Interactions in Online Courses and Student Academic Success

By

Deborah L. Taylor

BA, Mid-America Nazarene University

MA, The University of Kansas

Submitted to the graduate degree program in Educational Leadership and Policy Studies and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

____

Chairperson Dr. Ronald Aust

____

Dr. Young-Jin Lee

____

Dr. Susan Twombly

____

Dr. Bruce Frey

____

Dr. David Hansen

Date Defended: October 28, 2014

The Dissertation Committee for Deborah L. Taylor

certifies that this is the approved version of the following dissertation:

Interactions in Online Courses and Student Academic Success

____

Chairperson, Ronald Aust

Date approved: 12/05/14

Abstract

The Sloan Foundation reported a phenomenal 17% annual growth in online course enrollment from 2002 through 2012, while the overall enrollment in higher education has shown only a 2.5% annual growth. Despite the growth in enrollment, McFadden and others have reported that drop out rates were as much as seven times higher for online courses. To address these concerns, this study investigated how three types of interactions (student-student, student-content and student-instructor) in the first two weeks of fully asynchronous online courses were associated with student academic success.

This investigation analyzed archived tracking data from a learning management system for 1,703 students in 200 semester-long, fully online community college courses. Multinomial logistic regression analysis was used to consider the relationship between the three types of online interactions and academic success. The three types of independent interaction variables were measured as follows. Student-student interaction was the number of student posts to the discussion forums. Student-content interaction was the total number of pages the students accessed. Student-instructor interaction was a measure of the number of instructor posts and the number of instructor emails. The outcome measure, academic success, fell into three groups: successful completers (students completing the course with A, B, or C grades), low score completers (students completing the course with D or F grades), and non-completers (students who did not complete the course).

Odds ratios derived from the regression analysis were used to determine the percentage of change in academic success attributable to each type of interaction, holding all other factors constant. The multinomial logistic regression was statistically significant (chi square = 461.96 $p < .001$), indicating that the predictors reliably distinguished between the three outcome groups.

The findings suggest that increasing the number of times a student posts by one unit (5 posts) would increase the individual's odds for success by 74% for the non-completers group and by 71% for those in the low score completers group. The findings also suggest that if an individual would increase interaction with the content by one unit (49 online pages) the individual's likelihood for success would increase by 57% for non-completers and 39% for low score completers. The number of instructor posts had no effect on the outcome for low score completers, however increasing the number of instructor discussion posts by one unit (15 posts) increased the likelihood that the students would complete the course by 34%. An increase in one unit (151 messages) of instructor email was associated with a 45% decrease in the odds for success for non-completers and a 28% decrease in the odds for success for the low score completers.

This study provides support that student-student interaction and student-content interaction in the first two weeks of online courses contribute to student academic success with student-student interaction being most influential. More instructor postings in the discussion areas increased the likelihood that students would complete the course. The finding that the more instructors interacted with students by email the lower the academic success seems counterintuitive at first glance. This may be because instructor emails are often in response to increased requests for clarification by students and could be a reflection of poor course organization or insufficient course support materials.

Table of Contents

List of Tables

List of Figures

Acknowledgements

There are a number of people who I would like to thank for their support and/or assistance in the completion of my doctoral degree and this dissertation. First of all I would like to express my sincere appreciation and gratitude to my chair and advisor, Dr. Ronald Aust. Your guidance, patience, encouragement, and positive outlook were greatly appreciated throughout the process, but especially in the final stages when I am pretty sure I was a bit challenging to work with.

I am also indebted to the members of my committee for the assistance that each provided in his or her own special way. Thank you, Dr. Bruce Frey for meeting with me numerous times as we worked out the possible statistical methods and design as well as for encouraging me that, yes, this could be done, but also to keep my study focused so that it could actually be completed. Thanks to Dr. David Hansen for your enthusiasm, guidance, and for helping me organize my data. Thanks so much to Dr. Susan Twombly for your time providing your excellent comments and suggestions for my writing. Thank you Dr. Young-Jin Lee for providing suggestions and cautions in dealing with such a large data set.

There are a few friends and colleagues that I owe thanks to as well. I would like to thank Dr. Robert Isaacson for reading through several chapters and providing suggestions and guidance in my writing and for the encouragement you provided. I would like to thank my colleague, Ruben Noguera, for the technical assistance of retrieving the data from the learning management system, not once but several times as I figured out what it was that I really needed for the study. A very special thanks goes to my dear friend and colleague, Lorrain Green, who has been my unofficial "coach and cheerleader" throughout the entire process, encouraging me when I needed it and pushing me when I needed that as well. Your positive attitude and enthusiasm kept me on target and your focused questions helped me keep things in perspective, thank you so much!

Finally, to my husband, Jon, whose love, support, and encouragement were constant positive factors that allowed me to achieve something that I've dreamed of for many years. I can't thank you enough. I also want to thank my two sons, Graham and Raymond, for their support through this process as well.

CHAPTER I: INTRODUCTION

Technology has forever changed the face of higher education (Renes & Strange, 2011; Appama, 2008). Online learning has been edging into the mainstream of education with more than 6.7 million students enrolling in at least one online course during the Fall of 2011, an increase of nine percent, while the overall growth of the higher education student population has declined (Allen & Seaman, 2013). Students are embracing online courses for the flexibility they provide thus accommodating work and family demands (Allen and Seaman, 2010; Appama, 2008; Hart, 2012; Oblinger and Oblinger, 2005). For-profit and community colleges led the way in providing online courses to their students (Cejda, 2010). In a study of online classes and community colleges, Cejda (2010) noted “more students were enrolled in online courses at community colleges in fall 2006 than in all other types of institutions combined.” Online learning is growing in all sectors of higher education with 69.1% of the chief academic officers responding that online learning is critical to their long-term strategy, yet there is a perception that the lower retention rate for online courses remains a barrier to the growth of online instruction (Allen & Seaman, 2013).

The face of students has changed as well. Adult learners have displaced the traditional age student as the dominant learner in the 21st century (Boston, Ice, & Gibson, 2011; Lea, Stephenson, & Troy, 2003; Pew, 2007). The learning preferences of the adult learner are not the same as those of the traditional learner (Park & Choi, 2011; Pew, 2007). Successful online learners show characteristics and preferences that coincide with andragogy, a theory of adult learning, introduced to American educators by Malcolm Knowles in 1973. Knowles identified six principles of adult learning: (1) adults are internally motivated and self-directed, (2) adults bring life experiences and knowledge to learning experiences, (3) adults are goal oriented, (4) adults are relevancy oriented, (5) adults are practical, and (6) adults learners like to be respected. While Knowles originally identified andragogy with adult learning, many of the principles can be applied to any student who is somewhat independent and self-directed and the original definition has been broadened and is no longer just applied to adult learners (Pew, 2007). At the very core of online learning is the concept of autonomous learning in which the individual chooses when and where to study and to interact with the course. Those students who are described as self-directed learners do very well in the online environment. An attempt will be made to distinguish between andragogy and pedagogy throughout this writing, but at times the more broadly accepted term "pedagogy" will be used in a more global form and is not meant to specifically refer to the teaching techniques for children.

Several researchers have suggested that the biggest obstacle for research, policy, and practice for online learning is the inconsistent use of terminology to describe the components of online learning (Moore, Dickson-Dean, & Gaylen, 2011; Tallent-Runnels et al., 2006; Thurmond & Wambach, 2004). Even conducting research on the topic is challenging when the terms online learning, e-learning, online classes, virtual classes, asynchronous learning, web-based classes are often used interchangeably. For the purpose of this study, the term "online class" will refer to classes in which all of the content is delivered online and all contact is online (Allen and Seaman, 2011; Tallent-Runnels et al., 2006). Since these classes typically have no face-to face meetings and are conducted entirely online they are labeled as asynchronous online classes.

A Brief History of Online Education

Online education evolved from distance education and distance education by definition must be mediated technologically. As technology provided new means of communication, the mode of delivery changed, and distance education evolved. This resulted in the classification of three generations of distance education. The first generation of distance education was postal correspondence, followed by the second generation of mass media of television, radio and film production. The third generation of distance education introduced interactive technologies that incorporated the connectivity of the Internet and now comprises the dominant form of delivery methods today (Anderson and Dron, 2011, p. 81).

In correspondence courses, the students received a print copy of the "activities" for the course that primarily consisted of instructions for reading from the textbook, guidelines for study and essay questions to answer and turn in for a grade. The exams were proctored (by someone other than the instructor) and were essay in form. The instructor would receive the assignments by mail, grade them and return them, with this transactional feedback interaction requiring days and weeks to complete the cycle. It was rare to have any direct communication, i.e. via telephone conversations; for the most part, all communications between the student and instructor took place via the postal service. In contrast, current students can send emails and receive immediate response if the instructor is available, and best practice models suggest a 24-48 hour maximum window between student communication and instructor response. Students can also post in the discussion forum for all to share in the communication and responses can be immediate if another student or the instructor is online at that time. Chat tools are incorporated in most course delivery systems and synchronous communications via the Internet can also offer immediate interactions between students and instructors. As technology has improved, more interactive and immersive options have been made available to the online instructor.

Historically students taking distance education courses were those who lived in rural or remote locations, but with the advent of the Internet, its place in education has shifted. In many colleges and universities, online education is no longer considered “distance education” but rather a course option for all students. An online class refers to a class that is conducted entirely via the Internet with no face-to-face meetings; there are two types, asynchronous in which the participants are not all active at the same time, but that they interact with one another as part of the course activities, and synchronous in which everyone participates at the same time using various communication tools (Mupinga, Rora, &Yaw, 2006; Olivera, Tinoca, & Pereira, 2011; Rovai, 2002; U.S. Dept. of Education, 2009). Most online courses are delivered via learning management systems (LMSs) and communication tools are key components of most of those systems. The communication tools include a private email area allowing the option of individuals to communicate in a one-to-one manner and a discussion forum where everyone in the class has access to the information being communicated. In an asynchronous online course the participants are not all active at the same time, but interact with one another as part of the course assignments, thus fitting the classical description of being separated across time and space (Mupinga, Rora, and Yaw, 2006; Rovai, 2002).

Statement of the Problem

The population of students in higher education has changed dramatically in the past twenty years (Lea, Stephenson, & Troy, 2003). Institutions are offering more and more online courses to meet the demand of students and this growth is supported by the affordances of today’s technology. Students enter the online world because it offers a delivery format that is convenient, immediately accessible, and flexible (Mupinga, Nora, & Yaw, 2006). Students are demanding online classes and schools are attempting to meet those demands, yet many studies show that students withdraw from those courses more often than they do conventional courses. If institutions are going to meet the needs of the students by providing online classes, then studies need to be done to determine what factors contribute to success in online courses, incorporate those factors into the courses and provide instruction for the faculty teaching those courses as to how to implement the courses in a way that students can succeed.

Online courses exhibit a much higher attrition rate than do traditional courses (Carr, 2000; Lee & Choi, 2011; Morris & Finnegan, 2009). Students enroll, but they don't complete the classes, some even withdraw before the first assignment is due (Simpson, 2004). The dearth of empirical research on community college attrition in online classes is puzzling when a wealth of data is routinely collected by the learning management system as well as the institutional management system (Hachey, Conway, & Wladis, 2013). Higher Education is already being criticized for low retention and completion rates, and the American Graduation Initiative (Obama, 2009b) has zeroed in on community colleges and their responsibilities (Obama, 2009a). Online courses open up the educational field to those who might not otherwise be able to attend a traditional face-to-face classes due to distance issues, child-care needs, or work demands, yet if students don't complete the courses then nothing has been gained.

Purpose of This Study

The main purpose for this study was to determine if the level of student activity in the first two weeks of an online class is related to student academic success. To meet this purpose this research used a convenience sample of archived data from LMS activity logs. Three types of online course interactions—student-student, student-instructor, and student-content -- were investigated to see if any of the interactions were associated with student academic success. If the level of activity were associated with student academic success and if any mode of interaction were more effective at promoting retention and student academic success, this could provide guidelines for online course development and delivery.

By examining the student interaction patterns in the first two weeks, an early warning system might be established in which interventions could be provided early in the course to allow the student to adjust her or his behavior and increase chances for a successful outcome.

Research Questions

The key questions guiding this research were:

RQ1: Is interaction activity in the first two weeks of class associated with student academic success in the online classroom?

RQ2: Do the interaction types as a set reliably distinguish between the three levels of student academic success?

RQ3: Is there a relationship between each of the interaction types and student academic success?

RQ4: What is the effect of each of the interaction types on student academic success?

RQ5: Are some interaction types more important than others in promoting student academic success?

Rationale for the Study

There are both practical and theoretical reasons for conducting this study. Promoting student academic success has both institutional as well as personal advantages. With a higher success rate, more students can be served because they will move through the class rather than taking up additional seat time repeating the class and preventing other students from enrolling. As Lee and Choi (2011) argue, “If completion rates could be improved, institutions would make better use of resources without waste and administrators could plan budges for future fiscal years more efficiently” (p. 594). Additionally recent changes in federal financial aid has reduced the number of semesters that a student can receive aid, thus increasing the need for students to successfully complete classes on their first attempt. Identifying students at risk early in the semester would allow time for intervention in a timely manner and perhaps increase the completion rates.

On a theoretical level, this study will use Moore's Interaction Theory of three different types of interactions in the online classroom critical for learning.

Significance of the Study

Most of the studies examining student academic success in online classes have examined retention and student academic success at universities or four year schools, only a few have examined student academic success in the community college student. Those studies that did examine community college retention and student academic success focused on demographics and traditional retention factors and not learning analytics based on the tracking data generated in the learning management systems. The combination of high attrition rates in online classes, poor time to completion in both the two year and the four year schools, public concern with the quality of higher education, and the astronomical debt many students are generating while earning a college degree create a compelling reason for more studies on student retention and online success are needed all levels and types of institutions of Higher Education. Studies that uncover information that can be used to develop quality online courses promoting student retention and success would enable students to be successful in their coursework and complete their degrees in a timely manner.

This study addresses just one small area of the aforementioned issues, but if the early interaction data successfully predicts student academic success, then a guidance could be offered to help reduce the rate of attrition in online classes and thus increasing retention and student academic success. It is hoped that the information gleaned from this study could contribute to the body of knowledge identifying which of the interaction variables are most effective for promoting student academic success. This information could be utilized to help develop online course delivery and design to promote student academic success. While many instructors may recognize that low interactions on the part of a student leads to higher attrition rates, these anecdotal observations need empirical support to be sustained. This study used tracked interaction activity as a reflection of student behavior in community college online courses as a means of identifying activity promoting student academic success.

Hypotheses

Hypothesis 1– The amount of online course interactions (with course content, instructor, and other students) in the first two weeks of class is related to successful completion.

Hypothesis 2 – Student-content interaction promotes successful completion in online classes.

Hypothesis 3 – Student-student interaction promotes successful completion in online classes.

Hypothesis 4 – Student-instructor interaction promotes successful completion in online classes.

Hypothesis 5– Some types of interactions are better promoters of success than others.

Summary

Online education is here to stay and in order to meet student needs educators must continue to evaluate what components are necessary for student academic success in online classes. It is the goal of this study to determine if activity patterns of the three interaction modes occurring in online classes can predict student academic success, and if they do, to see if one is more important than the others in promoting success. It would be extremely beneficial if those predictors appear early enough in the class to indicate students at risk to allow the instructor to promote an intervention to increase the student's chance of success. This could become part of a model to predict at risk students in online classes and could provide guidance in course development and design to increase the student academic success rate in those classes.

Definition of Terms

Analytics: The science of examining data to draw conclusions, and when used in decision making, to present paths or courses of action

Asynchronous: The participants are not all active at the same time, but interact with one another as part of the course assignments, thus fitting the classical description of being separated across time and space

Interaction: "A dialogue or discourse or event between two or more participants and objects which occurs synchronously and/or asynchronously mediated by response or feedback and interfaced by technology" (Muirhead & Juwah, 2005, p. 13).

Instructional interaction: An interaction that takes place between a learner and the learner's environment that changes the learner's behavior towards an educational goal (Wagner, 1994, p. 8)

Interaction in distance education: "…the learner's engagement with the course content, other learners, the instructor, and the technological medium used in the course. . . . Ultimately, the goal of interaction is to increase understanding of the course content or mastery of the defined goals" (Thurmond & Wombach, 2004, p. 4).

Moore (1989) identified three types of interactions in online learning:

1) *Learner-instructor interactions* are those interactions between the learner and the instructor and will be determined by counting: emails (within the LMS), announcements (within the LMS), discussion posts made by the instructor (in the discussion forum of the LMS). Each of these types of interactions are tracked by the

LMS and the data from the log files form the LMS will be used to "count" each of these and they will be totaled to make one variable.

2) *Learner-content interactions* are those interactions between the learner and the content and tabulated within the LMS log files, each time the leaner "opens" a content file, this will constitute a "hit" and will be counted. Each time the student opens a page it will count as a "hit" with multiple visits resulting in multiple counts. It is the interaction with the content that is being tabulated.

3) *Learner-learner interactions* are those interactions between learners in the course, these interactions occur when students participate in the discussion forum. The activity is tracked by counting the number of initial posts, the number of reply posts and the number of other students' posts that the student reads and the counts will be combined to represent the learner-learner interaction variable.

Interaction treatments: The conditions or environment designed and arranged by teachers to encourage interaction behaviors (Bernard et al., 2008, p. 1248).

Learning Analytics: The measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs (Lockyer & Dawson, 2011)

Learning Management System: A collection of Web applications that integrate technological and pedagogical features of the Internet and the Web into a single, template-based authoring and presentation system that facilitates the design, development, delivery, and management of Web-based courses and online [distributed] learning environments (Dabbagh & Bannan-Ritland, 2005, p. 298)

Online learning: A distributed learning environment that uses pedagogical tools, enabled by

Internet and Web-based technologies, to facilitate learning and knowledge building through meaningful action and interaction (Dabbaugh & Bannam-Ritland, 2005, p. 15)

Process measures: Measures that capture data describing physically observable actions that can be collected through a LMS to identify student behaviors that imply learning (Nickels, 2005, p. 112).

Student academic success: In this study "student academic success" refers to students' ability to complete a course with a grade of A, B, or C. Student success consists of three groups:

Successful students are those students who completed the course with an A, B, or C grade.

Low score Students are students who completed the course and earned D or F

Non-completers are those students who did not complete the course and earned either a W, or WA grade.

CHAPTER II: REVIEW OF THE LITERATURE

Introduction

This study examines the relationship between online course activity and rates of student academic success in online courses. The literature base that has been used to frame this study comes from four groups of literature: 1) interaction in the online classroom, 2) learning analytics, 3) student behavioral factors, and 4) online course design for success. The theoretical framework is based upon Moore's Theory of Interaction (1989) and Anderson's Interaction Equivalency Theory (2003).

Context of the Study

For the purpose of this study an online class will refer to a class that is conducted entirely over the internet with no face-to-face meetings; there are two types, asynchronous in which the participants are not all active at the same time, but that they interact with one another as part of the course activities, and synchronous in which everyone participates at the same time using various communication tools (Mupinga et al., 2006; Olivera et al, 2011; Rovai, 2002a; U.S. Department of Education, 2011).

A major concern in higher education is the retention of students (Hart, 2012). Retention in online courses is much lower than that found in traditional classrooms (Allen & Seamen, 2013; Lee & Choi, 2011). However, student demand is high for online classes (Capra, 2011), raising the question of what can be done to reduce attrition and promote successful course completion? In order to determine what can be done, one must first discover what is going on in the online classroom. Many of the previous studies on student academic success and retention relied upon self-report studies, however with the advent of the use of learning management systems and their tracking capabilities, researchers no longer have to ask the students if they were engaged. Utilizing tracking records as indicators of student behavior online provides an additional source of information regarding student behaviors and how they relate to successful completion of online courses. An important component, however, is the necessity of providing opportunities in the online classroom for students to interact, not only with one another and the instructor, but with relevant, engaging content. Content availability is a prerequisite for content interaction and is represented by the number of content items in each course. In addition, it reflects one component of teaching presence in the composite for student-instructor interaction.

Interaction in the Online Classroom

Juwah (2006) asserts that interaction and interactivity are the key success factors underpinning the pedagogy of online education. Yet, the terminology is often confusing. Muirhead and Juwah (2007) explained the difficulty of understanding the constructs of interactivity and interaction in online education describing them as "complex multifaceted phenomena", and explaining that they have different meanings in a variety of contexts. Yet they agreed with a number of distance educator researchers (Anderson, 2002; Hirumi, 2002, Yacci, 2000) in noting that they are "critical in promoting and enhancing effective learning." Typically, an interaction is thought of as an active exchange of actions and information among individuals. Wagner (1994, p.8) defined interactions as "reciprocal events that require two objects and two actions and these interactions occur when the objects mutually influence one another" (i.e. teachers and students), but interaction has been extended to include student interaction with the curricular content in the educational environment (Moore, 1989). Yacci (2000) thought of interactivity as mutually coherent message loops with instructional interactivity in which learning occurs from the learner's point of view. Thurmond and Wambach (2003) captured the essence of interactivity in distance learning by saying it is "the learner's engagement with the course, content, other learners, and the technological medium used in the course. Ultimately the goal of interaction is to increase understanding of the course content of mastery of the defined goals" (p. 4). Building upon this Murihead and Juwah described interaction as "a dialogue or discourse or event between two or more participants and objects which occurs synchronously and/or asynchronously mediated by a response or feedback and interfaced by technology" (2004, p. 13).

Moore's Interaction Theory

The Theory of Interaction developed in Michael Moore's (1989) seminal work describes the three types of interactions that occur in the online learning environment. The first type of interaction is the learner-content interaction and it describes the transfer of information that occurs as the learner reads the instructions and is highly dependent upon the content being inviting, clear and visible. The second type of interaction is the learner-learner interaction and is a critical component in promoting the constructivist concepts of learning. The primary environment where this occurs in most online courses is in a discussion forum where students can comment and reply to one another. The third type of interaction is that between the student and instructor, Moore (1989) called it the learner-instructor interaction. This includes interaction that occurs in the discussion forum, but also would include email correspondence as well as announcements if the course design permits those communication tools.

Student-Student Interaction Research. The majority of research on interactions and student academic success and/or retention primarily has been focused on the student-student interactions and the importance of building a community in the online course (Anderson & Dron, 2010; Annand, 2011; Daspit & D'Souza, 2012; Dawson, 2006; Garrison, Anderson, & Archer, 1999; Kanuka, 2011). The concept of a community of learners primarily was based upon

Tinto's (1975) retention research, and Kember's (1995) work with adult learning in open online classrooms, both of which identified and emphasized the importance of student-student interactions. Rovai, building upon this, conducted a number of studies examining the role of community (Rovai, 2001; 2002a; 2002b; 2003; 2004) and found when learners feel a sense of community in the online environment they are less likely to drop out of an online class. He developed guidelines to promote a sense of community in the classroom as well as creating a measurement tool to determine the students' sense of community in online classrooms. Along those lines the Community of Inquiry research extends Moore's Interaction Theory by assigning presences to describe the collective nature of the interactions and not the interactions themselves (Garrison, et al. 1999). This theory has spawned voluminous research with much of it focused on the importance of the social presence created from the student-student interactions (Garrison & Arbaugh, 2007; Parscal, Sherman, Heitner, & Lucas, 2012; Richardson, Arbaugh, Cleveland-Innes, Ice, Swan, Garrison, 2012). In 2001, John Seely Brown was one of the first to suggest that universities use online activities to actively engage the student in the construction of their own learning experience using the content from the course (p. 77).

Student-Instructor Interaction Research. Student-instructor interaction has been thoroughly explored as well. Dawson and McWilliam assert, "instructor facilitation of discussion and an ongoing active presence in the discussion forum and other online communication media is essential for maintaining high levels of student engagement" (2008, p.31). Morris and Finnegan (2009a) concur, "When faculty are present ('visible' and active) in the online environment, students benefit and student participation increases (p. 61). The constructs of teacher immediacy and teacher presence have been shown to be valuable in student learning (Baker, 2010; Morris & Finnegan, 2009a; Ni & Aust, 2008; Russo & Benson, 2005).

Student-Content Interaction Research. Student-content interaction has been studied the least. Some attention has been paid to student-content interaction, but until recently, most interaction studies involved self-report student surveys, many of which were satisfaction surveys. From the results of reviewed studies, students who actively participated in learning interactions, especially with teachers and content, were more likely to complete and be retained in online courses (Lee & Choi 2011, p. 609). There have been some new directions for research in this area, but most were conducted in the university and may not generalize to the community college or even small private schools (Campbell, 2007; Dawson & McWilliam, 2008; Finnegan, Morris, & Lee, 2009; Morris, Finnegan, & Wu, 2005; Morris & Finnegan, 2009a, b; Morris, Wu, & Finnegan, 2005). Adding to the issue of the research on online learning often not being generalizable to community colleges and small schools, the absence of a common theory on how courses should be designed and delivered resulted in many diverse systems and evaluations making it difficult to create policy for online teaching and learning. "Fundamental confounds associated with different media, different pedagogies, different learning environments, and so forth, mean that causal inferences about the conditions of design, pedagogy, and technology use are nearly impossible to make with any certainty" (Bernard et al., 2009).

Andersons' Interaction Equivalency Theory

Building upon Moore's work on interaction, Terry Anderson developed the Interaction Equivalency Theory (2003) postulating that deep and meaningful formal learning is supported as long as one of the three forms of interaction (student-teacher, student-student, student-content) is at a high level. He feels that the other two interaction types may be offered at minimal levels or even eliminated, without degrading the educational experience. However, the second thesis of the theory postulates that high levels of more than one of these three models will likely provide a more satisfying educational experience, although these experiences may not be as cost- or time-effective as less interactive sequences. Anderson qualified this by observing that promoting the student-content interaction more effectively promoted learning than the other two types of interactions. This was supported in a meta-analysis by Bernard et al., (2008) in which a linear relationship was found between strong course design elements that promote the different interactions and course outcomes, yet it was only strengthening the student-content interaction treatment that improved the effect size of course outcomes.

At the *International Conference on E-Learning 2013,* Terry Anderson explained that student-content interaction provides two qualities that make that particular mode of interaction quite attractive to the higher education community. First, it is scalable, has the capacity to record once and then play back for many, thus moving the course from one-to-one interactions to one to very large numbers is a known cost-effective strategy. In addition, the flexibility of anytime/anywhere learning is extremely attractive to learners who have families and jobs. However, learning involving a high level of student-content interaction requires high degrees of autonomy and self-direction that many students lack (Anderson, 2013; Garrison, 1997). Therefore promoting behaviors that increase autonomy and self-directed learning may be needed to improve student retention and success.

In an attempt to discover if course design is important in directing interaction behaviors, Bernard, et al. (2008) conducted a meta-analysis of research on interaction treatments, the conditions or environment designed and arranged by teachers to encourage interaction behaviors, and not the interactions themselves (actual behaviors that can be observed and recorded). Their findings showed that the interaction treatments allow for the interactions to occur, and if the conditions are not present to encourage interaction of one type or another, then it is not possible for those interaction types to occur. Examining student-content interactions was not feasible in early online courses, however technological advances in LMSs now provide the ability to track student-content interaction thus providing new methods of research in online learning.

Teaching Online

The greatest challenge in delivering an asynchronous online class is that all of the interactions are computer mediated. The mode of delivery constrains the pedagogies that can be used in teaching. In a traditional college classroom, the primary mode of delivery is face-to-face verbal and nonverbal communication. The students know what to expect since this has been the primary classroom model for most if not all of their educational experiences. They prefer that the instructor deliver most of the content if not all by lecture method (Charbonneau, 2012). There may be some discussion in the classroom, but that is not the primary delivery method. Classroom interactions follow the traditional definition of interactions in which the instructor lectures, perhaps students ask questions, which the instructor answers – it is immediate and bidirectional, an exchange of actions and information occurs (Ni, 2013). In addition, both the student and the instructor are in the same place at the same time.

However, in an asynchronous online class, communication occurs via computer and the Internet with the students and instructor separated by time and space. This separation was identified by Moore (1989) as a transactional distance and he postulated that the highest level of learning occurring when that transactional distance is minimized. Kanuka (2011) suggested that minimizing that distance might be accomplished by utilizing the three variables and two dimensions of Moore's transactional theory when delivering a course online. The two dimensions are teaching and learning. Dialogue and structure are the variables on the teaching dimension and learner autonomy is the variable on the learning dimension. Dialogue is the interaction between the instructor and the learners, while the elements of course design compose the structure. When there is a high degree of structure in an online course and a high degree of interactive dialogue, then the transactional distance is minimized promoting student engagement and success. Learner autonomy is supported in a highly structured course that allows a student to decide how much guidance and directions are wanted and needed. "The praxis of this theory, then, involves determining the right mix of structure, dialogue, and autonomy for achieving successful distance learning transactions" (Kanuka, 2011, p. 154).

"Getting the mix right" is the holy grail of online course development. Distance educators were challenged to get the mix right between independent study and interactive learning system as early as 1979 (Daniel and Marquis, 1988) and the challenge has become more difficult with the advent of computer mediated interaction. There is no one right answer, but as the instructor works through the course development, it is imperative that there is an attempt to "get the mix right" so that student engagement, retention, and success are possible.

Learning Analytics

The 1st International Conference on Learning Analytics and Knowledge was held in March of 2011. In that conference learning analytics was defined as "the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs" (Booth, 2012, p. 1). The increased availability of the LMS (learning management system) for course delivery has opened up new possibilities for research on student academic success. When students log in to the system, all activity in the course is tracked and recorded, thus producing massive amounts of data. This enables new ways to assess and quantify student activity, particularly in the student-content interactions. In the early stages of LMS use for course delivery, it was possible to evaluate student activity in discussion forums. In addition, instructor presence could be quantified since there were recorded audit trails of course communication exchanges with students, therefore those types of interactions were the focus of much of the research attention. However, tracking student-content interaction was not possible until the learning management systems' data retrieval mechanisms were optimized (Murray, Perez, Geist, & Hedrick, 2012). Most, if not all, of the learning management systems now routinely track student activity resulting in available data but it is not always in a usable form. One of the earliest studies using log files from a learning management system was by Phillips, Baudains, & van Keulen, (2002) in which they collected data from usage logs to evaluate learning technology in the context in which it was used (a biology course) thus paving the way for utilizing this type of information to inform the design of instructional software environments and fostering more engagement in course material.

Nickels (2005) suggested that process measures be used to assess student activity and therefore work habits (at least for those activities that require online interaction) in order to identify student behaviors that imply learning. Process measures capture data describing physically observable actions. The login system of the LMS links the behavior of individual users in the server log, and permits detailed measures of behavior. The method is not without problems. For instance, when tracking how much time a student spends on a particular page, there is no way of verifying that the student is interacting with the content or has walked away from the computer if it is a text based page. This study will not use time logs for that reason. However, using tracking data to manage learning led to the field of learning analytics.

Analytics in Education

Analytics marries large data sets, statistical techniques, and predictive modeling

(Campbell, DeBlois, & Oblinger, 2007) and has been used in the business world for a number of years (Goldstein & Katz, 2005). Anthony Picciano (2012) explains that the generic definition of analytics is “the science of examining data to draw conclusions, and when used in decision making, to present paths or courses of action” (p. 12). Analytics in higher education was initially used for administrative purposes, yet as the methods were better understood the field expanded and analytics now describes the use of big data in all aspects of higher education; administrative, enrollment management, student academic success and outcomes, as well as that data from the learning management systems, and is referred to as academic analytics (Campbell et al., 2007).

The 2012 EDUCAUSE report on Analytics in Higher Education stated that “Analytics is an interest or a major priority for most colleges and universities, however the scope of usage is limited” (p. 25), and most of the responding institutions indicated that they were in the data-collection or data-monitoring stages and had not moved to the prediction or decision- making stages. John Campbell at Purdue developed one of the first academic analytics intervention systems, called *Signals*. *Signals* draws upon data from all areas of the university’s data bases for its prediction algorithms and is an excellent example of academic analytics in action.

A more applied and focused subset of academic analytics is learning analytics, in which the data is collected from the students’ interactions in the learning management systems to make pedagogically informed decisions and evaluations (Lockyer & Dawson, 2011). Real-time data in the learning environment is collected, processed and decisions made to identify students at risk. Learning analytics is in its infancy and the 2014 NMC Horizon Report states that it will be two or three more years before it is globally adopted in higher education (Adams et al., 2014). One of the reasons for this delay is that while it is possible for learning management systems to track activity, the tracking features are inadequate and the reports generated are insufficient to help instructors determine student progress and needs (Zhang, Almeroth, Knight, Bulger, & Mayer, 2007). Thus the application of learning analytics to that data can fill the gap and provide an assessment tool for engagement in the online classroom.

Related Research

There are several other initiatives to develop systems that collect real-time data (a necessity for interventions) and predict when students are at risk. A small number of institutions have implemented intervention programs to increase retention and promote student academic success in online learning. Each of the systems is unique and nearly all are at universities. As previously stated, Purdue's *Signals* (built by John Campbell) was one of the first activated systems, and is still being piloted, however it is an intervention program that is more of an academic analytics program in that it pulls data from all areas of the university's data bases for its prediction algorithms. However, Macfadyen & Dawson (2010) have criticized *Signals* for being a "one size fits all" model explaining that it is important to recognize that there may be significant differences at the course level (*Signals* currently treats all sections of a course as a single unit) due to different instructors' pedagogical competency and suggest that predictive models should be built at the individual course level. Macfadyen & Dawson (2010) created a model that looked at the actual activity within the online course and developed a predictive model using logistic regression that could correctly predict 81% of students who received a failing grade. This limited model examined activity in one class (biology) over several semesters, incorporating only 118 student logs, providing the groundwork for further studies.

In one of the few community college applications of analytics for predicting success, Rio Solada Community College, part of the Maricopa Community College System, is piloting a learning analytics application they developed called PACE (Grush, 2011). Michael Cottam, the associate dean for instructional design, reported:

> "… after crunching data from tens of thousands of students, we found that there are three main predictors of success: the frequency of a student logging into a course; site engagement--whether they read or engage with the course materials online and do practice exercises and so forth; and how many points they are getting on their assignments. All that may sound simple, but the statistics we encounter are anything but simple. And we've found that, overwhelmingly, these three factors do act as predictors of success (Grush, 2011, p.1)".

While this is a formidable accomplishment, Rio Solado uses a learning management system that they built themselves, possibly limiting the scalability of their intervention program without a major retrofit specific to the learning management systems being used at other institutions.

Morris et al., (2005) at the University Systems of Georgia built a complex model that predicts student success and evaluates student engagement through empirical analysis of learning behavior, tracking, time-stamping, and background information. This model has been used to predict student success in the online environment for core courses only, courses are developed by a content expert (faculty) and an instructional designer which are then for the most part are taught by contingent faculty.

The University of Maryland has developed a tool called "Check My Activity" which uses log information data and generates a report for students to use to compare their performance against others in the class, with the hope that the feedback to the students could motivate them to increase their engagement and promote success in the online class (Fritz, 2011). New techniques are emerging, but until a unified approach is used and organizations work together, it will still be sometime in the future before the benefits of learning analytics will be reaped by higher education, perhaps even longer than the 2014 Horizon Report prediction of three-to-five years (Adams, Estrada, Freeman, & Johnson, 2014).

Behavioral Factors

A student entering an online class brings a number of behavioral factors into the classroom that affect the interactions with the course content, instructor and other students. The internal student-level factors that influence student behavior in an online course will be examined from the perspective of motivational theory.

Engagement. Kuh (2003, p. 25) defined engagement as "The time and energy that students devote to educationally sound activities inside and outside the classroom, and the policies and practices that institutions use to induce students to take part in these activities." Engagement in an online course is seen as interaction with content, with other students and with the instructor (Abrami et al., 2011; Cho, 2011; Moore, 1989; Rovai, 2002b). "Student engagement is important because it makes learning possible, directly causes positive student outcomes, fully mediates and explains the motivation-to-outcomes relation, and is directly proportional to subsequent motivation toward a repeated activity" (Reeve, 2012). Dawes and Larson (2011) explain that when one becomes psychologically engaged they are motivated in such a way that their attention is absorbed in the tasks and challenges of an activity, and refer to this state of absorption as flow. Interest theory notes that for engagement to occur, the activities or tasks must be personally meaningful (Dawes & Larson, 2011; Hidi, 2000; Park & Choi, 2009). "Students who actively participated in learning interactions, especially with teachers and contents, were more likely to complete and retain in online courses" (Lee & Choi, 2011, p. 609). The instructor must create a class that is interesting, engaging and can have personal meaning to the students, as well as create a community of learners in the online classroom. Xie, DeBacker, and Ferguson (2006) used self-determination theory as a framework to investigate the effect of online discussion in an online course on motivation of students and found that an increase in motivation occurred as a result of feedback from instructors and the discussion topics being made more practical and applicable to the real world.

Persistence. For students to succeed in an online course, they must engage in the course, persist, and successfully complete the course. Hart (2012) examined persistence in online courses and came up with a list of factors that contributed to persistence. Since there are multiple definitions for persistence, she used the following definition: "persistence is the ability to complete an online course despite obstacles or adverse circumstances" (p. 30). Some of the factors identified that promote online persistence leading to success in online learning include: communication with the instructor, satisfaction with the course, relevance of the course, self-efficacy, and social connectedness or presence. Street (2010) conducted a similar study and noted, "internal factors of self-efficacy, self-determination, autonomy, and time management along with external factors of family, organizational, and technical support were found to be significant" (p.1). Ryan and Deci's Self-Determination Theory (SDT) (2000a), interest development and the construct of the self-regulated learner can be used to explain the roles of these internal factors.

Self-Determination Theory. Self-determination theory (SDT) is a theory of motivation that describes the continuum of intrinsic and extrinsic motivation and its relationship to self-regulation (Ryan & Deci, 2000a,b). Self-determination theory states that motivation and well-being are determined by the extent to which three basic needs are satisfied: the need for autonomy (i.e. the extent to which a learner feels in control), the need for relatedness (i.e. the extent to which a learner feels included), and the need for competency (i.e. the extent to which a learner feels competent with respect to tasks and learning activities).

This motivational theory provides insight into online student learning in that it brings the contribution of the learning environment into the individual's experience and identifies the three critical psychological needs of autonomy, competence, and relatedness necessary to foster interest (Chen & Jang, 2010). With interest comes engagement in activities thus enhancing performance and persistence. "Self-determination is enhanced where supportive social-contextual conditions exist to promote feelings of competence or self-efficacy" (Zepke & Leach, 2010, p. 171). "Self-determination theory is unique in that it emphasizes the instructional task of vitalizing students' inner motivational resources as a key step in facilitating high-quality engagement" (Reeve, 2012, p. 152). Self-determination has been shown to be critical for success in online learning settings.

Autonomy. Autonomy is forecaster of success in an online course and can be used as a predictor of online achievement (Capra, 2011; Lim & Kim, 2003; Chen, Lambert, & Guidry, 2010; Rienties et al., 2012). Autonomy is defined as a the extent to which a learner feels in control, and as stated above is one of the basic human needs according to SDT theory (Deci & Ryan 1985, 2002). SDT states that if environmental conditions are such that they support an individual's autonomy then more autonomous forms of motivation will be promoted (Ryan & Deci, 2000a, b). In distance education, the amount of autonomy is a critical variable in reducing transactional distance and thus promoting learning (Moore, 1989). The online environment is ideal for providing an autonomous learning experience for students. The students decide when to log on and when to complete assignments. This is a key factor in self-directed learning. The perceived autonomy is the most significant factor promoting students' motivation and time on task and thus engagement (Chen, et al., 2010). However, autonomy support and appropriate scaffolding are necessary to promote success in online learning for not only novice learners but to expert learners as well (Rienties et al., 2012). Clear navigation through the course is critical. Reushle, et al. (1999) note that when navigation is not clear then students can become "disoriented" in the course site and experience excessive cognitive load which interferes with learning. Effective online learning is the result of a well-planned instructional design effort along with high quality instruction that work together to meet pedagogical needs (Murray et al., 2012; Rovai, 2003).

Self-regulation. Self-regulation is a motivational construct and falls into the category of individuals' techniques and strategies as defined by Pintrich and De Groot (1990, p. 17). Self-regulation is an active constructivist process whereby learners set goals for their learning and monitor, regulate and control their cognition (Nicol & Macfarlane-Dick, 2006, p 202). In the classroom it is described as self-regulated learning and has been shown to be a predictor of academic success (Boekarts, 1997; Miltiadou & Savenye, 2003; Pintrich & DeGroot, 1990; Rosen, et al, 2010). Pintrich and Zusho (2002) provide the following working definition of self-regulation: "Self-regulated learning is an active constructive process whereby learners set goals for their learning and monitor, regulate, and control their cognition, motivation, and behavior, guided and constrained by their goals and the contextual features of the environment" (p. 64). The self-regulated learner is one who is internally motivated. "Self-regulation is not a mental ability or an academic performance skill; rather it is the self-directive process by which learners transform their mental abilities into academic skills" (Zimmerman, 2002, p. 65). "Each self-regulated process or belief, such as goal setting, strategy use, and self-evaluation, can be learned from instruction and modeling by parents, teachers, coaches, and peers" (Zimmerman, 2002, p.

69). Self-regulation of learning is process driven and it is possible to teach self-regulatory processes resulting in increases in students' motivation and achievement (Shea & Bidjerano, 2010; Shunk & Zimmerman, 1998). Pintrich & Zusho (2007) provide a general working definition of self-regulated learning stating that it is "an active, constructive process whereby learners set goals for their learning and then attempt to monitor, regulate, and control their cognition, motivation, and behavior, guided and constrained by their goals and the contextual features in the environment" (p. 741). Self-regulation can be taught with explicit instructions, directed reflection, and metacognitive discussions (Paris & Winograd, 2001, p. 8). A student who is a self-regulated learner is able to activate strategies that enhance their learning (McMahon & Oliver, 2011). Self-regulated learning is best examined at the classroom or course level (Pintrich & Zucho, 2007), be it a traditional class or an online course. A well-designed course promotes the learning of processes and strategies that promote self-regulated learning in the very structure of the online course (Dabbagh & Kisantas, 2013; Hense & Mandl, 2012; Shea & Bidjerano, 2009).

Relating activity to behavior

The learning management system provides a method to track all activities in the system. This allows examination of student activity (and instructor activity) and permits drawing conclusions about behavior, activity, performance, learning, motivation, engagement, retention, and student academic success. The following is short review of recent studies utilizing these activity monitoring and tracking features.

Tracking data generated from student activity in an LMS can be used as an indicator of learning performance (Dawson & McWilliam, 2008). Fritz (2011) built upon that concept to explore the relationship between student performance and activity. A direct correlation was found between online course activity and performance in a study that investigated the relationships between available tools used in an online course, motivation, participation, and performance on a final exam (Giesbers, Rienties, Tempelaar & Gijselaers, 2012). The "Check my Progress" tool at the University of Maryland empowers students to take control of their own learning and evaluate their progress by building student awareness of the connection between online activity and student performance as well as promoting autonomous behavior (Fritz, 2011) thus increasing their chances for student academic success. Whenever tools can be provided that encourage student autonomy and self-regulated learning everyone benefits.

Interaction Treatments and Course Design

"In the past ten years, Web access, the nature of the Web, and contexts for learning have been transformed, along with the emergence of desired technological competencies for learners, teachers, and administrators" (Greenhow, Robelia, & Hughes, 2009, p. 246). These changes allow greater affordances and a broad range of collaborative and cooperative opportunities for online learning. The Web changed allowing more user contributions and interactions, the nature of the Web itself changed. The first generation (Web 1.0) was viewed as an educational and communication resource (Dunlap & Lowenthal, 2011; Greenhow et al., 2009). The design promoted authenticated knowledge in which experts presented findings and conclusions. It was a one-way transfer of information for most users, with user contributions primarily limited to text-based forums and archived listserves (Dede, 2008; Greenhow et al., 2009).

Many see the next generation, often referred to as Web 2.0, as a transformation from the predominantly read-only Web 1.0 into a read-and-write Web 2.0 facilitating participatory, collaborative and distributed practices (An & Williams, 2010; Greenhow et al., 2009). User contributions exploded into the web as a result with Web 2.0 users sharing information in a number of media rich formats: sharing personal information in social networks such as Facebook and LinkedIn; media sharing in sites such as YouTube and Flickr, and social bookmarking, using sites like Delicious and CitULike where they share with others sites that they find informative. They can also participate in collaborative knowledge development through wikis (Wikipedia, PBWorks, Wikispaces); share creative works, with podcasts, videocasts, blogs and, microblogs (Twitter, Blogger). These are only a sampling of the user based sharing sites allowing the two-way transmission of information via the Web (Conole & Alevizou, 2010; Dunlap & Lowenthal, 2010; Greenhow et al., 2009; McGreal & Elliot, 2008). In a study on Web 2.0 tool usage, most participants reported that these tools helped build a sense of community, increased interaction and communication among the instructor, students, and other people (An & Williams, 2012). No longer was the web limited to the delivery of information, but advances of Web 2.0 capabilities made it easy for anyone to upload information and contribute to online content thus allowing individuals to share information as well as receive it.

Summary

A recent search resulted in only one study specifically focused on interactions in the community college online setting in which Aragon & Johnson (2008) attempted to determine if demographics affected whether students completed online courses or not. They did not find any statistical significance between the two groups based on demographics. One reason this study focused primarily on online activity is the student background data in community college settings is quite different from that of universities and four-year schools. Additionally, many of the other studies used ACT and/or SAT scores in their models as well as previous success in classes, which are not often available to community college data sets (Campbell, 2007; Morris & Finnegan, 2009a; Morris et al., 2005; Dawson & McWilliam, 2008; Murray et al., 2012). Most community colleges have an open enrollment policy, and students typically do not take the ACT or SAT exams so much of the educational data used in complex university systems is unavailable in community college environments. Only first-time enrolling students are required to take the placement exams, so students who had attended college prior to enrollment would not have this data. Many students drop in for one class (often because they can't get into the class on their home campus) and then move on. So these variables are not often part of the mix when looking at community college data, it made sense to focus on what is available and that is the data from the tracking logs in the learning management system.

CHAPTER III: METHODOLOGY

Introduction

The primary goal of this study was to determine if interactions in the first two weeks of a fully online class were associated with student academic success, and if any of the specific interaction types increased the odds for student academic success. This information could be used to guide policy for instructors in the design of online courses, as well as identify interaction behaviors to encourage in the online environment in order to promote student academic success. This study was made possible by the affordance of learning management systems allowing all activity within the learning environment to be continuously tracked and archived, thus providing a robust data source for research.

The methodology used to test the research questions is listed in this chapter. The following is a list of the individual sections: 1) introduction, 2) research questions 3) research design including a brief explanation of the statistical method, 4) participants, 5) data collection and processing, 6) data analysis, and 7) summary.

Research Questions

RQ1: Is interaction activity in the first two weeks of class associated with student academic success in the online classroom?

RQ2: Do the interaction types as a set reliably distinguish between the three levels of academic success?

RQ3: Is there a relationship between each of the interaction types and student academic success?

RQ4: What is the effect of each of the interaction types on student academic success?

RQ5: Are some interaction types more important than others in promoting student academic success?

Research Design

This study employed multinomial logistic regression using Statistical Package for Social Sciences (SPSS) 20.0 to examine associations between behaviors in the online classroom and three levels of student academic success: successful completers, low score completers, and non-completers. Moore's Theory of Interaction in online learning (Moore, 1989) provided the theoretical framework for the study. In the initial hypothesized model, six predictor variables represented the three primary interaction types in an online classroom: student-student, student-content, and student-instructor. The student-student interactions variable was defined by counting the number of posts made in the discussion forum by each student. The student-content interactions were those interactions that occur when the student clicks on the various components that comprise the course curricular content. Student-instructor interactions (a more complex interaction type) were initially represented by four independent variables that constitute the various options for instructor-student interactions: the number of emails the instructor received and sent, the number of discussion posts the instructor made in the course, the number of quizzes and exams in the course, and the number of instructor graded assignments all representing some facet of instructor activity and thus student-instructor interaction in the course. Analysis of the model showed that one of the variables, the number of instructor graded assignments, did not contribute significantly to the model. Additional analysis also revealed that there were confounding issues with the assessment variable. The researcher removed both of these variables from the final model, resulting in four independent variables instead of six. The outcome variable is the categorical variable, student academic success, and is subdivided into three groups: successful completers, low score completers, and non-completers.

Multinomial logistic regression was selected for the analysis because it works well for describing and testing hypotheses about relationships between a categorical outcome variable with more than two levels and multiple continuous predictor variables (Garson, 2012). It is a useful tool for educational data sets because it has less stringent requirements than other regression models. It does not require normally distributed variables, nor does it assume homoscedasticity. It also does not assume a linear relationship between the raw values of the independent variables and raw values of the dependent variable. There are a couple of assumptions that it does require. The observations must be independent and if the independent variables are continuous they must be linearly related to the logit of the dependent variable. A classification table can be generated that assesses the predictive success of the model. Since logistic regression uses maximal likelihood estimates to determine the odds of a certain event occurring instead of ordinary least squares, it can be used to determine the effect size of the independent variables on the dependent and to rank the relative importance of the independent variables. The impact of predictor variables is explained in terms of odds ratios. Odds ratios are the exponents of the parameter estimates. "The odds ratio is the increase (or decrease if the ratio is less than one) in odds of being in one outcome category when the value of the predictor increases by one unit" (Tabachnick & Fidell, 2001, p. 548).

Participants

The data used for this study consisted of a convenience sample of students enrolled in the Fall 2012 semester in entirely online, full semester (16 weeks) courses at a medium-sized community college in the Midwest. Archived data was retrieved from the Learning Management System (LMS) as well as from the Enrollment Management System. All personal identifying information was stripped from both sets of data and the two sets were merged and sorted using a generated student identifier number and SPSS version 20.0 automated procedures. Activity tracking records representing 6,447 students in 409 courses were pulled from the LMS data set. Academic services provided grades for 4,876 students enrolled in 409 courses. Data for courses that were not fully online and full semester courses were removed from both data sets. An assumption of logistic regression is that no individual can be in the data set more than once, yet many students take more than one online class at a time. A randomized process was developed and used to select one course per student thus ensuring that there was no duplication of cases in the data for this study.

Tracking records representing 3345 cases in 240 courses were still in the data set, but that number was reduced to 1,702 students in 200 courses after removing cases for students enrolled in more than one course. A complete list of the departments represented and how many cases were from each department is presented in Appendix A. The 1,702 unique students represented 22.8% of the overall student population.

A total of 1,187 (69.8%) females and 514 (30.2%) males comprised the sample (one individual did not designate gender), whereas the overall enrollment at the college was 4,625 (61.78%) females and 2,832 (37.83%) males. Student ethnicity in this sample was: African American 414 (24.3), Asian 38 (2.2%), Hispanic 94 (5.5%), Other 87 (5.1%), Unknown 66 (3.9%), and White 1,003 (58.9%). This population sample aligns well with the overall enrollment demographic of ethnicity of African American 2,121 (28.33%), Asian 184 (2.46%), Hispanic 661 (8.83%), Other 690 (6.76%), Unknown 308 (4.11%), and White 3,706 (49.51%). Table 1 provides a summary of the student demographics for the study compared to the college's overall student demographics.

Table 1: Categorical Student Demographics

Ethnicity	Study Demographics	College Average
African American/Black	24.3%	28.3%
Asian	2.2%	2.5%
Hispanic	5.5%	8.8%
Other	5.1%	6.8%
Unknown	3.9%	4.1%
White	58.9%	49.5%
Gender		
Female	69.8%	61.8%
Male	30.2%	37.8%
Unknown	0%	0.4%
Full Time/Part Time		
Full Time	23.6%	35.4%
Part Time	76.3%	64.6%

Human Subjects Issues

The Review Board at the community college granted permission for the study. The Human Subjects Research Committee at the University of Kansas approved this study as well. This study consisted of a secondary analysis of data electronically retrieved from two existing data sources (course management and enrollment management). All personalizing information was removed from the data providing an anonymous data set.

Data collection and processing

The research division of the community college provided the student demographic, enrollment and grade information in an EXCEL document. This file was imported into SPSS 20.0 in preparation to be merged with the LMS data for analysis. The data set contained student data for 409 courses, many of which were not fully online, full semester courses. Courses in that file that were not classified as full semester, fully online courses were removed from the data set. Only students that were enrolled in full semester, fully online courses were used in this study.

Collecting the LMS Data. Previously archived data from the learning management system in the form of raw log data was retrieved from the LMS discussion areas, content areas, and mail areas via multiple queries. The sample used for analysis was extracted from data sets consisting of more than two million individual tracking records. Data processing included extracting student tracking activity for individuals enrolled in only fully online, full semester courses and truncating those records to restrict the data to only the activity in the first two weeks of classes.

The structural framework of types of interactions in an online environment guided the collection of data from the data set provide by online services. Data identified by the literature as representative of the three interaction types was identified and retrieved. This study focused on information found in the tool areas most commonly used by online instructors at the community college in an effort to normalize data. It was hoped that data from the announcement tool could be used, but that data was unreliable and was disregarded as a possible data set. While there are other tools available in the LMS used by some instructors, this researcher felt that using the data from infrequently used areas would result in the generation of too much missing or incomplete data. Including only the most frequently used tools would provide the most consistent data in order to construct the independent variables.

The data were collected from archived records from the LMS in four batches. The first query collected the discussion activity logs for students and instructors. The second retrieved the content activity logs for the students. The third contained the mail activity counts for the instructors and the fourth retrieved information regarding the number of and types of content items in each course.

Processing the data. The raw LMS data were processed both at the individual level and course level. Data for the entire semester was collected and then the data from the first two weeks of the semester was extracted and used. The data was cleaned, recoded when necessary, and the discussion data was segregated into student data and instructor data.

As mentioned above, the initial LMS data set contained 3395 students many of whom were enrolled in more than one course. The assumption of independent observations is one requirement of logistic regression. This means that that an individual can be represented in the data set only once. In order to meet this requirement, a random selection method was used to allow each student to be included in the data set only once. All remaining cases for any student taking more than one fully online course were removed from the data set. This resulted in the reduced data set containing 2062 primary cases. Courses having no discussion component presented null data in the data set. Therefore, any course having no discussion component was removed from the data. This resulted in a final tally of 1702 students in 200 courses in the working data set. All data processing was performed using SPSS data management functions.

Removal of extraneous data. Data from the LMS initially included data from students in non-fully online, not full semester classes, and classes not counting towards graduation were removed from the original data sets. In addition, any data that did not have a corresponding grade in the academic records data were also excluded.

Constructing the Independent Interaction Variables

The two student level interaction variables were constructed as follows: the first variable, student-content interactions, was constructed by counting each time a student clicked on a link to a content item and opened that page in the course. Every time a student opened a page it counted as a “hit”, even if the student had visited that page before. The “hits” were totaled for each student in the first two weeks of the active data set. The second variable, student-student interactions, was constructed by calculating the total number of postings each student made in the two-week period in the discussion forums. In order to capture the student-instructor construct, initially four instructor level variables were created and used in the initial model; the number of instructor discussion posts, the number of instructor emails (sent and received), the number of instructor graded assignments, and the number of quizzes/exams in the course. The variable representing the number of instructor-graded assignments was removed from the final analysis since it was found to be not significant for the model. In addition the variable representing the number of exams/quizzes was also removed due to concerns of reliability of the measure.

A unique student identifier was created and used to match each student’s tracking records with the grade earned.

Standardization of Variables

The student-student and student-content variables were standardized within individual course sections. The counts were transformed into Z-scores based upon the mean activity for each course section. The resulting transformed scores had a mean of zero and a standard deviation of one.

The student-instructor data were standardized across all course sections, thus normalizing the data and possibly removing the inter-rater bias that might occur since the different courses were taught by a number of instructors. The counts were transformed into Z-scores based upon the overall mean activity for each of the student-instructor variables.

The data from the enrollment management system was provided by the CRDC staff and was stripped of personal identifiers and was merged with the LMS data. The data was processed and missing data was addressed. Missing data resulted in the removal of classes that did not include discussions as part of the course. In addition the data was checked for duplicate cases (students enrolled in more than one online class) and data from only one class was randomly selected for use for each individual in the data set.

Missing Data

The files were checked for missing data. Missing data was addressed with case-wise deletion. There were not many instances of missing data so it was felt that deleting those data should not have any effect on the results.

Interaction Variables (Independent Variables)

The learning management system collects tracking information and generates footprints of all activity within each course. Variables were constructed from the data retrieved to represent the types of interactions in distance education based on Moore's Interaction Types (Moore, 1989). (See Tables 2 and 3 for the complete list of independent variables.) The interaction independent predictor variables used in the final model were constructed as follows:

1) Student-content interaction was tracked within the LMS log files, in that each time the student "opened" a content file, it constituted a "hit" and was counted. It is the total interaction with the content rather than initial contact that is important for this analysis; therefore, each time the student opened a page it counted as a "hit" with multiple visits resulting in multiple counts. It is the total interaction with the content that is being tabulated. The standardized value for the individual student content counts for each course was calculated in order to normalize the value across courses.

2) Student-student interaction occurs when students interact with one another in the

discussion area of the course. The activity was tracked by counting the number of posts that the student made in the course and the standardized value was calculated for each student in each course in order to normalize the values across courses.

The student-instructor interaction was difficult to quantify therefore four predictors based upon the literature foundation were chosen to capture this interaction in the initial model. The four student-instructor predictors were constructed as follows.

3) Student-instructor interaction (Discussion) was calculated by counting the discussion posts made by the instructor (in the discussion area of the LMS) and was standardized across all courses included in the study.

4) Student-instructor interaction (Mail) was constructed by standardizing the total number of instructor emails both sent and received (within the LMS) for all courses.

Table 2: Student-based Predictor (Interaction) Variables

VARIABLE	DESCRIPTION	TYPE	CLASS
Student-Content Interaction	The total number of content items opened by the student.	discrete	interval
Student-Student Interaction	The total number of discussion posts made by the student.	discrete	interval

Table 3: Instructor-based Predictor (Interaction) Variables

VARIABLE	DESCRIPTION	TYPE	CLASS
Student-Instructor Interaction: Discussion	The total number of discussion posts by the instructor.	discrete	interval
Student-Instructor Interaction: Mail	The total number of emails (sent and received) by the instructor.	discrete	interval
Student-Instructor Interaction: Graded Assignments	The total number of instructor graded assignments.	discrete	interval

The mean counts for the interaction variables for each of the outcome groups and for the study as a whole are shown in Table 4 below.

Table 4: Mean Counts for the Predictor Variables for Each Outcome Group

Interaction Variables	Successful Completers (A,B,C)	Low Score Completers (D,F)	Non-completers (W, WA, I)	All
Student-student	5.54	3.15	2.26	4.76
Student-content	60.03	43.75	32.81	54.05
Instructor posts	11.07	10.79	7.47	10.52
Instructor email	198.21	222.45	243.41	151.19

The Dependent Variable

The purpose of this study was primarily to determine if the tracked class interactions could correctly predict student academic success grouping and secondarily to contribute to developing a model to identify students who could benefit from early intervention in order to promote student academic success. The dependent variable, Student Academic Success, is a categorical variable (Table 5) and is broken down into three groups:

1) Successful Completers are students who earned A, B, or C grades

2) Low score Completers are students who earned D or F grades

3) Non-completers are students who earned W, WA, or I grades.

Table 5: Outcome (Dependent) Variable Groups

GROUP	DESCRIPTION	TYPE	CLASS
Successful Completers	Received a grade of A, B, or C in course	discrete	categorical
Low score Completers	Received a grade of D or F	discrete	categorical
Non-completers	Received a grade of W, WA, or I	discrete	categorical

Data Analysis

Multinomial logistic regression. A four-predictor multinomial logistic model was fitted to the data to test the research hypothesis regarding the relationship between online course interaction (with other students, instructor, and content) in the first two weeks of class and successful completion of the course. Multinomial logistic regression was selected for this analysis because the dependent variable is non-metric and has three categories: successful completers (A,B,C), low score completers (D,F), and non-completers (W, WA, I)(shown in Table 7). The five independent variables were all based on counts of activity and were mean standardized values (z-scores) for actual counts of actions or items. The goal of the study was to see if the five independent variables could accurately predict student academic success category placement.

Multinomial logistic regression was conducted in SPSS 20.0 to determine if the predictor variables (representing online behavior) could classify each case into the correct group. In addition, the output was examined to determine if any of the predictor variables was a stronger predictor of success than any of the others. Multinomial logistic regression models the relationships in by comparing two of the categories defined by the dependent variable with the third group identified as the reference category. In this process SPSS treats the analysis as two binary logistic regressions using the same reference category for each. Typically the reference category is the most highly populated dependent variable category.

In this study, the successful completers group was the largest category with 72.7% of the students in the study falling in that category. This resulted in two comparisons for the study: non-completers versus successful completers and low score completers versus successful completers.

Testing for outliers. Outliers were detected by running two binary logistic regressions, using case selection to compare group 1 to group 3 and group 2 to group 3, generating studentized residuals. Studentized residuals were generated in SPSS by dividing a residual by an estimate of its standard deviation. This identified a list of cases with studentized residuals greater than +/- 2.0. The next step was to test the multinomial solution with these class excluded and compared with the full model with all cases. An accuracy rate of less than 2% more accurate is required to interpret the model including all cases. An accuracy rate of greater than 2% more accurate indicates that the outliers should be removed from the model (Schwab, 2002). The accuracy rate was 2.2% and the outliers were removed.

External and internal validity. External validity examines if the study can be generalized to other populations. This study is limited to one semester's data set (discussed in limitations), but there are two methods that can be used to validate the study.

Bootstrapping is a fairly recent statistical method of resampling that is useful when one is limited to one data set. The data set was treated as a population and the data was resampled with replacement one thousand times, thus simulating a new data set.

Cross-validation was performed by randomly dividing the data into two subsets: a training sample and a holdout sample. This study used a 75/25 cross-validation strategy and the two subsets were: a training sample containing 75% of the cases (randomly selected) and a holdout sample containing the remaining 25% of the cases. The training sample was used to derive the multinomial logistic regression model. The holdout sample was classified using the coefficients for the training model. SPSS does not classify cases that are not included in the training sample so the classification table for the hold out sample was calculated independently. The classification accuracy for the holdout sample was used to estimate how well the model based on the training sample will perform for the population represented by the data set. According to Schwab (2002) it is expected that the classification accuracy for the validation sample will be lower than the classification for the training sample. However the difference, identified as shrinkage, should be no larger than 2%. To further validate the model, in addition to satisfying the classification accuracy, this study required that the significance of the overall relationship and the relationships with individual predictors for the training sample matched the significance result for the model using the full cleaned data set after removing the outliers

Limitations

This study investigated data collected from only one medium sized community college in the Midwest during one semester, which may skew the samples of students due to the specific population and it also limits the breadth and generalizability of the study. Future studies using the same procedure on at other types of higher education institutions in different regions could be used to improve the study.

This study did not control for gender, ethnicity, academic ability, discipline, or course design. The study was an attempt at stripping down the data to only what happened in the online courses and purposefully did not control for the aforementioned fields often because academic ability which is often assessed in higher education studies using ACT scores and/or SAT scores. Most community college students either don't take those exams or are returning students and their scores are out of date. The other factors were not considered because the outcome was based upon what the student was actually doing in the class regardless of those factors.

Reliability

Reliability is the consistency of measurement within a study. For this study, the reliability for the course management, demographic, and course grade were addressed.

All course management data is found within a common central server. Due to the nature of the system, all data collected by the course management system was automatically generated and was not changed by the researcher or the system users. Great care was taken when combining the information to create the predictor variables. Verification for reliability included examining the data for the researcher's classes and comparing that data with the information contained in the actual courses.

The demographic data was provided by the college's Research Department and is considered validated by that unit. Great care was taken when merging the demographic and course activity data to ensure reliability.

The grade data was archived from the enrollment management system and is reliable in that respect, however the courses are taught by different instructors using different grading scales, but it is hoped that the overall degree of inter-rater reliability will be countered by the large data sets and will also be softened by using two categories for the criterion variable -

successful student and low score student rather than using course grades which can vary greatly from instructor to instructor and course section to course section.

In an attempt to focus the examination on interaction data only, no controls for types of courses or student characteristics were used. This will be addressed further in the limitations section.

Assumptions for Multinomial Logistic Regression

Most statistical tests rely upon certain assumptions about the variables used in the analysis. Therefore, the following section focuses on the assumptions for multinomial logistic regression.

Assumption 1. Independent observations

Meeting assumption 1 of multinomial logistic regression-"the observations must be independent" was met by the inclusion of only one data sample per student. This was accomplished by the randomization and selection of students in multiple classes as mentioned above.

Assumption 2. Linearity of independent variables with logit of the dependent

In order to meet the second assumption, the study examined the independent variables supplied by the course management system to see if they were correlated with student academic success in undergraduate courses. Bivariate correlational techniques were used to analyze the degree of the relationship between two variables. All independent variables were compared with the dependent variable to determine the strength of the correlation. The data met the requirement of that the independent variables be linearly related to the logit of the dependent variable

Multicollinearity. Multicollinearity is when the independent variables are highly correlated. This can be detected in the multinomial regression solution by examining the standard errors for the b coefficients. A standard error larger than 2.0 indicates numerical problems (Schwab, 2002). None of the independent variables in this analysis had a standard error larger than 2.0.

Significance Tests for Logistic Regression

Model fitting information. Multinomial logistic regression was used to analyze the relationships between the non-metric dependent variable (student academic success) and the metric independent variables. The following significance tests will be applied to determine if the model adequately describes the relationships between the dependent variable and the independent variables. Analysis of the model will include the overall test of the relationship, testing the strength of the multinomial logistic regression relationship, evaluating the usefulness of the model by computing the by chance accuracy and comparing that to the accuracy rates of the model. The various tests include: Goodness of Fit, Log likelihood ratio, chi square test, pseudo r-square, parameter estimates, a classification table and a classification accuracy test.

Goodness of fit. The overall test of relationship between the independent variables and the groups defined by the dependent variable is based on the reduction in the likelihood values for a model that does not contain any independent variables and the model that contains the independent variables. The goodness of fit table is identified by SPSS as the "Model Fitting Table" and produces a chi-square value where the presence of a relationship between the dependent variable and combination of independent variables can be identified. A significance level of less than 0.05 is an acceptable proof that the model adequately represents the relationship between the dependent variable and the combination of the independent variables.

Strength of the multinomial logistic regression ratio. Multinomial logistic regression does compute correlation measures to estimate the strength of the relationship and these values are identified as pseudo R square measures. SPSS reports the Nagelkerke's R*2*, Cox and Snell's

R*2* and McFadden's R*2*, however it is important to note that they are not interpreted in the same manner as in linear regression. There is considerable debate regarding the usefulness of these values and while they are reported they are not considered to represent variance explained, but rather are used to are interpreted the relationships as weak, moderate, or strong (Tabachnick & Fidell, 2001).

A more useful measure for estimating effect size was recommended by Schwab (2002) in which classification accuracy was assessed, comparing predicted group membership based on the logistic model to the actual, known group membership. The benchmark recommended to use was that of a 25% improvement of accuracy over the rate of accuracy achievable by chance alone. Chance accuracy is the accuracy of the prediction of group membership some percent of the time even if the independent variables had no relationship to the groups defined by the dependent variable. This study applied a method to estimate the chance accuracy rate by using the proportional by chance accuracy rate which was computed by calculating the proportion of cases for each group based on the number of cases in each group in the "Case Processing Summary" table, and then squaring and summing the proportion of cases in each group. This was converted to a percentage and multiplied by 1.25% to derive the proportion by chance accuracy criterion. This value was compared to the classification accuracy rate provided in the "Classification" table in the SPSS output and when the classification accuracy rate is greater than the proportion by chance accuracy rate then the model is classified as "useful".

Relationship between individual independent variables and the dependent variables. Two types of tests were applied to the individual independent variables. The first was the likelihood ratio test that evaluated the overall relationship between the individual independent variables and the dependent variable. A significance value of less than 0.05 supports the existence of a relationship. The results from this statistic will be explained in the results section, however the majority of the independent variables did show an adequate significance level. The second test performed was the Wald test to evaluate whether or not the independent variable was statistically significant in differentiating between the two groups in each of the binary logistic comparisons performed in the multinomial logistic procedure. It is important to note that an independent variable might have an overall relationship to the dependent variable, but it may or may not be statistically significant in differentiating between the two pairs of groups defined by the dependent variable. The first step was to identify if a significance existed supporting the relationship between the independent variable and the dependent variable using the "Likelihood Ratio Tests" output and then determining if the independent variable could significantly distinguish between the two categories of the dependent variable by examining the "Parameter Estimates".

It is possible that there could be different interpretations for the independent variable depending upon which group comparisons were made. In this study, the reference group was "Successful Completers" since that was the largest group.

Model modification. After running the analysis, it was determined that the instructor assignment variable was not significant, nor could it be supported by the literature therefore it was removed from the model. Further examination of the assessment variable showed that it was not reliable and was also removed from the model. The final reduced model contained four independent variables instead of six, with two variables representing student-instructor interaction instead of four.

Summary

The researcher conducted a forced entry multinomial logistic regression to determine if the predictors (listed in table 3.1) were related to successful course completion and if any were more potent predictors of student academic success than the others. The researcher extracted the first two weeks of data from the tracking logs and that data was used to generate the predictor variables for analysis (see above). The study used multinomial logistical regression because the dependent variable was categorical with three levels and there were multiple independent variables. Multinomial logistical regression determines the impact of multiple independent variables presented simultaneously to predict membership in each of the three dependent variable categories. A classification table was generated to look at the proportion of cases that were classified correctly and a classification analysis was performed. Outliers were removed, bootstrapping as well as a 75% cross-validation test were performed.

Hypotheses

Hypothesis 1– The amount of online course interactions (with course content, instructor, and other students) in the first two weeks of class is related to successful completion.

Hypothesis 2 – Student-content interaction promotes successful completion in online classes.

Hypothesis 3 – Student-student interaction promotes successful completion in online classes.

Hypothesis 4 – Student-instructor interactions promote successful completion in online classes.

Hypothesis 5– Some types of interactions are better promoters of success than others.

CHAPTER IV: RESEARCH RESULTS

The results of the data analysis for each research question will be presented in this chapter. These results are organized into several sections. The first section is an introduction to the study. The second section presents an examination of the characteristics of the study population. The third section describes the results of the first research question correlating the interaction types with the levels of student academic success. The fourth section validates the model. The fifth section examines each of the independent variables and their individual contribution to the model as well as identifying the strength of effect for the independent variables and distinguishing between behavior patterns of the two outcome categories of low score students. The sixth section summarizes the results.

Introduction

The purpose of this study was to determine if student tracking data representing interaction in the online classroom could be used to identify future student academic success in online classes, and if possible, were any of the types of interaction more influential than others in predicting placement. A hypothesized model was generated based upon previous research.

The hypotheses guiding this research are:

Hypothesis 1: The amount of online course interactions (with course content, instructor, and other students) in the first two weeks of class is related to successful completion.

*Hypothesis 2:*Student-content interaction promotes successful completion in online classes.

*Hypothesis 3:*Student-student interaction promotes successful completion in online classes.

Hypothesis 4: Student-instructor interaction promotes successful completion in online classes.

Hypothesis 5: Some types of interactions are better promoters of success than others.

A forced entry multinomial logistic regression analysis was performed on archived data from 1,700 students in one semester of 200 full semester length classes using SPSS 20.0. A classification table was produced to build a classification model as a means to assess if interaction in the first two weeks of class could be used to predict correct placement in the dependent groups and if so, to determine if any of the interaction types was more effective than the others.

The initial hypothesized model included six interaction independent variables: one student-student interaction variable (student posts to the discussion forums), one student-content interaction variable (total number of clicks on content items), and four student-instructor interaction variables (number of instructor email counts, number of instructor discussion counts, number of instructor graded assignments, and number of quizzes and/or exams in the course). The model was modified when analysis showed that one of the variables (number of instructor graded assignments) did not contribute significantly to the model, the variable was removed from the model, thus resulting in five independent variables instead of six. In addition, it was discovered that the exams/quizzes count interaction variable had some reliability issues and it was removed from the model. These actions reduced the number of independent interaction variables to four: two student interaction variables -- student-student interaction (number of posts to the discussion forums) and student-content interaction (total number of clicks on content items) -- and two student-instructor interaction variables –the number of instructor discussion posts and the quantity of instructor email.

Construction of the dependent variable, student academic success, was guided by criteria that the college had in place for achievement tracking (successful students earned a grade of A, B, or C and unsuccessful students earned a grade of D, F, I, W, or WA). However, for this study, the researcher felt that that there might be a difference in students who do not complete a class and those who complete a class but earn a D or an F. The unsuccessful students were divided into two groups: low score completers and non-completers. Dividing the dependent variable into three groups provided a means to identify possible behavior differences between students who do not complete a course and those who complete the course but with grades that either do not transfer or do not count towards a degree. Therefore, for this study, the dependent variable, student academic success, was constructed to have three levels: successful completers, low score completers, and non-completers.

Descriptive Statistics

Characteristics of the independent variables. The independent variables were constructed using counts from the archived data and then modified so that the independence of assumption rule is met and that each student can only appear once in the data set. The process for doing this was described in the methods section. The descriptive statistics for the predictive interaction variables used in the study are shown in Table 6.

Table 6: Descriptive Statistics for Predictor Variables

Predictor	N	Minimum	Maximum	Mean	Std. Dev.
Student-Student Interactions	1702	0	41	4.76	4.571
Student-Content Interactions	1702	0	393	54.05	48.863
Instructor Post Interactions	1702	0	124	10.52	14.736
Instructor email Interactions	1702	8	1131	207.79	151.184

Z-scores were computed for each of the variables to be used in the analysis in order to accommodate the vast range in data between the different variables.

Characteristics of the dependent variable. Students receiving final grades of A, B, or C were assigned to the successful completer group (72.7%), those with a grade of D or F were in the low score completer group (13.1%), and those students earning W, WA, or I were classified as non-completers (14.2%). Figure 7 displays the composition of the dependent variable groups.

Table 7: Descriptive Statistics for the Dependent Variable

Outcome Group	N	Percentage
Successful Completers (A,B,C)	1238	72.7%
Low Completers (D,F)	223	13.1%
Non-completers (W, WA, I)	241	14.2%
Total	1702	100%

Results of Testing the Research Questions

Research Question One

Is interaction activity in the first two weeks of class associated with student academic success in the online classroom?

A forced entry multinomial logistic regression analysis was performed through SPSS 20.0 to assess prediction of membership in one of three categories of outcome (successful completers, low score completers, and non-completers) initially using six interaction independent variables: student-student interaction (the number of discussion postings), student-content interaction (the number of content hits), and four student-instructor interaction variables (instructor email counts, instructor discussion posts, instructor graded assignments and the number of quizzes/exams in the class. However, upon further examination of the model, it was discovered that the independent variable representing the number of instructor graded assignments did not contribute significantly to the overall functioning of the model, and that the variable representing the number of quizzes/exams was found to have some issues with reliability, both variables were removed for the remainder of the study.

A second forced entry multinomial logistic regression analysis was performed upon the reduced model using SPSS 20.0 to assess prediction of membership in one of three categories of outcome (successful completers, low score completers, and non-completers) using four interaction independent variables: student-student interaction (the number of discussion postings), student-content interaction (the number times content items were opened), and two student-instructor interaction variables (instructor email counts and instructor discussion posts).

Effect size in logistic regression is provided by the classification table and in this study the classification table supports the predictive success of the logistic regression, showing that a 73.9% correct prediction rate exceeding the proportional by chance accuracy criteria of 70.7% by 3.2%, thus satisfying the criteria for classification accuracy (Schwab, 2000). In other words, the model is more accurate than predicting by chance alone, thus supporting research question number 1. Table 8 illustrates a classification table that documents the validity of the predicted probabilities. A classification table shows the degree to which predicted probabilities agree with actual outcomes and Hosmer and Lemeshow 2000) assert, “the classification table is most appropriate when classification is a stated goal of the analysis” (p. 60).

The values on the diagonal are the correct predictions and the values not on the diagonal are the incorrect predictions. According to Table 8 with the cutoff set to 0.5 by SPSS, the prediction for successful completers was far more accurate than for either low score completers or non-completers. This could be considered a limitation if the goal is to identify those at risk students, but could be considered a positive finding for the objective of identifying interactions that promote student academic success.

Table 8: Classification Table for Observed and Predicted Frequencies

Observed	Predicted			
	Non-completers	Low score Completers	Successful Completers	Percent Correct
Non-completers	67	2	172	27.8%
Low Score Completers	33	2	188	0.9%
Successful Completers	46	4	1188	96.0%
Overall Percentage	8.6%	0.5%	91.0%	73.9%

Research Question Two

Do the interaction types as a set reliably distinguish between the three levels of student academic success?

A test of the full model against a constant only model was statistically significant, indicating that the predictors as a set reliably distinguished between the three dependent variable groups, successful completers, low score completers, and non-completers, (chi square = 461.91, $p<.00$, $df=8$). Non-significant Deviance and Pearson scores of 1.0 and a Nagelkerke's R^2 of .302 (indicating a moderate relationship between prediction and grouping) further support the theory that the logistic model was more effective than the null model. Table 9 shows the results of the chi square test.

Table 9: Evaluation of the Model

Test	X^2	df	p
Overall Model Evaluation			
Likelihood ratio test	461.961	8	0.0000
Goodness of Fit test			
Pearson	2796.009	3310	1.000
Deviance	2111.579	3310	1.000

Note: Nagelkerke $R^2 = .302$

Pseudo R^2 in logistic regression is interpreted differently than R^2 in linear regression, it does not quantify variance explained. In logistic regression, Nagelkerke's R^2 indicates how useful the explanatory variables are in predicting the response variable and can be referred to as measures of effect size. A Nagelkerke's R^2 of .302 is identified as moderate, and thus acceptable for this model.

Research Question Three

Is there a relationship between each of the independent variables and student academic success?

As mentioned above the goodness of fit test showed that the independent variables as a whole contributed to providing a model that significantly improved the classification of the dependent variable. In order to test for the significance of each of the independent variables in the model, a likelihood ratio test was performed. The likelihood of ratio test looks at each of the independent variables and tests for the contribution of each effect to the model. All five independent variables were determined to contribute significantly to the full final model. Table 10 illustrates the likelihood ratio tests showing each of the independent interaction variables as significant contributors to the model.

Table 10: Likelihood Ratio Tests of the Predictor Variables

Effect	X^2	df	ρ
Student-student	196.412	2	.000000
Student-content	59.417	2	.000000
Instructor posts	19.595	2	.000056
Instructor email	27.599	2	.000001
Constant	1271.053	2	.000000

Note: The chi-square statistic is the difference in -2 log-likelihoods between the final model and a reduced model. The reduced model was formed by omitting an effect from the final model. The null hypothesis is that all parameters of that effect are 0.

Research Question Four

What is the effect of each of the interaction types on the level of student academic success?

Multinomial logistic regression analysis runs two binary logistic regressions in order to compare the categories of the dependent variable, one category is designated the reference category (usually the largest) and the analysis results in two equations comparing each of the other categories (called the alternative categories) against that reference category. The reference category for this study was the successful completers. Parameters with significant negative coefficients decrease the likelihood of the outcome category with respect to the reference category and parameters with positive coefficients increase the likelihood of the outcome category.

The logistic regression analysis results in the generation of a value called the odds ratio, the natural logarithm (logic) of the standardized variable B. The odds ratio serves as an effect measure of the contribution for each of the independent variables in each of the equations. The farther from one that each of the odds ratios falls for each predictor variable (interaction variable), the greater the effect that variable has on the outcome. Values less than one favor the reference category and values greater than one favor the alternate outcome category.

As shown in table 11 below, all of the predictors had a significant effect on the placement into the outcome categories of non-completers against the successful completers group. An increase of one unit for all of the predictors, except instructor email, contributed positively toward the likelihood of an individual being in the successful completers group. Increasing instructor email was associated with a decreased likelihood of being in the successful completers group with a one unit increase, holding all other variables constant.

Table 11: Variables in the Equation for Non-completers Compared to Successful Completers

Predictor	B	SE	Wald X^2 (df=1)	OR	95 % CI for OR	
					Lower	Upper
Student-student*	-1.342	.131	105.718	.261	.202	.338
Student-content*	-.854	.129	44.007	.426	.331	.548
Instructor posts*	-.422	.103	16.864	.656	.536	.802
Instructor mail*	.372	.073	25.777	1.451	1.257	1.676
Constant*	-2.333	.113	405.718			

Reference group is Successful Completers
CI, confidence interval; df, degrees of freedom; OR, odds ratio; SE, standard error
*denotes sig. <0.0005

Table 12 shows the evaluation of the analysis for low score completers versus successful completers. All of the variables except the number of instructor posts were significant contributors to the outcome. Student-student interaction was the strongest predictor of student academic success. Student-content interaction also promoted success in this comparison. The number of instructor emails was associated with a less likelihood of being in the successful completers category for both non-completers and low score completers.

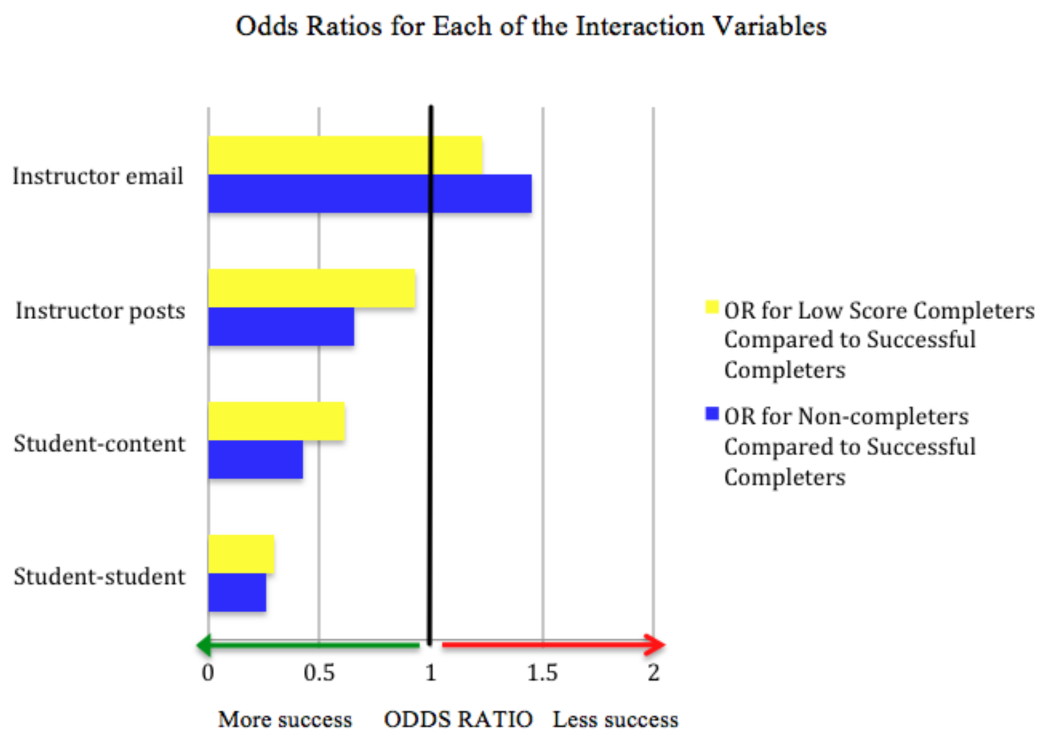
Table 12: Variables in the Equation for Low Score Completers Compared to Successful Completers

Predictor	B	SE	Wald X^2 (df=1)	OR	95 % CI for OR	
					Lower	Upper
Student-student*	-1.225	0.127	93.721	.294	.229	.377
Student-content*	-0.492	0.118	17.920	.612	.487	.768
Instructor posts ^{ns}	-0.072	0.084	0.776	.930	.792	1.093
Instructor mail*	0.245	0.080	10.102	1.277	1.098	1.486
Constant*	-2.144	0.100	444.518			

Reference group is Successful Completers
CI, confidence interval; df, degrees of freedom; OR, odds ratio; SE, standard error.
*denotes sig. < 0.0005; ns = not significant

Figure 1 illustrates the odds ratios for each of the interaction variables for the two comparisons. The odds ratio for each interaction variable is given for the comparison of Non-completers with Successful Completers and Low Score Completers with Successful Completers. Note that a value of one means that a one unit change will have no effect on the outcome, and that values less than one promote student academic success whereas values larger than one do not.

Figure 1: Comparison of Odds Ratios for Each of the Interaction Variables for Both Equations



In an effort to make the odds ratio value more understandable, a calculation method has been developed in which the odds change is noted and the percentage change as a result of a one unit increase in the standardized interaction variable and holding all other variables constant can be computed. Typically, one would interpret the change in likelihood of being placed in the alternate outcome category, but for this study the goal is for the student to be in the reference category, successful students, so the findings were calculated to determine what the odds change and likelihood that a student would be predicted to be in the successful student category. Therefore, for the variable student-student interactions, the odds ratio was 0.261 and to calculate the change in odds one would subtract 1 from 0.261 and the value would be -0.739 indicating that an increase in one unit of discussion posting (5 posts) would result in a decrease of 73.9% chance of being in the alternate outcome category, in this case the non-completers category. However, the goal of this study is to determine what interactions result in promoting student academic success so the following tables calculate the chance of being placed in the successful student category, in order to do this, the odds calculations were reversed. For example, the odds change of being placed in the successful student category (the reference category) was calculated resulting in a 73.9% increase when an odds ratio value of 0.261 was used.

Table 13 shows the Odds Ratios for the non-completers compared to successful completers as well as the change in odds and percent likelihood change that would occur if the standardized independent interaction variables were to increase by one unit.

Table 13: Non-completers Compared to Successful Completers: Odds Ratios, Odds Change and Change in Likelihood for One Unit Increase for Each Variable Holding All Other Variables Constant

Variable	OR	Odds change for becoming a successful student	Likelihood to become a successful student (reference category)
Student-student	0.261	Increasing	73.9% more likely
Student-content	0.426	Increasing	57.4% more likely
Instructor posts	0.656	Increasing	34.4% more likely
Instructor mail	1.451	Decreasing	45.1% less likely

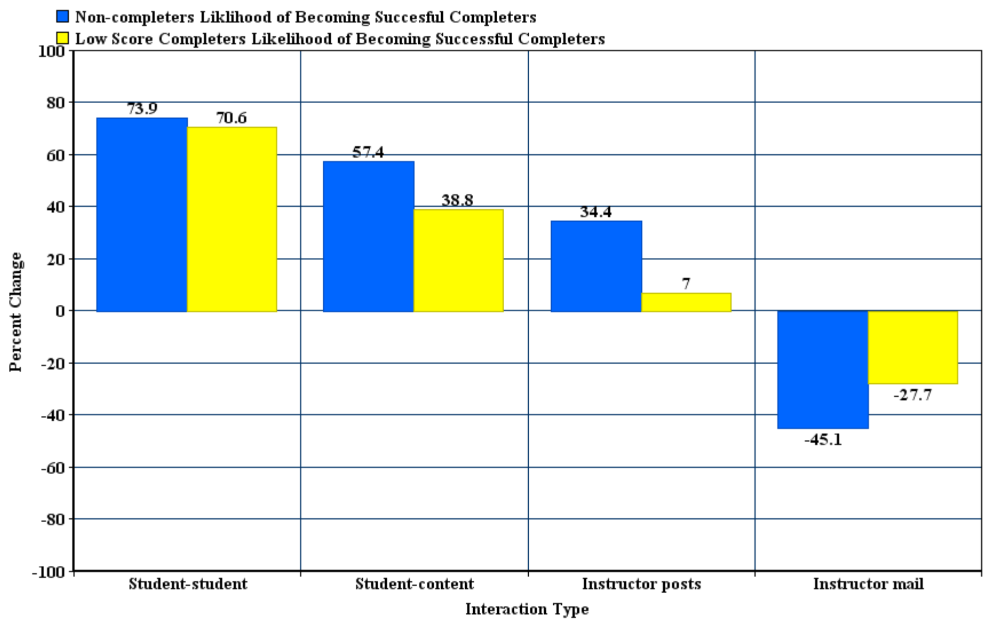
Table 14 shows the Odds Ratios for the low score completers compared to successful completers as well as the change in odds and percent likelihood change that would occur if the standardized independent interaction variables were to increase by one unit.

Table 14: Low Score Completers Compared to Successful Completers: Odds Ratios, Odds Change, and Change in Likelihood for One Unit Increase Holding All Other Variables Constant

Variable	OR	Odds change for becoming a successful completer	Likelihood to become a successful student (reference category)
Student-student	0.294	Increasing	70.6% more likely
Student-content	0.612	Increasing	38.8% more likely
Instructor posts	NS	No effect	No effect
Instructor mail	1.277	Decreasing	27.7% less likely

Figure 2 illustrates the percent change in likelihood that would occur if each of the interaction types (predictor variables) were increased by one unit, holding all other variables constant and converted to show the likelihood of being in the reference group, successful completers rather than the alternate outcome groups, non-completers or low completers.

Figure 2: Possible Change in Likelihood for Interaction Types



Interpretation of the Odds Ratio for each independent variables:

Non-completers compared to successful completers.

Student posts. The OR (odds ratio) of 0.261 implies that for each unit increase in student posts the odds of the student being in the non-completer category decreased by 73.9% (0.261-1 = -0.739) holding all other variables constant, and increased the chances of being in the successful completers category by 73.9%. More generally, if the subject would increase her number of postings by 5 posts, she would be 73.9% more likely to become a successful completer.

Student-content activity. The OR of 0.426 implies that for each unit increase in student content interaction the odds of the student being in the non-completer category decreased by 57.4% (0.426-1=- 0.574) holding all other variables constant, thus increasing the chances of being in the successful completer category by 57.4%. More generally, if the subject would increase her interactions with the content items, she would be more likely to become a successful completer.

Instructor posts. The OR of 0.656 implies that for each unit increase in instructor posts the odds of the student being in non-completer category decreased by 34.4% (0.656-1= -0.334) holding all other variables constant, and increased the chances of being in the successful completers category by 33.4%. More generally, instructors posting in the discussion area, increases the likelihood that the student is more likely to become a successful completer.

Instructor mail. The OR of 1.451 implies that for each unit increase in instructor emails the odds of the student being in the non-successful category increased by 45.1% (1.41-1=0.451) holding all other variables constant, and the chances of being in the successful category decreased by 45.1%. More generally, the greater the number of instructor mail interaction, the less likely the student would be placed in the successful completer category.

Low score completers compared to successful completers.

Student posts. The OR of 0.294 implies that for each unit increase in student posts the odds of the student being in the low score completers category decreased by 70.6% (0.294-1= -0.706) holding all other variables constant, and increased the chances of being in the successful completers by 70.6%. More generally, if the subject would increase her posting score, she would be more likely to become a successful completer.

Student-content activity. The OR of 0.612 implies that for each unit increase in student content interaction the odds of the student being in the low score completers category decreased by 38.8% (0.612-1=-0.388), holding all other variables constant therefore the chances of being in the successful completers category increased by 38.8%. More generally, if the subject would increase her content activity, she would be more likely to become a successful completer.

Instructor posts. This variable was found to not be significant in contributing towards the prediction for low score completers.

Instructor mail. The OR of 1.277 implies that for each unit increase in instructor emails the odds of the student being in the non-completers category increased by 27.7% (1.277-1=0.277) holding all other variables constant. More generally, the greater the number of instructor mail interaction, the less likely the student would be placed in the successful completer category.

Figure 2 above illustrates the percentage change in odds for each of the independent variables for non-completers versus successful completers and low score completers versus successful completers.

Research Question Five.

Are some interaction types more important than others in promoting student academic success?

The odds ratio is the unit of effect and an odds ratio of one means that the independent variable has no effect on the outcome category. The further the odds ratio is from one, the greater the effect size. An odds ratio greater than one indicates that the outcome was favored, whereas an odds ratio less than one indicates that the reference group is favored. In both regression equations in this study, the student-student interaction variable was the strongest contributor with an OR (odds ratio) value of .261 for the non-completers and .294 for the low score completers, translating to an increase in the odds of being a completer of 73.9% and 70.6% for each unit increase in the interaction variable respectively, holding all other variables constant. The second most important contributor was the student-content interaction (.426 for non-completers and .612 for low score completers) increasing the odds of becoming a successful completer (57.4% and 38.8% respectively) holding all other variables constant.

CHAPTER V: DISCUSSION

Introduction

This chapter includes a summary of the entire study. It is divided into the following sections: a brief overview of the study including a summary of the analysis, discussion of the findings, limitations of the study, implications for practice, future research, and conclusions.

Summary

Retention and student academic success has been the focus of higher education research for decades and continues to be an issue. Recently, President Obama refocused the nation's attention on retention and student success when he implemented the *American Graduation Initiative of 2009*. Paradoxically, the phenomenal growth of online learning escalates these concerns in that student retention and success in online learning is lower than that in the traditional classroom (Allen & Seaman, 2013; Lee and Choi, 2011; Patterson & McFadden, 2009). There is a pressing need to identify factors contributing to retention and success and to develop interventions promoting those factors. The online classroom provides a unique opportunity to study online student behavior without unduly interfering with the physical teaching and learning environment because the data is automatically captured and is not intrusive. The information gathered from this data could then be used to promote student academic success by providing guidelines for quality design and delivery of online classes.

Purpose of the study. The purpose of this study was to determine if online classroom activity in the first two weeks of a full semester class is related to student academic success, and if so, to determine if some types of activity are better predictors of academic success than others. This information might be able to be used to develop an intervention strategy to identify students at risk early enough in the semester to be able to help them to become successful completers.

The theoretical framework used for this study was that of interaction in the online classroom, following Moore's Interaction Theory (1989) for guidance. The definition used for interaction was, "a dialogue or discourse or event between two or more participants and objects which occurs synchronously and/or asynchronously mediated by response or feedback and interfaced by technology" (Muirhead & Juwah, 2005, p. 13).

The hypotheses guiding this study were:

Hypothesis 1: The number of online course interactions (with course content, instructor, and other students) in the first two weeks of class is related to successful completion.

Hypothesis 2: Student-content interaction promotes successful completion in online classes.

Hypothesis 3: Student-student interaction promotes successful completion in online classes.

Hypothesis 4: Student-instructor interactions promote successful completion in online classes.

Hypothesis 5: Some types of interactions are better promoters of success than others.

Literature Review. There are a number of initiatives, mostly at universities, to use and analyze the captured data, along with student demographic and academic data to identify factors that are associated with student academic success (Campbell, 2007; Grandzol & Grandzol, 2010; Lee & Choi, 2013; Maltby & Mackie, 2009; Morris, Finnegan, & Wu, 2005). Many of those studies utilize demographic and academic characteristics of their students in order to identify students at risk (Campbell, 2007), yet community college students as a group differ from students attending four-year institutions (Stafford, 2014). Therefore, many of the social and academic factors used in those models do not apply well to community college students (Wild & Ebers, 2002). In fact, Aragon and Johnson (2008) found no demographic difference between completers and non-completers in a study of community college students, supporting the theory that using demographic factors may not be useful in studies of community college students.

Methodology. This study examined archived tracking data retrieved from a learning management system for 1,703 students in 200 semester-long, fully asynchronous online community college courses. Moore's (1989) three categories of interactions were used for this study: student-student interaction, student-content interaction, and student-instructor interaction. Student-student interaction was captured using activity logs for the discussion areas in the learning management system; each time the student posted a comment it counted as an activity. Student-content interactions were captured by the learning management system. Each time a student opened a page it counted as a "hit". Initially the student-instructor interaction was represented by four independent variables: one for the number of emails the instructor sent and received, one for the number of posts to the discussion forum by the instructor, one for the number of instructor graded assessments, and one for the number of quizzes and/or exams in the course.

A forced entry multinomial logistic regression analysis was performed through SPSS 20.0 to assess placement of individuals in one of three categories of outcome (successful completers, low score completers, and non-completers). The category, successful completers, was designated as the reference group. Multinomial logistic regression in SPSS estimated two models, one for non-completers relative to successful completers and the second model for low score completers relative to successful completers.

The results showed that the original hypothesized overall model was insufficient in that one of the variables representing student-instructor interaction (instructor graded assignments) was found to not contribute significantly to the overall model and was thus removed. Further analysis revealed that the variable representing the number of quizzes and exams in the course was unreliable, and that variable was also removed. The final overall model then contained four independent variables rather than six. The final reduced model included four interaction independent variables: one student-student interaction, one student-content interaction, and two student-instructor interaction variables. The final overall model maintained the predictive power, and a 75/25 validation test met the criteria for internal validity.

The final overall reduced model contained one dependent variable with three levels: successful completers, low score completers, and non-completers. The four independent variables for the final model were: student-student interaction, student-content interaction, instructor discussion activity, and instructor mail counts,.

Summary of the analysis. Student-student interaction served as the strongest predictor for both generated models, followed by student-content interaction. Student-instructor interaction variables were confounding in their influences. The more instructor email there was, the less likely the students would be in the successful completer category. More discussion posts by the instructor resulted in increasing the odds for success for the non-completers and had no effect on the low score completers.

This study provides support that the level of student interaction in online courses with other students and with the content contributes to student academic success. The tracking data generated in the first two weeks of class can be used to identify students who are not participating at least at a minimal level and then allow for intervention to promote student succeeds. However, the notification probably should be something other than email from the instructor, since that tended to have a negative influence on student academic success.

Discussion of the Findings

Five research hypotheses provided guidance for this study. The findings from the data analysis previously presented will now be used to explore the meanings and implications as they apply to the guiding hypotheses.

Hypothesis One

The first hypothesis states, "the number of online course interactions (with course content, instructor, and other students) in the first two weeks of class is related to successful completion". This hypothesis was supported. The primary goal of this study was to identify students at risk of not succeeding in an online course in a timely manner so that interventions could be offered in time to affect the final outcome in the class. A test of the final full overall model with all five predictors against a constant only model was statistically reliable, indicating that the predictors as a set reliably distinguished between the three variable groups, successful completers, low score completers, and non-completers (chi square = 461.961, $p<1.0$, $df=8$). Non-significant Deviance and Pearson scores of 1.0 and a Nagelkerke's R^2 of .302 indicated a moderate relationship between prediction and grouping which further supports that the logistic model was more effective than the null model. The likelihood ratio test found that each of the independent variables was determined to be significant in the final full overall model indicating that each of the variables contributed to the prediction model and therefore played some role in student academic success.

Data from two weeks of activity in the online course was sufficient to generate a classification table that was correct 73.9% of the time and was more accurate than chance alone, thus providing a model that could be put in place to identify students at risk by the end of the second week of class. The researcher selected a time period of two weeks as the cut-off point as a time that would be early enough in the semester that there would be time for the interventions to be effective yet differing levels of student activity would be adequate to predict students at risk.

Odds Ratio and Percent Change in Odds

In order to discuss hypotheses two through five a brief explanation of odds ratio and percent change in odds is given here. As stated in the results section, the odds ratio is the measure of effect for each of the independent variables and it might be helpful to review the role of the odds ratio as a measure of effect for the predictor variables in multinomial logistic regression. The interpretation of odds ratio assumes that all other variables are held constant. Odds ratios are interpreted as follows: the odds ratio is the ratio of the odds for an individual being in an outcome category verses the odds of being in the reference category. In both analyses, the reference category was successful completers. An odds ratio less than one means that increasing the independent variable by one unit decreases the odds of the individual being placed in the alternate outcome category and increases the odds of the individual being in the reference category. An odds ratio of one indicates that changes in the independent variable would have no effect on category placement. An odds ratio greater than one indicates that increasing the independent variable by one unit would increase the chance of the individual being in the alternate outcome category and decrease the chance of the individual being in the reference category. Odds ratio is a complex concept and a more understandable interpretation is that of percent change in odds. The percent change in odds is calculated by subtracting one (1) from the OR and multiplying by 100 to give a percentage of likely change of being in the alternate outcome category. For this study, since the goal was to identify factors that promote student academic success, the percent change was adjusted to show the effect of the predictor variable on the chances of being in the reference group, successful students. Table 13 compares the effect size (OR) for each independent variable, the odds change for each, as well as at the likelihood of becoming successful completers for the non-completers for each unit increase in the variable considered while holding all other variables constant. Table 14 compares the effect size (OR) for each independent variable, the direction of odds change for each, as well as the likelihood of becoming successful completers for the low score completers for each unit increase in the variable considered while holding all other variables constant. Figure 1 shows the odds ratios for each of the variables for the models. Figure 2 illustrates the effect of a unit change in each of the predictors on the likelihood of percent change for the individual to be in the reference group (successful completers) for the two equations.

Hypothesis Two

The second hypothesis states, "student-content interaction promotes successful completion in online classes". This hypothesis was supported by the results of the analysis. In both models generated by SPSS (non-completers relative to successful completers and low score completers relative to successful completers), the Odds Ratio for student-content interaction indicated that increasing student-content interaction would increase the odds by 57.4% for non-completers and 38.8% for low score completers to be in the successful completer category, when holding all other variables constant.

In a totally asynchronous online class, everything the student needs to know about the mechanics of the class and what to do is embedded in the class website itself, whether it involves looking at the course calendar, reading the directions, accessing assignments, reading the syllabus, or taking quizzes, the student must enter the class area and navigate through it by clicking on links. Bernard et al. (2008) found a linear relationship between strong course design elements that promote the different interactions and course outcomes, yet in that study it was only strengthening the student-content interaction treatment that improved the effect size of course outcomes. If the course is well organized the student will have no problems finding out what they should be doing, thus increasing the self-efficacy of the student. However, courses with poor navigation and poor design frustrate students and many times the students just give up (Clay, 2009). In a study of student preferences in online courses, Sheridan, Kelly & Benz (2013) found that students want a well-designed and organized course and this promotes interaction with the content.

Hypothesis Three

The third hypothesis states "student-student interaction promotes successful completion in online classes". This hypothesis was supported by the results of the analysis. In both models, the Odds Ratio for student-student interaction showed that increasing student-student interaction (by making more discussion posts) would increase the odds 70.6% (low score completers) and 73.9% (non-completers) for the student being in the successful completer category rather than either the low score completer or non-completer categories.

Student-student interaction is an important component of online learning. The discussion areas provide a place to build a community of inquiry and to promote constructivist learning opportunities. This study confirms previous studies on the role of student-student interactions' importance in promoting student academic success in the online classroom (Anderson & Dron, 2010; Annand, 2011; Daspit & D'Souza, 2012; Dawson, 2006; Garrison, Anderson, & Archer, 1999; Kanuka, 2011). Swan (2002) cites discussion as one of the most effective practices for online courses. The huge body of evidence for the creation of a community in the online classroom based upon the early work by Rovai (2004) is supported by this study. Activity in the discussion area promotes the building of a community and furthers student academic success.

Hypothesis Four

The fourth hypothesis states "student-instructor interaction promotes successful completion in online classes". This hypothesis was only partially supported. The construct was difficult to capture and instead of one variable representing this interaction type, there were two in the final model. The two independent variables representing student-instructor interaction were the number of instructor posts to the discussion area and the quantity of instructor email in the class.

The number of instructor posts to the discussion areas promoted student academic success as indicated by the ability to increase the odds 34.4% for the non-completer being in the successful completer category when the interaction variable was increased by one unit. This finding supports the research of the importance of the instructor maintaining a presence in the discussion areas as a facilitator and promoter of student academic success (Dawson & McWilliam, 2008; Morris & Finnegan, 2009a; Ni & Aust, 2008). However, it is important to note that contrary research exists showing that if the instructor posts too early or too often, discussion often shuts down (Mazzolini & Maddison, 2003).

The number of instructor posts had no significant effect on the low score completers' group placement and this could be due to the possibility that these students were not interacting in the discussion area and were unaware of the instructor's activity. It also hints at a different nature of students who continue in a course even when they know they are not passing and those who withdraw (or are withdrawn) from a course.

This study found that increasing email by one unit resulted in a decreased likelihood of being in the successful student group by 45.1% for the non-completers and 27.7% for the low score completers group. At first this was confusing to the researcher since frequent contact with an instructor often makes up part of the instructor's presence, a proven promoter of student academic success. However, in this study, the email count was not directly tied to individual student interactions, but to that of an overall group thus changing the construct. The results made more sense when Carr's (2014) assertion that "Email frequency needs to be monitored as it can be a sign of insufficient information in the course for students or a lack of content support" (p.102) was brought into play. Further research could be conducted in which the courses with a high email count could be compared with those having a low email count to see how the course designs compare. It is also possible that instructors send more email to students who are not succeeding thus increasing total email count. A study where the data would be drilled down to individual counts including a possible content analysis might show differing results.

Hypothesis Five

Hypothesis five states, "some types of interaction are better promoters of success than others". This hypothesis was supported. In both models, non-completers relative to successful completers and low score completers relative to successful completers student-student interaction was the most potent contributor, followed by student-content interactions. This supports many studies that also show that student-student interaction and student-content interaction are critical to student academic success and that those components are vital to building a course that promotes student academic success.

Breaking the student-instructor interaction into two variables allowed for a closer examination of differing types of interactions between the students and the instructor.. Unfortunately, some of the components that may have been an important part of that construct were not included in this model. Yet, by constructing a variable that isolated the number of instructor emails and looking at that influence brought about the information as to how a poorly constructed course results in an increase in instructor email and could possible serve as a quick check item for course quality.

Limitations of the Study

There are a number of limitations in this study. This study did not control for gender, ethnicity, academic ability, discipline, or course design. The study was an attempt at stripping down the data to only what happened in the online courses and purposefully did not control for the aforementioned fields often because academic ability in higher education studies is typically based on ACT scores and/or SAT scores. Most community college students either don't take those exams or are returning students and their scores are out of date. Previous grades may or may not have a bearing on the study, many community college students stop out of college (non-continuous enrollment) and the grades may not be reflective of their abilities at the time of taking the course. The other factors were not considered because the outcome was based upon what the student was actually doing in the class regardless of those factors.

Another limitation was the use of only one sample to build and verify the model. The researcher was unable to acquire a second semester of data because the school was moving to a new learning management system and there was not enough manpower for someone to take time to pull the data. However, several methods were used to validate the data set used to accommodate this limitation.

An additional limitation was the inexperience of the individual extracting the data from the learning management system. This was a new procedure for him and it took multiple attempts to get data that was reliable (several checks were made by the researcher to validate the data). She was able to pull data from her courses and compare the tallies comprising the variables to the data provided by online services for her classes. This inexperience resulted in recognizing that the data for the announcements in the class was corrupted as was that for the assessments, and therefore were unusable, thus limiting the scope of that construct.

The inexperience of the researcher added greatly to the limitations and the process was a learn-as-you-go approach. A stronger background in advanced statistical analysis would have benefitted the process, or at least streamlined it considerably.

SPSS 20.0 was used for the analysis, and although it was adequate, a more robust statistical package may have fitted the research agenda better.

The complexity of the problem being examined and the challenge of dealing with a very large data set also added greatly to the limitations. This is a preliminary study and it is hoped it will provide a foundation for future studies.

Implications for Practice

The greatest challenge in delivering an asynchronous online class is that all of the interactions are computer mediated. How a course is taught is constrained by the mode of delivery. In order to increase student academic success, attention needs to be paid to what contributes to students doing well in online classes and what does not. Carr (2014) asserts, "Online course delivery requires an organized course format and delivery; an instructor who is knowledgeable in the environment; and students that are aware of the responsibilities and demands of the online setting" (p. 108). This trifecta requires careful course design, competent instructor training, and informing students as to what the online course expectations and responsibilities are even before they enter the online classroom.

This study found that student-student interaction was the most potent contributor to student academic success. The primary means of student-student interaction is in the discussion forums. The discussion forums allow for sharing of ideas and can promote the feeling of being in a traditional classroom and serves as a meeting place where community can be built.

Discussion can be structured in a manner that facilitates student-to-student interactions. The first discussion forum should be a place for building community by allowing the students to introduce themselves to one another and to the instructor (Swan, 2004). Discussion should be guided, for example, in the first introductions forum specific information should be requested (Al-Shalchi, 2009). Directions for the introduction area could read, "Please introduce yourself to the class here. Tell us what your major is and what your career goals are. Please also share one thing interesting about yourself that you are comfortable sharing." Providing very specific instructions guides the students so that most students will post about the same amount of information and the last statement invites them to share a personal statement with the other students, thus helping to start building the community of the class. Students need to be instructed to reply to another post (or two) to further the community building; again the directions must be very precise (Al-Shalchi, 2009).

Graded discussion forums enforce student participation and a rubric for grading can provide guidelines for expectations (Rovai, 2007; Swan, 2004; Al-Shalchi, 2009). Creating discussion areas for each module builds structure into the class. Students know that there will be a discussion on the topics for that module and if the discussion postings instructions promote early first postings it works best. The following example has worked well for this instructor. The posting instructions are "post one new thing that you learned from reading the chapter". This type of posting has two goals, one is to enhance the assigned reading and the second is that by writing about something covered in the chapter, it encourages them to reflect on the readings as well as reinforcing the information covered by the reading. An important guideline for these types of postings is that they cannot post their initial post on the same topic that someone else has posted. This motivates students to post early so that they can post on their first choice. Then the other students can reply to the post and add to the information shared. The instructor has an important role of facilitator by monitoring and managing the course discussions and by participating in the discussion area, but not too early in the discussion, nor too often (Mazzolini & Maddison, 2003; Morris, et al, 2005; Rovai, 2007). Anderson, Rourke, Garrison, and Archer (2001) assert that an important role for the instructor is to facilitate discussion by “moving the discussion along” and insuring effective and efficient use of time. The value of the discussions is for the students to interact with one another on a regular basis and to share their take on the material from the readings or activities. In the online environment there is time for students to reflect before responding thus allowing time for developing critical thinking skills and constructing thoughtful responses.

This study found the second most important promoter of student academic success to be that of student-content interaction. Interaction with content is pervasive in the online classroom experience. As mentioned above, everything a student does in the online class involves interaction with the content. Course design is a critical component and promoting clear navigation is imperative. The findings of this study showing that increased instructor emails lead to decreased chances of success link directly to poor course design (Carr, 2014). In order to promote student academic success, great care must be put into the design of online classes. There are a number of design models available, yet many institutions encourage instructors to move their classes into the online environment without assistance or guidance. Effective faculty training is critical. Sheridan, Kelly, and Bentz (2013) noted that the most important aspects are related to creating a well-designed course and being responsive to students’ needs. These characteristics include direct instruction, clarity, feedback, facilitation, discussion participation, flexibility, helpfulness, and understanding.

It is also important to note that learning involving a high level of student-content interaction requires high degrees of autonomy and self-direction that many students lack (Anderson, 2013; Garrison, 1997). Therefore, promoting behaviors that increase autonomy and self-directed learning may be needed to improve student retention and success. The implications for course design are obvious and significant, pointing to the benefits of weaving course content into a cohesive, compelling tapestry (Murray, et al., 2012, pp.137-138).

Perhaps the most significant finding in this study is that increasing the number of instructor email count reduced the odds for student academic success. Initially, this was quite confounding, but upon further research the link between high instructor emails and poor course construction was made (Carr, 2014). Aragon and Johnson (2008) reported in a study of student academic success in online community college courses that when students were contacted regarding why they did not complete their online courses twenty-eight percent indicated that it was due to course design or lack of communication (p. 155). This furthers the cause of careful construction of course sites.

The Web 2.0 environment has the potential to provide faculty the leverage that is needed to develop engaging courses that meets the learning needs of students. This change could revitalize higher education, however as Zemsky (2009) proclaims, “e-learning’s blossoming will require a fundamental shift in the culture of teaching.” But he also cautions, “For e-learning to reach its full potential, faculty will first have to recognize the technology as a means of solving problems that traditional pedagogies too often ignore” (p. 155).

Instructors need to be more creative in the construction of their courses, and not just transfer lectures to the online course but utilizing the Web 2.0 affordances develop engaging and empowering opportunities for students. This required shift in the instructor’s role from dispensers of knowledge (sage on the stage) to that of a content expert and facilitator (guide on the side) is well documented (Arbaugh, 2004; Desai, Hart, & Richards, 2008; Garrison, Anderson, & Archer, 1999; Palloff & Pratt, 2001). New methods of learning in the online environment need to be pursued and integrated into online classes.

Students' roles need to change as well, away from passively receiving knowledge to actively constructing and generating their own learning, both in isolation and in collaboration with other learners (Arbaugh, 2004; Garrison et al., 2000; Palloff & Pratt, 2002). For adult learners who may have been immersed in undergraduate classroom experiences based on the dispenser/recipient approach, this change in roles and expectations can be quite an adjustment (Greenhow et al., 2009). Student-centered learning is ideally suited to the online environment and the constructivist theory supports student-centered learning. One of the tenets of the constructivist theory is the instructor must also become a learner (Buckley, 2002). This is almost inevitable when instructors incorporate technology into the teaching arena. They too have to learn how to use the technology and then have to find ways to intertwine the technology with the content of the class (Churches, 2009). Institutions should provide guidance and training for instructors moving their classes online to facilitate this movement. Establishing a community of practice at the school in which new learning methods are explored and applied to new and current courses could be a starting point. Institutions will need to provide quality training and support for course design, they will need to ensure that instructors have time to develop content and to stay current with the Web 2.0 tools as well as learn about emerging technological web-based tools.

Learners must have access to appropriate and relevant content, know how to find it, and must have the opportunities to apply and practice what they have learned. Bates & Sangra

(2011) contend “learning has to be a combination of content, skills, and attitudes and increasingly the need to apply to all areas of study” (p. 50). Learning in the online environment is very different from that in a traditional lecture based classroom, online learning provides the mechanism to move from the instructor centered model to a student centered one. Students don’t necessarily have the skill sets that they need to flourish in the online environment.

There is a positive relationship between Web-based learning technology use and student engagement and desirable learning outcomes (Chen et al., 2010). Students are more likely to make use of deep approaches of learning like higher order thinking, reflective learning and integrative learning in their study and they reported higher gains in general education, practical competence, and personal and social development (Arbaugh, 2004, p.1230). “This cohort likely has been surrounded by technology from a very young age, they are apt to arrive on campus expecting technology to be an integral part of their college experience. Thus, it appears logical to integrate technology into the learning environment for these digital natives” (McCabe & Mueter, 2011, p. 149).

Recommendations for Future Research

There is a common theme running through the literature regarding future research – future research needs to examine the kinds of instructor and student roles in online interactions that enhance class discussions and encourage critical thinking and construction of knowledge (Tallent-Runnels et al., 2006, p. 17). This study focused on one method of looking at interactions and provides a preliminary examination of those factors. There are a number of studies that need to be done to follow up on this initial investigation.

The first study that should be done would utilize the same data, but control for gender, discipline, and course design. An attempt should be made to see if a measure of academic ability could be acquired, and if so it also should be used in the study as a control.

A continuation of this study using another semester's data from the community college to provide additional validation would be the first suggestion for further research. After that, it would be beneficial to apply the model to data from another community college to support external validation, and then perhaps expand to incorporate data from a university in hopes of extending the generalizability of the model.

In addition, this study could be enriched utilizing other methods of data analysis, possibly factor analysis or applying data mining methods to further examine interactions and the effects on student academic success.

The findings that increased instructor email was associated with a decreased chance of student academic success should be investigated. The first step would be to see if the courses with the most emails did, indeed, have content and structure issues. The next step would be do look at a few courses and do a content analysis of the email to determine if the emails were in response to course structure and content issues.

Another direction of future research would be to examine course structure and content and compare courses with high success rates to courses with low success rates to see if any specific guidelines could be constructed to promote student academic success.

Conclusion

This work contributes to the literature on interaction in the classroom several ways. The first contribution is that it suggests that examining data from the first two weeks is useful in predicting course success, although it wasn't one hundred percent correct, it could be used to identify students in need of intervention early enough in the class to perhaps help them become successful completers.

The findings form this study support the theory that building a community through student-student interaction promotes student academic success in the online environment (Anderson & Dron, 2011; Rovai, 2002a; Kanuka, 2011). The findings also indirectly support the positive aspects of using discussion activities as a means for constructivist learning opportunities (Rovai, 2004). The odds ratios generated by the multinomial logistic regression analysis suggest that if low score completers and non-completers would increase the number of posts by one unit then they could increase their odds of becoming successful students by 70.6% and 73.9% respectively, holding all other variables constant. In order to increase the odds for completion, students must be made aware of the importance of participating in the discussion areas early in the course. Course construction following best practice guidelines would include specific directions and rubrics so that students know what is expected of them in the online environment.

This study also contributes to the literature on student-content interaction, showing that increasing student-content interaction could result in increasing the odds of a student by 38.8% in becoming a successful completer rather than an low score completer and by 57.4% for non-completers in becoming successful, holding all other variables constant. The results from this study seem to counter Anderson's 2003 study in which he found that student-content interaction more effectively promoted learning that the other types of interaction, yet the only other interaction type that contributed more was the student-student interaction. It is important to note that this study only examined activity in the first two weeks of a full semester course, and it is very likely that as the course progresses through the semester, the student-content interaction becomes more valuable. The discussion postings are important in developing the concept of community and perhaps become less important as the course progresses and moves from the phase of community building to that of mastering content. However, the focus of this study was to determine if identification of students at risk of not completing was possible in the first two weeks in the class. The results from this study suggest that this is so.

The results regarding student-instructor interactions were confounding. The instructor discussion component supported Dawson and McWilliam's (2008) insistence on the importance of the instructor facilitating discussions and exhibiting an active presence in the discussion forum for maintaining an ongoing active presence for the non-completing students. Yet, the effect was non-significant in the analysis of the low score completers group. This could be due to the inherent characteristics of low score completers, specifically those characteristics that would encourage a student to continue participating in a course in which they were not passing. Perhaps they were not attending to the activities in the discussion area, this would be something worth following up on.

The quantity of instructor mail was inversely associated with student academic success for both comparison groups was initially a conundrum, if faculty email was a component of teaching presence then it should have been a positive factor. Yet, the odds ratios from the analysis indicated that increasing instructor email was associated with more students being in the non-completer and low score completer categories. This may occur because, as mentioned above, the course may be poorly designed and students needed more direction from the instructor than in a course that is well designed (Carr, 2014).

These finding suggest that a more robust model of interactions in online learning is needed to guide further investigation concerning the impact of online learning on student academic success in fully online courses. The current model purposely only considered the participants in the interaction and what was occurring in the course. A more complete model that considers both the participants and the nature of the interactions might be useful in guiding further studies on success in online courses.

This study found an increase in the email messages between the instructor and students resulted in a decrease in the likelihood that students would succeed in an online class. Whereas, the greater number of instructor posts the more likely the students were to succeed. This may be because the nature of the interaction is key in indicating student academic success. This study focused on isolating the interactions from their collective nature in order to distill out the essence of factors contributing to student academic success and identify if any were more important than the others. However, it appears that the Community of Inquiry (Garrison, et al. 1999) theoretical model that assigns presences to describe the collective nature of the interactions and not the interactions themselves must be applied to these results in order to understand the complex process for the student-instructor interactions.

The Community of Inquiry model represents a process of creating a deep and meaningful (collaborative-constructivist) learning experience through the development of three interdependent elements: social, cognitive and teaching presence. Social presence is "the ability of participants to identify with the community (e.g., course of study), communicate purposefully in a trusting environment, and develop inter-personal relationships by way of projecting their individual personalities" (Garrison, 2009). Social presence is made possible through the student-student interactions occurring in the discussion forums (Garrison & Arbaugh, 2007; Parscal, et al., 2012; Richardson, et al., 2012) and very clearly contributes to student academic success. Cognitive presence is the extent to which learners are able to construct and confirm meaning through sustained reflection and discourse (Garrison, et al., 2001). Teaching presence is "the design, facilitation, and direction of cognitive and social processes for the purpose of realizing personally meaningful and educationally worthwhile learning outcomes" (Anderson, et al., 2001, p.5). This validates using two independent variables in this study to capture the essence of the student-instructor interactions. Facilitation could be represented by the email in that the email interactions between the instructor and student that deal with course logistics such as questions about assignments, due dates, locating correct information and grading expectations may reflect potential problems in the organization of the course. On the other hand, instructor posting in the discussion areas may be directed more toward enhancing student understanding and promoting instructor presence. Ni and Aust (2008) found that the nature of the interaction in online courses can influence student satisfaction in online courses and it is very possible that the interactions that contribute to teaching presence contribute to student academic success.

In an attempt to identify specific behavior contributing to student academic success, a less robust model was used than that of the Community of Inquiry. The Community of Inquiry model is so complex that it is challenging to pinpoint the exact nature of the behavior capture by the learning management system without assuming a great deal about what the student is experiencing. Yet the model used in this study was limited to only looking at interactions. It may be that the nature of the interactions plays an important role in indicating student academic success. For example, email interactions between instructors and students that deal with course logistics such as questions about assignments, due dates, locating correct information and grading expectations may reflect potential problems in the organization of the course. On the other hand, instructor postings in the discussion forums may be directed more toward enhancing student understanding. Previous research (Ni & Aust, 2008) has also found that the nature of the interaction in online courses can influence student satisfaction in online courses. Dividing the student-instructor interactions by the nature of the communication (email versus discussion postings) provided some insight into the complexity of the interactions. Further studies examining the nature of the communications including content analysis might provide further insight that could more clearly identify the role of the instructor in student academic success.

The demand for online classes is still strong and identification of methods that promote success in online learning is needed to inform institutions, faculty, and developers as to what effective practices can be employed to promote student academic success.

This study finds that interactions in the first two weeks of an online class are associated with student academic success and while it was not strong in identifying characteristics of students at risk, it was a strong predictor of behavior that promoted student academic success and could serve as a support for encouraging those behaviors in the online classroom. Interventions could then be offered to guide students and possibly increase their chances of becoming successful completers.

References

Abrami, P. C., Bernard, R. M., Bures, E. M., Borokhovski, E., & Tamim, R. M. (2011). Interaction in distance education and online learning: Using evidence and theory to improve practice. *Journal of computing in higher education, 23*(2), 82-103.

Adams Becker, S., Estrada, V., Freeman, A., & Johnson, L. (2014). NMC Horizon Report: 2014 Higher Education Edition. *Austin, Texas: The New Media Consortium*.

Al-Shalchi, O. (2009). The effectiveness and development of online discussions. *Journal of Online Learning and Teaching, 5*(1), 104-108.

Allen, I., & Seaman, J. (2014). Grade Change: Tracking Online Education in the United States. Babson Survey Report.

Allen, I. E., & Seaman, J. (2013). *Changing Course: Ten Years of Tracking Online Education in the United States*: ERIC.

Allen, I. E., Seaman, J., Group, B. S. R., & Consortium, S. (2010). *Class differences: Online education in the United States, 2010*: Babson Survey Research Group.

An, Y. J., & Williams, K. (2010). Teaching with Web 2.0 Technologies: Benefits, Barriers and Lessons Learned. *International Journal of Instructional Technology and Distance Learning, 7*(3), 41-48.

Anderson, T. (2002). *An updated and theoretical rationale for interaction.* Paper presented at the Athabasca University: IT Forum Paper.

Anderson, T. (2003). Getting the mix right again: An updated and theoretical rationale for interaction. *The International Review of Research in Open and Distance Learning, 4*(2).

Anderson, T. (2013). Finding a place in the new world of networked teaching and learning. *International Conference on E-learning*.

Anderson, T., & Dron, J. (2011). Three generations of distance education pedagogy. *The International Review of Research in Open and Distance Learning, 12*(3), 80-97.

Anderson, T., Rourke, L., Garrison, D. R., & Archer, W. (2001). Assessing teacher presence in a computer conferencing context.

Annand, D. (2011). Social presence within the community of inquiry framework. *The International Review of Research in Open and Distance Learning, 12*(5), 40-56.

Appana, S. (2008). A review of benefits and limitations of online learning in the context of the student, the instructor, and the tenured faculty. *International Journal of Online Learning, 7*(1), 5.

Aragon, S. R., & Johnson, E. S. (2008). Factors influencing completion and noncompletion of community college online courses. *The Amer. Jrnl. of Distance Education, 22*(3), 146-158.

Arbaugh, J. (2004). Learning to learn online: A study of perceptual changes between multiple online course experiences. *The Internet and Higher Education, 7*(3), 169-182.

Arend, B. (2007). Course assessment practices and student learning strategies in online courses. *Journal of Asynchronous Learning Networks, 11*(4), 3-13.

Bates, A. W. T., & Sangra, A. (2011). *Managing technology in higher education: Strategies for transforming teaching and learning*: Jossey-Bass.

Bernard, R. M., Abrami, P. C., Borokhovski, E., Wade, C. A., Tamim, R. M., Surkes, M. A., & Bethel, E. C. (2009). A meta-analysis of three types of interaction treatments in distance education. *Review of Educational Research, 79*(3), 1243-1289.

Brown, J. S. (2001). *Learning in the digital age.* Paper presented at the The Internet and the University: Forum.

Boekaerts, M. (1997). Self-regulated learning: A new concept embraced by researchers, policy makers, educators, teachers, and students. *Learning and instruction, 7*(2), 161-186.

Boston, W., Díaz, S. R., Gibson, A. M., Ice, P., Richardson, J., & Swan, K. (2014). An exploration of the relationship between indicators of the community of inquiry framework and retention in online programs.

Boston, W. E., Ice, P., & Gibson, A. M. (2011). Comprehensive assessment of student retention in online learning environments. *Online Journal of Distance Learning Administration, 14*(4).

Buckley, D. P. (2002). In pursuit of the learning paradigm. [4-8-12]. *Educause Review, 37*(1), 29-38.

Campbell, J. P. (2007). *Utilizing student data within the course management system to determine undergraduate student academic success: An exploratory study.* (Ph.D.), Purdue. (3287222)

Campbell, J. P., DeBlois, P. B., & Oblinger, D. G. (2007). Academic analytics: A new tool for a new era. *Educause Review, 42*(4), 40.

Capra, T. (2011). Online Education: Promise and Problems. *Journal of Online Learning and Teaching, 7*(2), 299-293.

Carr, M. (2014). The Online University Classroom: One Perspective for Effective Student Engagement and Teaching in an Online Environment. *The Journal of Effective Teaching*, 99.

Cejda, B. (2010). Online education in community colleges. *New Directions for Community*

Colleges, 2010(150), 7-16.

Charbonneau, L. (2012). Students prefer good lectures over the latest technology in class, *UA/AU University Affairs, Affairs University*. Retrieved from http://www.universityaffairs.ca/students-prefer-good-lectures-over-the-latest-technology-in-class.aspx

Chen, K.-C., & Jang, S.-J. (2010). Motivation in online learning: Testing a model of self-determination theory. *Computers in Human Behavior, 26*(4), 741-752.

Chen, P. S. D., Lambert, A. D., & Guidry, K. R. (2010). Engaging online learners: The impact of Web-based learning technology on college student engagement. *Computers & Education, 54*(4), 1222-1232.

Cho, T. J. (2011). The Impact of Types of Interaction on Student Satisfaction in Online Courses. *International Journal on E-learning, 10*(2), 109-125.

Churches, A. (2009). Bloom's digital taxonomy. *Unpublished Dissertation.* http://edorigami.wikispaces.com/Bloom's+Digital+Taxonomy., *15*.

Conole, G., & Alevizou, P. (2010). A literature review of the use of Web 2.0 tools in Higher Education. *A report commissioned by the Higher Education Academy*.

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

Dabbagh, N., & Kitsantas, A. (2005). Using web-based pedagogical tools as scaffolds for self-regulated learning. *Instructional science, 33*(5-6), 513-540.

Dabbagh, N., & Kitsantas, A. (2013). Using Learning Management Systems as Metacognitive Tools to Support Self-Regulation in Higher Education Contexts *International Handbook of Metacognition and Learning Technologies* (pp. 197-211): Springer.

Daniel, J. S., & Marquis, C. (1979). Interaction and independence: Getting the mixture right. *Teaching at a Distance, 14*, 29-44.

Daspit, J., & D'Souza, D. (2012). Using the Community of Inquiry Framework to Introduce Wiki Environments in Blended Learning Pedagogies: Evidence from a Business Capstone Course. *Academy of Management Learning & Education*. doi: 10.5465/amle.2010.0154

Dawes, N. P., & Larson, R. (2011). How youth get engaged: Grounded-theory research on motivational development in organized youth programs. *Developmental Psychology, 47*(1), 259.

Dawson, S. (2006). A study of the relationship between student communication interaction and sense of community. *The Internet and Higher Education, 9*(3), 153-162.

Dawson, S., & McWilliam, E. (2008). Investigating the application of IT generated data as an indicator of learning and teaching performance. *Queensland University of Technology and the University of British Columbia, 41*.

Dede, C. (2009). Immersive interfaces for engagement and learning. *Science, 323*(5910), 66-69.

Desai, M. S., Hart, J., & Richards, T. C. (2008). E-Learning: Paradigm Shift in Education. *Education, 129*(2), 327-334.

Dunlap, J. C., & Lowenthal, P. R. (2011). Learning, unlearning, and relearning: Using Web 2.0 technologies to support the development of lifelong learning skills. *E-infrastructures and technologies for lifelong learning: Next generation environments. Hershey, PA: IGI Global*.

Education, U. S. D. o. (2011). The Condition of Education 2011. NCES 2011-033. *National Center for Education Statistics*.

Finnegan, C., Morris, L. V., & Lee, K. (2009). Differences by course discipline on student behavior, persistence, and achievement in online courses of undergraduate general education. *Journal of College Student Retention: Research, Theory and Practice, 10*(1), 39-54. doi: 10.2190/CS.10.1.d

Fritz, J. (2011). Classroom walls that talk: Using online course activity data of successful students to raise self-awareness of underperforming peers. *The Internet and Higher Education, 14*(2), 89-97.

Garrison, D., Anderson, T., & Archer W. (1999). Critical inquiry in a text-based environment: Computer conferencing in higher education. *The Internet and Higher Education, 2*(2), 87-105.

Garrison, D. R. (1997). Self-directed learning: Toward a comprehensive model. *Adult Education Quarterly, 48*(1), 18-33.

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.

Giesbers, B., Rienties, B., Tempelaar, D., & Gijselaers, W. (2012). Investigating the relations between motivation, tool use, participation, and performance in an e-learning course using web-videoconferencing. *Computers in Human Behavior*.

Goldstein, P. J., & Katz, R. N. (2005). *Academic analytics: The uses of management information and technology in higher education*: Educause.

Grandzol, C. J., & Grandzol, J. R. (2010). Interaction in online courses: More is not always better. *Online Journal of Distance Learning Administration, 13*(2), 2010.

Greenhow, C., Robelia, B., & Hughes, J. E. (2009). Learning, Teaching, and Scholarship in a Digital Age Web 2.0 and Classroom Research: What Path Should We Take Now?

Educational Researcher, 38(4), 246-259.

Grush, M. (2011). Monitoring the PACE of student learning: Analytics at Rio Salado Community College. *Campus Technology*.

Hachey, A., Conway, K., & Wladis, C. (2013). Community Colleges and Underappreciated Assets: Using Institutional Data to Promote Success in Online Learning. *Online Journal of Distance Learning Administration, 16*(1), n1.

Hart, C. (2012). Factors Associated With Student Persistence in an Online Program of Study: A Review of the Literature. *Journal of Interactive Online Learning, 11*(1), 19-42.

Hense, J., & Mandl, H. (2012). Learning" in" or" with" Games? Quality Criteria for Digital Learning Games from the Perspectives of Learning, Emotion, and Motivation Theory. *International Association for Development of the Information Society*.

Hidi, S. (2000). An interest researcher's perspective: The effects of extrinsic and intrinsic factors on motivation.

Hirumi, A. (2006). Analysing and designing e-learning interactions. *Interactions in online education: Implications for theory and practice*, 46-71.

Hosmer Jr, D. W., & Lemeshow, S. (1989). *Applied Logistic Regression*. United States of America: John Wiley & Sons.

Juwah, C. (2006). *Interactions in online education: implications for theory and practice*: Routledge.

Kanuka, H. (2011). Interaction and the online distance classroom: Do instructional methods effect the quality of interaction? *Journal of computing in higher education, 23*(2), 143-156.

Ke, F. (2010). Examining online teaching, cognitive, and social presence for adult students.

Computers & Education, 55(2), 808-820.

Kember, D. (1995). *Open learning courses for adults: A model of student progress*: Educational Technology.

Knowles, M. (1973). The adult learner: A neglected species.

Kuh, G. D. (2003). What we're learning about student engagement from NSSE: Benchmarks for effective educational practices. *Change: The Magazine of Higher Learning, 35*(2), 24-32.

Lea, S. J., Stephenson, D., & Troy, J. (2003). Higher education students' attitudes to student-centred learning: beyond'educational bulimia'? *Studies in higher education, 28*(3), 321-334.

Lee, Y., & Choi, J. (2011). A review of online course dropout research: implications for practice and future research. *Educational technology research and development, 59*(5), 593-618.

Levy, Y. (2007). Comparing dropouts and persistence in e-learning courses. *Computers & Education, 48*(2), 185-204.

Lim, D. H., & Kim, H. (2003). Motivation and learner characteristics affecting online learning and learning application. *Journal of Educational Technology Systems, 31*(4), 423-439.

Lockyer, L., & Dawson, S. (2011). *Learning designs and learning analytics.* Paper presented at the Proceedings of the 1st International Conference on Learning Analytics and Knowledge.

Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an “early warning system” for educators: A proof of concept. *Computers & Education, 54*(2), 588-599.

Maltby, A., & Mackie, S. (2009). Virtual learning environments-help or hindrance for the'disengaged'student? *ALT-J, 17*(1), 49-62.

Mazzolini, M., & Maddison, S. (2003). Sage, guide or ghost? The effect of instructor intervention on student participation in online discussion forums. *Computers & Education, 40*(3), 237-253.

McCabe, D. B., & Meuter, M. L. (2011). A Student View of Technology in the Classroom Does It Enhance the Seven Principles of Good Practice in Undergraduate Education? *Journal of Marketing Education, 33*(2), 149-159.

McGreal, R., & Elliott, M. (2008). Technologies of online learning (E-learning). *Theory and practice of online learning*, 115.

McMahon, M., & Oliver, R. (2001). *Promoting self-regulated learning in an on-line environment.* Paper presented at the World Conference on Educational Multimedia, Hypermedia and Telecommunications.

Miltiadou, M., & Savenye, W. C. (2003). Applying social cognitive constructs of motivation to enhance student success in online distance education. *AACE Journal, 11*(1), 78-95.

Moore, J. L., Dickson-Deane, C., & Galyen, K. (2011). e-Learning, online learning, and distance learning environments: Are they the same? *The Internet and Higher Education, 14*(2), 129-135.

Moore, M. G. (1989). Three types of interaction. *The American Journal of Distance Education, 3*(2), 1-6.

Morris, L. V., Finnegan, C., & Wu, S. S. (2005). Tracking student behavior, persistence, and achievement in online courses. *The Internet and Higher Education, 8*(3), 221-231.

Morris, L. V., & Finnegan, C. L. (2009a). Best practices in predicting and encouraging student persistence and achievement online. *Journal of College Student Retention: Research, Theory and Practice, 10*(1), 55-64.

Morris, L. V., & Finnegan, C. L. (2009b). University System of Georgia's eCore: Virtual general education. *New Directions for Higher Education, 2009b*(146), 25-34.

Morris, L. V., Wu, S. S., & Finnegan, C. L. (2005). Predicting retention in online general education courses. *The American Journal of Distance Education, 19*(1), 23-36.

Muirhead, B., & Juwah, C. (2004). Interactivity in computer-mediated college and university education: A recent review of the literature. *Educational Technology & Society, 7*(1), 12-20.

Mupinga, D. M., Nora, R. T., & Yaw, D. C. (2006). The learning styles, expectations, and needs of online students. *College Teaching, 54*(1), 185-189.

Murray, M., Pérez, J., Geist, D., & Hedrick, A. (2012). Student Interaction with Online Course Content: Build It and They Might Come. *Information Technology Education: Research, 11*, 125-140.

Ni, A. Y. (2013). Comparing the Effectiveness of Classroom and Online Learning: Teaching Research Methods. *Journal of Public Administration Education 19(2): 199-215*, 199.

Ni, S. & Aust, R. (2008). Examining Teacher Verbal Immediacy and Sense of Classroom Community in Online Classes. *International Journal on E-learning, 7*(3), 477-798.

Nickels III, G. M. (2005). Identifying measures of student behavior from interaction with a course management system. *Journal of Educational Technology Systems, 34*(1), 111-126.

Nicol, D. J., & Macfarlane‐Dick, D. (2006). Formative assessment and self‐regulated learning: A model and seven principles of good feedback practice. *Studies in higher education, 31*(2), 199-218.

Obama, B. (2009a, July 12, 2009). Rebuilding something better, *The Washington Post*. Retrieved from http://www.washingtonpost.com/wp-dyn/content/article/2009/07/11/AR2009071100647.html

Obama, B. (2009b). The American Graduation Initiative Retrieved 9-19-13, from http://www.whitehouse.gov/blog/Investing-in-Education-The-American-Graduation-Initiative

Oblinger, D., & Oblinger, J. (2005). Is it age or IT: First steps towards understanding the Net Generation. *Educating the Net generation*, 2.1-2.20.

Oliveira, I., & Tinoca, L. (2011). Online group work patterns: How to promote a successful collaboration. *Computers & Education*.

Palloff, R. M., & Pratt, K. (2001). *Lessons from the cyberspace classroom: The realities of online teaching*: Wiley. com.

Paris, S. G., & Winograd, P. (2001). The role of self-regulated learning in contextual teaching: Principles and practices for teacher preparation. *Washington, DC, US Department of Education. Retrieved April, 16*(2009), 2001-2004.

Park, J. H., & Choi, H. J. (2009). Factors influencing adult learners' decision to drop out or persist in online learning. *Educational Technology & Society, 12*(4), 207-217.

Parscal, T., Sherman, K., Heitner, K., & Lucas, G. (2012). *A New Model to Measure Student Engagement with a Mobile Learning Tool Integrating CoI and FRAME.* Paper presented at the Global TIME.

Pew, S. (2007). Andragogy and pedagogy as foundational theory for student motivation in higher education. *InSight: A Journal of Scholarly Teaching, 2*(1), 14-25.

Phillips, R., Baudains, C., & van Keulen, M. (2002). An evaluation of student learning in a web-supported unit on plant diversity.

Picciano, A. G. (2012). The Evolution of Big Data and Learning Analytics in American Higher Education. *Journal of Asynchronous Learning Networks, 16*(3), 9-20.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of educational Psychology, 82*(1), 33-40.

Pintrich, P. R., & Zusho, A. (2002). Student motivation and self-regulated learning in the college classroom *Higher education: Handbook of theory and research* (pp. 55-128): Springer.

Pintrich, P. R., & Zusho, A. (2007). Student motivation and self-regulated learning in the college classroom *The scholarship of teaching and learning in higher education: An evidence-based perspective* (pp. 731-810): Springer.

Reeve, J. (2012). A self-determination theory perspective on student engagement. *Handbook of research on student engagement*, 149-172.

Renes, S. L., & Strange, A. T. (2011). Using technology to enhance higher education. [joournal article]. *Innovative Higher Education, 36*(3), 203-213. doi: 10.1007/s10755-010-9167-3

Reushle, S., Dorman, M., Evans, P., Kirkwood, J., McDonald, J., & Worden, J. (1999). *Critical elements: Designing for online teaching.* Paper presented at the ASCILITE.

Richardson, J. C., Arbaugh, J. B., Cleveland-Innes, M., Ice, P., Swan, K. P., & Garrison, D. R. (2012). Using the Community of Inquiry Framework to Inform Effective Instructional Design. *The Next Generation of Distance Education: Unconstrained Learning*, 97.

Rienties, B., Giesbers, B., Tempelaar, D., Lygo-Baker, S., Segers, M., & Gijselaers, W. (2012). The role of scaffolding and motivation in CSCL. *Computers & Education, 59*(3), 893-906.

Rosen, J. A., Glennie, E. J., Dalton, B. W., Lennon, J. M., & Bozick, R. N. (2010). Noncognitive skills in the classroom: New perspectives on educational research. *RTI International*.

Rovai, A. P. (2002a). Sense of community, perceived cognitive learning, and persistence in asynchronous learning networks. *The Internet and Higher Education, 5*(4), 319-332.

Rovai, A. P. (2002b). Building sense of community at a distance. *The International Review of Research in Open and Distance Learning, 3*(1).

Rovai, A. P. (2003). In search of higher persistence rates in distance education online programs. *The Internet and Higher Education, 6*(1), 1-16.

Rovai, A. P. (2004). A constructivist approach to online college learning. *The Internet and Higher Education, 7*(2), 79-93.

Rovai, A. P. (2007). Facilitating online discussions effectively. *The Internet and Higher Education, 10*(1), 77-88.

Russo, T., & Benson, S. (2005). Learning with invisible others: Perceptions of online presence and their relationship to cognitive and affective learning. *Educational Technology & Society, 8*(1), 54-62.

Ryan, R. M., & Deci, E. L. (2000a). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist, 55*(1), 68-78.

Ryan, R. M., & Deci, E. L. (2000b). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psychology, 25*(1), 54-67.

Schwab, R. (2002). SPSS Multinomial Logistic Regression Problem Retrieved Feb. 26, 2014, 2014, from http://www.utexas.edu/courses/schwab/sw388r7/

http://www.utexas.edu/courses/schwab/sw388r7/Tutorials/SPSSMulitNomialLogisticRegression Problem_doc_html/

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

Sheridan, K., Kelly, M., and Bentz, T. (2013). A follow-up study of the indicators of teaching presence critical to students in online courses. In D. R. Zehra Akyol and Garrison (Ed.), *Educational Communities of Inquiry: Theoretical Framwork, Research and Practice* (pp. 67-83). Hershey, PA: Information Science Reference.

Simpson, O. (2004). The impact on retention of interventions to support distance learning students. *Open Learning: The Journal of Open, Distance and e-Learning, 19*(1), 79-95.

Smith, L. (2006). Best Practices in Distance Education. *Distance Learning, 3*(3), 59-66.

Smith, V. C., Lange, A., & Huston, D. R. (2012). Predictive Modeling to Forecast Student Outcomes and Drive Effective Interventions in Online Community College Courses. *Journal of Asynchronous Learning Networks, 16*(3), 51-61.

Stafford, S. (2014). Comparison Between Community Colleges and Four-Year Institutions, from http://www.education.com/reference/article/compare-community-college-four-year/

Street, H. (2010). Factors influencing a learner's decision to drop-out or persist in higher education distance learning. *Online Journal of Distance Learning Administration, 13*(4).

Swan, K. (2002). Building learning communities in online courses: The importance of interaction. *Education, Communication & Information, 2*(1), 23-49.

Swan, K. (2004). Relationships between interactions and learning in online environments. *The Sloan Consortium*.

Tabachnick, B. G., & Fidell, L. S. (2001). Using multivariate statistics.

Tallent-Runnels, M. K., Thomas, J. A., Lan, W. Y., Cooper, S., Ahern, T. C., Shaw, S. M., & Liu, X. (2006). Teaching courses online: A review of the research. *Review of Educational Research, 76*(1), 93-135.

Thurmond, V., & Wambach, K. (2004). Understanding interactions in distance education: A review of the literature. *International Journal of Instructional Technology and Distance*

Learning, 1(1).

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research, 45*(1), 89-125.

Wagner, E. D. (1994). In support of a functional definition of interaction. *American Journal of Distance Education, 8*(2), 6-29.

Wild, L., & Ebbers, L. (2002). Rethinking student retention in community colleges. *Community College Journal of Research & Practice, 26*(6), 503-519.

Xie, K., DeBacker, T. K., & Ferguson, C. (2006). Extending the traditional classroom through online discussion: The role of student motivation. *Journal of Educational Computing Research, 34*(1), 67-89.

Yacci, M. (2000). Interactivity demystified: A structural definition for distance education and intelligent CBT. *Educational Technology, 40*(4), 5-16.

Zemsky, R. (2009). *Making reform work: The case for transforming American higher education*: Rutgers Univ Pr.

Zepke, N., & Leach, L. (2010). Improving student engagement: Ten proposals for action. *Active Learning in Higher Education, 11*(3), 167-177.

Zhang, H., Almeroth, K., Knight, A., Bulger, M., & Mayer, R. (2007). *Moodog: Tracking students' online learning activities.* Paper presented at the World Conference on Educational Multimedia, Hypermedia and Telecommunications.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into practice, 41*(2), 64-70.

Appendices

Appendix A: Cases by Course Department

Department	Number of Cases	Percent
ADCN	58	3.4
ALHT	110	6.5
BIOL	347	20.4
BUSN	160	9.4
CHEM	55	3.2
CHLD	47	2.8
CIST	52	3.1
CRJS	50	2.9
ECON	27	1.6
EDUC	21	1.2
ENGL	158	9.3
EXSC	54	3.2
FNAR	10	0.6
FRSC	24	1.4
GEOG	19	1.1
HIST	91	5.3
HUDV	9	0.5
JOUR	5	0.3
MATH	43	2.5
MUSC	16	0.9
NASC	54	3.2
NURS	66	3.9
PHIL	75	4.4
PSYC	87	5.1
RSCR	11	0.6
SOSC	21	1.2
SPCH	20	1.2
VTSP	12	0.7
Total	1702	100.0