To further unpack these associations, Table 3 provides the percentage of cases in agreement and disagreement (by one, two and three or more points) among the automated and hand-scored systems. The automated and hand scoring systems agreed in roughly two thirds of all cases.

Table 2
Correlations among Automated and Hand Scoring System

	1	2	3	4	5	6	7	8
1. Automated: Strategies	1							
2. Hand-scored: Strategies	.82*	1						
3. Automated: Examples	.24*	.22*	1					
4. Hand-scored: Examples	.15	.15*	.87*	1				
5. Automated: Breadth	.92*	.74*	.24*	.15*	1			
6. Hand-scored: Breadth	.75*	.94*	.21*	.13	.73*	1		
7. Automated: Match	.92*	.79*	.28*	.19*	.85*	.72*	1	
8. Hand-scored: Match	.70*	.89*	.28*	.31*	.64*	.81*	.72*	1

Note: * = $p < .05$

Table 3
Percentage of Cases in Agreement and Disagreement among Automated and Hand Scoring Systems

	Exact Match	Disagree by 1	Disagree by 2	Disagree by 3+
Strategy	65.7%	25.7%	8.0%	0.6%
Example	62.9%	30.9%	5.1%	1.1%
Breadth	66.3%	24.6%	9.1%	0.0%
Match	65.1%	26.3%	6.9%	1.7%

Last, we explored the extent to which the automated system could be used to detect intervention effects. As previously mentioned, the first intervention phase was a course, and the second phase was a coaching intervention. As a result, we tested the extent to which the automated scoring system could be used to detect intervention effects by comparing the means for the treatment and control groups separately by phase (Table 4). For the course phase, three of the four components, as well as the mean score of the automated VAIL were found to be significantly different by intervention group. In particular, there were significant differences in scores for the treatment and control groups for strategy, breadth, match, and overall scores, such that the treatment group was significantly more likely to receive credit for these VAIL components than the control group. Associated effect sizes were moderately large, ranging from .54 to .60 (Table 4). There was no significant difference between groups for example scores. Consistent with findings from prior research using the hand-scoring method, there were no significant mean differences for any of the VAIL scores for the coaching phase. As shown in Table 4, the automated and hand-scored results yielded relatively similar effect sizes. In sum, these results suggest that the automated VAIL system replicated findings from the hand-scoring method.

Table 4

Mean Differences and Effect Sizes by Treatment Condition& Intervention Phase

			Automated			Hand
	t	*p*	*M (SD)* Treatment	*M (SD)* Control	Cohen's *d*	Cohen's *d*
Phase I	*df* = 145		*n* = 72	*n* =75		
Strategy	3.50	<.001	1.29 (1.61)	.56 (.81)	.57	.56
Example	.49	.89	3.65 (1.52)	3.53 (1.46)	.08	.05
Breadth	3.57	<.001	1.21 (1.30)	.57 (.81)	.59	.55
Match	3.69	<.001	1.02 (1.34)	.37 (.67)	.61	.49
Total	3.29	<.001	1.79 (1.19)	1.26 (.72)	.54	.51
Phase II	*df* = 173		*n* = 88	*n = 87*		
Strategy	.36	.55	.93 (1.17)	.79 (1.35)	.11	.07
Example	.66	.42	3.55 (1.57)	3.66 (1.49)	-.07	-.07
Breadth	.14	.71	.93 (1.09)	.76 (1.12)	.18	.14
Match	1.16	.28	.66 (.87)	.63 (1.20)	.02	.04
Total	.38	.97	1.52 (.96)	1.45 (1.04)	.07	.06

Discussion

Automated scoring systems offer a systematic and sustainable alternative to hand scoring techniques, which may be particularly valuable in large-scale interventions (Shermis & Burstein, 2013). The purpose of this study was to explore the extent to which an automated VAIL scoring system was able to replicate scores produced by the previously validated hand scoring system. Findings show that the automated VAIL scores were strongly related to the hand coded VAIL scores, suggesting that the automated VAIL consistently replicated previous scores. Similarly, the automated VAIL system detected similar intervention effects as the hand-scoring method. Taken together, these findings suggest that the automated VAIL appears to be a valid and reliable means of measurement.

It is worth noting that the strength of the associations between automated and hand-scored systems varied by what was assessed. More specifically, strategy and example scores had the strongest associations while match and breadth scores had the weakest. This is not surprising given that the automated dictionary was specifically built around the language of strategies and examples. In the VAIL, breadth and match are only scored if the strategy is found to be "correct", as neither breadth nor match are relevant if the strategy is incorrect. This has implications in our study because exact matches of *breadth* and *match* among automated and hand coding systems required an exact match on *strategy* as well. This likely accounts for the slightly higher discrepancy among the two systems for breath and match scores, although the correlations are still high.

The automated system replicated roughly two thirds of hand scores, which is slightly lower than previous work (Leacock & Chodorow, 2003). Of course, exact replication of the hand scored system is irrelevant in building an automated system given that the hand scores contain some degree of human error. Instead, we sought to develop a better system for measurement that would address issues in scalability.

Utilizing the automated VAIL or similar measures may be particularly advantageous in large-scale interventions, such as online courses, which are pervasive in various fields (U.S. Department of Education, National Center for Education Statistics, 2016). Automated measures allow intervention designers and instructors to incorporate open-ended questions into learner assessments, rather than simply relying on the more restrictive, close-ended alternative. Previously, open-ended questions posed logistical challenges for scoring in large-scale programs; however, automated systems offer a feasible alternative to manual hand scoring (Shermis & Burstein, 2013). Aside from open-ended questions being a better metric of student learning, this solution is also cost-effective and efficient (Williamson et al., 2012). The time spent designing and using an automated system is front-loaded, meaning more time is spent in the construction of the system than in the actual process of running the system to obtain scores (Brew & Leacock, 2013). By comparison, the hand scoring systems require more time to be spent maintaining highly reliable coders. Although human coders are likely to become more efficient over time, the possibility of making a mistake is ever-present (Ramineni & Williamson, 2013).

Although the VAIL is most applicable to teacher education, the findings from this study may be relevant to other fields. For this reason, it is useful to consider why the VAIL was particularly conducive to adaptation and how this information can be used in other studies. In this study, correct responses were finite and language-specific. More specifically, there were a fixed number of correct strategies that teachers could identify, and thus, the building of the automated dictionary encompassed words taken directly from the pre-constructed manual, as well as a limited number of associated synonyms. This was due in large part to the fact that VAIL responses were prompted by the viewing of a short video clip followed several short-answer questions. In other words, respondents were not asked to generate examples of Positive Climate from their potentially limitless repertoires of personal experience; instead, they were asked to identify concrete, specific examples from the video. Furthermore, there is less controversy surrounding the use of automated systems for scoring short answer responses, as compared to essays. AES has been criticized for oversimplifying the assessment to abstract concepts that may be difficult to quantify (Condon, 2013; Perelman, 2014).

Course designers in various fields could use relevant video, short writing excerpts, or images to elicit explicit concepts or facts conducive to automated scoring. Consistent with the use of videos to demonstrate how to *do* (or not do) something, videos could also be used to assess a learners' abilities to effectively *see* a skill or concept in action (Hamre, Downer, et al., 2012). For instance, kinesiology students could watch a video of a person exercising and analyze his gait or specific muscles that are being activated. Assuming that the video could be used to generate a finite number of correct responses, this information lends itself well to automated scoring, which would be particularly useful if a large number of students were enrolled in the course. In other fields, such as English or history, it may be most relevant to utilize writing excerpts or images to test learners' understanding of specific content. Ultimately, as online coursework continues to be scaled up, across fields, it is necessary to think creatively about how to effectively assess learning. Automated scoring systems offer one possible technique, especially for short-answer assessments that have a limited number of correct answers.

Limitations

In the present study, VAIL scores for the CLASS domain of interest were only collected at the conclusion of the intervention study. As a result, it was not possible to examine changes in VAIL scores from the beginning to the end of the intervention. In addition, the timing and delivery of the VAIL at the end of the post-intervention survey may have contributed to a lowered response rate. Those who self-selected to complete this survey proved to be different than those who chose not to complete the survey in terms of ethnicity and educational attainment.

Future Directions

Future work should continue to develop and test automated systems across fields. Given the availability of various automated software programs and the ecological validity of using automated systems, it is important for researchers to share their methods in developing and testing automated systems. The process of automating assessment scoring should be driven by the characteristics of the assessment questions and answers. Therefore, sharing techniques for automating various types of assessments, and establishing validity, can advance the use of automated scoring systems across fields.

In terms of the automated VAIL specifically, it would be beneficial to conduct similar inquiries using different samples of teachers enrolled in different interventions. Future studies should continue to utilize different samples to test the validity of the automated VAIL to verify that the current results consistently replicate. In light of recent findings that suggest automated scoring techniques differentially benefit native English speakers (Reilly et al., 2016), it would be advantageous to test the automated VAIL, and similar measures, in linguistically diverse samples to ensure validity. As previously mentioned the present study considered in-service teachers who were enrolled in coursework and/or coaching. Given the utility of the automated VAIL in large-scale interventions, future inquiries should also examine the extent to which the automated VAIL can be used to detect intervention effects using other intervention designs, such as online coursework. Additionally, based on previous work suggesting the VAIL is valuable in assessing teacher skill in both in-service and pre-service teachers (Wiens et al., 2013), it would also be informative to utilize samples of pre-service, as well as in-service teachers.

Conclusion & Implications

The design and delivery of scalable interventions and programs is of central importance to many fields utilizing online coursework. In teacher education, a variety of professional development programs for teachers have focused on training teachers to engage in positive interactions with children (Bierman et al., 2008; Domitrovich et al., 2009; Pianta, Mashburn, et al., 2008). However, such large-scale interventions pose challenges to assessment, especially the assessment of more complex open-ended responses. As a result, the automated VAIL appears to offer a valid and practically useful means of assessing teachers' skills in observing teacher-child interactions with implications for pre-service training, higher education, and professional development.

Acknowledgments

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grants R305B090002 and R305A060021 to the University of Virginia. The opinions expressed are those of the authors and do not represent views of the funding agencies.

References

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Upper Saddle River, NJ: Prentice-Hall, Inc.

Bierman, K. L., Nix, R. L., Greenberg, M. T., Blair, C., & Domitrovich, C. E. (2008). Executive functions and school readiness intervention: Impact, moderation, and mediation in the Head Start REDI program. *Development and Psychopathology, 20*, 821-843. doi: 10.1017/S0954579408000394

Biggs, J., & Tang, C. (2011). *Teaching for quality learning at university*. New York: McGraw-Hill International.

Boston, C. (2002). The concept of formative assessment. ERIC Digest. Retrieved from ERIC database. (ED470206).

Brew, C. & Leacock, C. (2013). Automated short answer scoring: Principles and prospects. In M. D. Shermis & J. Burstein (Eds.) *Handbook of automated essay evaluation: Current applications and new directions* (pp. 136-152). New York: Routledge.

Burchinal, M., Howes, C., Pianta, R., Bryant, D., Early, D., Clifford, R., & Barbarin, O. (2008). Predicting child outcomes at the end of kindergarten from the quality of pre-kindergarten teacher-child interactions and instruction. *Applied Developmental Science, 12*(3), 140-153. doi: 10.1080/10888690802199418

Burstein, J., Chodorow, M. & Leacock, C. (2003). Criterion online essay evaluation: An application for automated evaluation of student essays. Retrieved from American Association for Artificial Intelligence: http://www.aaai.org/Papers/IAAI/2003/IAAI03-001.pdf

Condon, W. (2013). Large-scale assessment, locally-developed measures, and automated scoring of essays: Fishing for red herrings? *Assessing Writing*, *18*, 100-108. doi: 10.1016/j.asw.2012.11.001

Dikli, S. (2006). An overview of automated scoring of essays. *The Journal of Technology, Learning, and Assessment*, *5*(1).

Domitrovich, C. E., Gest, S. D., Gill, S., Jones, D., & DeRousie, R. S. (2009). Individual factors associated with professional development training outcomes of the Head Start REDI Program. *Early Education & Development, 20*, 402-430. doi: 10.1080/10409280802680854

Downer, J. T., Pianta, R. C., Burchinal, M., Field, S., Hamre. B. K. …Scott-Little, C. (in press). Coaching and coursework focused on teacher-child interactions during language/literacy instruction: Effects on teacher beliefs, knowledge, skills, and practice.

Foddy, W. (1993). Constructing questions for interviews and questionnaires: Theory and practice in social research. Cambridge: Cambridge University Press.

Franks, R. P. & Schroder, J. (2013). Implementation science: What do we know and where do we go from here? In T. Halle, A. Metz, & I. Martinez-Beck (Eds.) Applying implementation science in early childhood programs and systems (pp 5-19). Brookes Publishing: Baltimore.

Gill, W. E. (2011). The Ready to Teach program. A federal initiative in support of online courses for teachers. Retrieved from http://files.eric.ed.gov/fulltext/ED530966.pdf

Hamre, B. K., Downer, J. T., Jamil, F. M., & Pianta, R. C. (2012). Enhancing teachers' intentional use of effective interactions with children. In R. C. Pianta (Ed.) *Handbook of early childhood education* (pp 507-532). New York: The Guilford Press.

Hamre, B. K., Pianta, R. C., Burchinal, M., Field, S., LoCasale-Crouch, J., Downer, J. T., . . . Scott-Little, C. (2012). A Course on Effective Teacher-Child Interactions: Effects on Teacher Beliefs, Knowledge, and Observed Practice. *American Educational Research Journal*, 88-123. doi: 10.3102/0002831211434596

Jamil, F. M., Sabol, T. J., Hamre, B. K., & Pianta, R. C. (2015). Assessing teachers' skills in detecting and identifying effective interactions in the classroom: Theory and measurement. *The Elementary School Journal*, *115*(3), 407-432. doi: 10.1086/680353

Landauer, T. K., Laham, D. & Foltz, P. (2003). Automatic essay assessment. *Assessment in Education, 10*(3), 295-308. doi: 10.1080/0969594032000148154

Leacock, C. & Chodorow, M. (2003). C-rater: Automated scoring of short-answer questions. *Computers and the Humanities, 37*, 389-405.

Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2009). Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies. Retrieved from U.S. Department of Education website: https://www2.ed.gov/rschstat/eval/tech/evidence-based-practices/finalreport.pdf

Miller, K. (2011). Situation awareness in teaching: What educators can learn from video-based research in other fields. In M. Sherin, V. Jacobs, & R. Phillip (Eds.) *Mathematics teacher noticing: Seeing through teachers' eyes* (pp. 51-65). New York: Routledge.

Palloff, R. M. & Pratt, K. (2008). *Assessing the online learner: Resources and strategies for faculty.* San Francisco: Jossey-Bass.

Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2007). *Linguistic Inquiry and Word Count: LIWC2007 Operator's Manual*. Retrieved from http://www.LIWC.net.

Perelman, L. (2014). When "the state of the art" is counting words. *Assessing Writing*, *21*, 104-111. doi: 10.1016/j.asw.2014.05.001

Pianta, R. C., Hamre, B. K., & Hadden, D. S. (2012). Scaling up effective professional development. In C. Howes, B. Hamre, & R. Pianta (Eds.) *Effective early childhood professional development: Improving teacher practice and child outcomes* (pp.191-212). Baltimore: Brookes Publishing Company.

Pianta, R., La Paro, K., & Hamre, B. K. (2008). *Classroom Assessment Scoring System*. Baltimore: Brookes Publishing Company.

Pianta, R., Mashburn, A., Downer, J. Hamre, B., & Justice, L. (2008). Effects of Web-mediated professional development resources on teacher-child interactions in pre-kindergarten classrooms. *Early Childhood Research Quarterly, 23*(4), 431-451. doi: 10.1016/j.ecresq.2008.02.001.

Ramineni, C., & Williamson, D. M. (2013). Automated essay scoring: Psychometric guidelines and practice. *Assessing Writing, 18*, 25-39. doi: 10.1016/j.asw.2012.10.004.

Reilly, E. D., Williams, K. M., Stafford, R. E., Corliss, S. B., Walkow, J. C., & Kidwell, D. K. (2016). Global times call for global measures: Investigating automated essay scoring in linguistically-diverse MOOCs. *Online Learning*, *20*(2), 97-109.

Shermis, M. D., & Burstein, J. (2013). *Handbook of automated essay evaluation: Current applications and new directions.* New York: Routledge.

Thomason, A.C. & La Paro, K.M. (2009). Measuring the quality of teacher-child interactions in toddler child care. *Early Education & Development*, *20*(2), 285-304. doi: 10.1080/10409280902773351.

U.S. Department of Education, National Center for Education Statistics. (2016). *Digest of Education Statistics*, *2014* (NCES 2016-006), Table 311.15.

Vale, K & Littlejohn, A. (2014). Massive open online courses: A traditional or transformative approach to learning? In A. Littlejohn & C. Pegler (Eds.) *Reusing open resources: Learning in open networks for work, life and education* (pp. 138-153). New York: Routledge.

Wiens, P., Hessberg, K., LoCasale-Crouch, J., & DeCoster, J. (2013). Using a standardized video-based assessment in a university teacher education program to examine pre-service teachers knowledge related to effective teaching. *Teaching and Teacher Education*, *33*, 24-33. doi: 10.1016/j.tate.2013.01.010.

Williamson, D. M., Xi, X., & Breyer, F. J. (2012). A framework for evaluation and use of automated scoring. *Educational Measurement: Issues and Practice*, *31*(1), 2-13. doi: 10.1111/j.1745-3992.2011.00223.x

Yoshikawa, H., Weiland, C., Brooks-Gunn, J., Burchinal, M. R., Espinosa, L. M., Gormley, W. T., & Zaslow, M. J. (2013). Investing in our future: The evidence base on preschool education. Retrieved from Society for Research in Child Development website: http://www.srcd.org/sites/default/files/documents/washington/mb_2013_10_16_investing_in_children.pdf