

RELATIONSHIP BETWEEN PEDAGOGIC AND COURSE FACTORS
AND STUDENT OUTCOMES IN COMMUNITY COLLEGE ONLINE COURSES

A dissertation presented to the faculty of the Graduate School of
Western Carolina University in partial fulfillment of the
requirements for the degree of Doctor of Education.

By

Marlowe Geoffrey Mager

Director: Dr. Bianca Montrosse
Assistant Professor
Department of Human Services

Committee Members:
Dr. Frederick Buskey, Department of Human Services
Dr. Mary Jean Ronan Herzog, School of Teaching and Learning
Dr. Jana Ulrich, Department of Institutional Success, Stanly Community College

September 2012

ACKNOWLEDGEMENTS

I wish to thank my committee members, Dr. Frederick Buskey, Dr. Mary Jean Herzog, and Dr. Jana Ulrich, for their time, effort, and valuable input. I would especially like to thank my director, Dr. Bianca Montrosse, for her invaluable assistance and unwavering support.

I also owe a great debt to all of the professors whom have guided me through this program, especially Dr. Ann Alexander who introduced me to community college leadership as a discipline, rather than as a vocation. A large number of people have supported me in this endeavor as well, particularly many of my classmates in the doctoral program and many of my coworkers at Haywood Community College, and I owe them my thanks as well. I owe particular thanks to my two reviewers who assisted with the pilot study.

Finally, and most importantly, I could not have achieved this goal without the love and support of my wonderful wife.

TABLE OF CONTENTS

	Page
List of Tables	6
List of Figures	7
List of Abbreviations	8
Abstract	9
Chapter One: Introduction	11
Background	11
Purpose of the Study and Theoretical Framework	14
Significance of the Study	15
Methodology	17
Predictive and Outcome Variables.....	17
Research Questions.....	20
Study Design.....	20
Pilot Assessment Procedure.....	22
Analyses.....	22
Remaining Chapters	23
Summary	24
Chapter Two: Literature Review	25
The Rise of Online Distance Learning.....	26
Online Distance Learning as Disruptive Innovation.....	29
Educational Leadership and Online Distance Learning.....	32
The Nature of Online Distance Learning.....	35
Research on Online Distance Learning.....	37
Parity of Online and Face-to-Face Instruction.....	38
Other Online Distance Learning Research Topics.....	41
Theoretical Framework: A Predictive Model of Student Success in Online Courses ..	42
Factors Un-related to Instructor Behaviors.....	43
Factors Related to Instructor Behaviors: Pedagogic and Course Factors	44
Pedagogic Factors	45
Student Directed Learning	45
Interactivity	47
Building Student Capacity	52
Formative Assessment	53
Course Factors	55
Measurable Learning Objectives	55
Varied Teaching Activities	56
Varied Sensory Modalities.....	59
Preprogrammed Instructor Communication	60
Synchronous Communication.....	61
Immediacy.....	62
Hybrid Instruction.....	62
Outcome Measures.....	63
Learning Outcomes.....	64

Student Satisfaction	65
Retention	65
Higher Learning	66
Purpose of the Study and Research Questions	66
Chapter Three: Methodology	69
Research Design	72
Sample and Population	74
Selection of Courses	74
Selection of Students and Population Demographics	77
Operational Definitions	77
Counts of Predictive Variables	78
Student Directed Learning	78
Interactivity	79
Building Student Capacity	80
Formative Assessment	81
Measurable Learning Objectives	82
Varied Teaching Activities	83
Varied Sensory Modalities	84
Preprogrammed Instructor Communication	84
Synchronous Communication	85
Immediacy	86
Learning Outcomes	87
Student Satisfaction	87
Retention	88
Pilot Assessment	89
Phase 1 - Selection of Operational Definitions from the Literature	89
Phase 2 - Operational Definition Testing by the Author	90
Phase 3 - Operational Definition Testing by Reviewer 1	91
Phase 4 - Consensus	92
Phase 5 - Operational Definition Testing by Reviewer 2	93
Phase 6 - Ongoing Validation of Operational Definitions ("Spot-checking")	93
Data Collection and Analysis	96
Summary	98
Chapter Four: Results	99
Descriptive Results	99
Predictive and Outcome Variables	102
Student Directed Learning	102
Student-Student Interaction	103
Student-Teacher Interaction	104
Student-Content Interaction	106
Building Student Capacity	106
Formative Assessment	107
Measurable Learning Objectives	108
Varied Teaching Activities	108
Varied Sensory Modalities	109
Preprogrammed Instructor Communication	110

Synchronous Communication.....	111
Immediacy.....	111
Course Grade	112
Assignment Grade.....	113
Retention.....	114
Student Satisfaction	114
Comparisons Between Courses.....	114
Intercorrelations Among Predictive and Outcome Variables.....	118
Inferential Results	121
Research Question 1: Course Grade	121
Model 1: Mean Course Grade.....	122
Model 2: Mean Course Grade with Level-1 Controls.	124
Model 3: Mean Course Grade with Final Level-1 Controls.	126
Model 4: Mean Level-1 and Level-2 Effects on Mean Course Grade.....	128
Research Question 2: Assignment Grade	131
Model 1: Mean Assignment Grade.....	132
Model 2: Mean Assignment Grade with Level-1 Controls.....	135
Model 3: Mean Level-1 and Level-2 Effects on Mean Assignment Grade.....	136
Research Question 3: Retention.....	139
Model 1: Mean Retention.	141
Model 2: Mean Retention with Level-1 Controls.....	142
Model 3: Mean Level-1 and Level-2 Effects on Mean Retention.	143
Research Question 4: Student Satisfaction	145
Analysis Limitations and Assumptions.....	146
Summary	148
Chapter Five: Discussion, Implications, and Limitations.....	149
Answers to Research Questions	149
Research Question 1	151
Research Question 2	154
Research Question 3	158
Research Question 4	159
Relationships between the Research Question Answers	159
Recommendations for Educators and Educational Administrators.....	161
Limitations of the Study and Recommendations for Future Research.....	165
Conclusions	170
References.....	173
Appendices.....	192
Appendix A: Original Operational Definitions for Use in Pilot Phase 2.....	193
Appendix B: Instructions Given to Reviewer 1 for Use in Pilot Phase 3	195
Appendix C: Sample of Data Produced by Reviewer 1 in Pilot Phase 3	197
Appendix D: Instructions Given to Reviewer 2 for Use in Pilot Phase 5	198
Appendix E: Data from Pilot Phase 6	200

LIST OF TABLES

Table	Page
1. Concise Operational Definitions by Level of Analysis	69
2. Pilot Assessment Phase 6 Interrater Reliability	96
3. Descriptive Statistics for Level-1 Predictive and Outcome Variables.....	100
4. Descriptive Statistics for Level-2 Predictive and Outcome Variables.....	101
5. Presence of Each Predictive Variable at Least Once per Course.....	116
6. Intercorrelations for All Predictive Variables.....	119
7. Intercorrelations for All Outcome Variables	120
8. Fixed Effects Estimates for Models of the Predictors of Course Grade	122
9. Fixed Effects Estimates for Models of the Predictors of Assignment Grade	133
10. Fixed Effects Estimates for Models of the Predictors of Retention.....	140
11. Answers to Research Questions.....	150

LIST OF FIGURES

Figure	Page
1. Variability in the Presence of Each Predictive Variable across Courses.....	117
2. Statistical Significance and Direction of the Variables that Contributed to the HLM Models Used to Answer Research Questions 1, 2, and 3151	151

LIST OF ABBREVIATIONS

AGd - Assignment grade
BSC - Building Student Capacity
CGd - Course grade
FOR - Formative Assessments
IMM - Immediacy
MLO - Measurable Learning Objectives
NFOR - Number of Formative Assessments
PIC - Preprogrammed Instructor Communication
RET - Retention
SC - Synchronous Communication
SCI - Student-Content Interactivity
SDL - Student Directed Learning
SS - Student Satisfaction
SSI - Student-Student Interactivity
STI - Student-Teacher Interactivity
VSM - Varied Sensory Modalities
VTA - Varied Teaching Activities

ABSTRACT**RELATIONSHIP BETWEEN PEDAGOGIC AND COURSE FACTORS AND
STUDENT OUTCOMES IN COMMUNITY COLLEGE ONLINE COURSES**

Marlowe G. Mager

Western Carolina University (September 2012)

Director: Dr. Bianca Montrosse

The purpose of this study was to identify instructor behaviors that lead to positive student outcomes in online courses. The study investigated the relationship between 12 predictive variables and three measures of student success (assignment grade, course grade, and student retention) in online courses. Archived online courses at a rural community college were analyzed for the presence of the predictive variables, with each variable counted within the course and within each course activity completed by each student. Outcome variables were determined through the college's data warehouse and the online courses' gradebooks. Hierarchical linear modeling was used to analyze the predictive value of each predictive variable as it relates to the three outcome measures; the fourth was described in terms of correlations. Student-student interaction and immediacy were significant predictors of all three outcome variables, while other variables were inconsistent across outcomes or were not statistically significant predictors. Course grade was positively correlated with student-student interaction, student formative behaviors, and immediacy. It was negatively correlated with building student capacity. Assignment

grade was positively correlated with student-student interaction, student-teacher interaction, student formative behavior, immediacy, and varied teaching activities. It was negatively correlated with number of formative activities in the course and preprogrammed instructor communication. Student retention was positively correlated with student-student interaction and immediacy. Possible explanations for these findings are discussed in relationship to the literature. Recommendations for future research and online instructional practice are suggested.

CHAPTER ONE: INTRODUCTION

Distance learning through online technology has become an increasingly common means of instruction (Beaubien, 2002; Farnsworth & Bevis, 2006; Sims, Dobbs, & Hand, 2002; U.S. Department of Education [USED], 2009). According to the National Center for Education Statistics (2008), during the 2006-2007 academic year, 66% of 2- and 4-year Title IV degree-granting institutions offered some form of distance learning courses. Githens, Crawford, and Sauer (2010) reported that 47.5% of community colleges offer at least one program primarily or entirely online. However, little research has been conducted on the ways that specific instructor behaviors affect student learning and achievement. The purpose of this study was to identify instructor behaviors that lead to positive student outcomes in online courses.

Background

Only 10% of postsecondary students took an online course in 2002, but 25% did so in 2008, 29% did so in 2009, and it is estimated that 50% will do so in 2014 (Christensen, Horn, Caldera, & Soares, 2011, p. 31). Clearly, this rapid growth of a new way of instruction requires close scrutiny; 94% of institutions surveyed report that they develop their own distance learning courses (National Center for Education Statistics, 2008, p. 3). It would benefit these institutions to know what instructor behaviors lead to positive student outcomes in online courses.

Research into online teaching and learning has taken a variety of tracks, but few studies have attempted to directly correlate instructor behaviors with student learning. Generally, outcome measures have addressed perceived student learning or other indirect

measures (see for instance Sher, 2004), rather than objective measures of student outcomes, such as grades or retention. Additionally, it is quite rare that individual teachers approach online teaching differently than they do face-to-face instruction. Despite the ability of networked computers to provide a wide range of media and interactivity, many distance learning teachers continue to treat online instruction as a "translation" of traditional face-to-face lectures into a new media (Stevens-Long & Crowell, 2002, p. 153).

In a 2009 meta-analysis of instructional and media elements designed to facilitate online student learning, Bernard et al. (2009) summarized extant research with three major conclusions, the first of which is that "distance education can be much better and much worse than classroom instruction" (p. 1246). Distance learning is not a singular experience; it can be made highly effective or highly ineffective and educational leaders must be able to shape online courses such that they are effective ones. Unfortunately, the focus has been less on making distance learning better and more on managing the growth of online education at the expense of pedagogical concerns (Picciano & Seaman, 2010, p. 24).

The need to integrate effective online pedagogy with growing online offerings is evident, and the importance of this need is increasing steadily. The impact of the rise of online distance learning cannot be overstated. For instance, the North Carolina Community College System estimates that 90% of its courses will be delivered online by 2018 (NCCCS, 2010). Indeed, Christensen, Horn, Caldera et al. (2011) equate the advent of online education to the same "disruptive innovation" that changed consumer computers from expensive, place-bound devices to relatively inexpensive mobile ones.

"Disruptive innovations fundamentally transform a sector by replacing expensive, complicated, and inaccessible products or services with much less expensive, simpler, and more convenient alternatives" (Christensen, Horn, & Johnson, 2011, p. 17). In other words, disruptive innovation leads to a substitution model in which a new way of doing something replaces an older way. In this context, online students, courses, and programs are rapidly replacing traditional face-to-face students, courses, and programs (Christensen, Horn, Caldera et al., p. 31).

This rapid substitution of online for face-to-face courses is not occurring in a vacuum. Educational leaders are confronted with a number of challenges related to economic changes, often referred to as the new "knowledge economy" (Batson, 2010; Lumina Foundation, 2011; Symonds, Schwartz, & Ferguson, 2011). Though "knowledge economy" has been defined in varying ways, common elements include the increasing importance of intellectual capital over that of physical and natural resources, the rapid obsolescence of existing technologies and business strategies, and the accelerated pace of technological change (Powell & Snellman, 2004); coupled with the increasing role of globalization on all aspects of business and education (Houghton & Sheehan, 2000). While the knowledge economy has altered what educators do in order to prepare students, the expectations of students and the general public concerning higher education have also changed dramatically in recent years (Tulinko, Glasser, Heus, Isaacs, & Wald, 2005). It has not been easy for higher education leaders to manage these changes. "One of the primary challenges facing higher education is the seemingly unending spiral of expectations regarding changes in the ways colleges and universities must operate" (Ulrich, 2009, p. 10). It has been suggested that effectively adapting to the rapid rise of

the knowledge economy and rapid changes in student and public expectations requires the prompt yet careful use of online learning and extant Web 2.0 tools (Batson; Christensen, Horn, Caldera et al., 2011).

Purpose of the Study and Theoretical Framework

Online courses are not a "magic bullet" for the challenges facing educational leaders. As Hill, Domizi, and Collier (2011) stated, "we need to build a deeper understanding of how to design effective [distance learning] environments to enhance and extend the learning process" (p. 91). In response to this need, the purpose of this study was to identify instructor behaviors that lead to positive student outcomes in online courses.

However, there are so many factors that might affect student outcomes, and so many outcomes that may be unrelated to the nature of online instruction, that a theoretical framework was required in order to narrow the focus of this study. Menchaca and Bekele (2008) suggested one such framework. They identified five kinds of factors that affect success in online learning: human factors (e.g., student motivation), leadership factors (e.g., faculty training), technology factors (e.g., internet connection speed), pedagogic factors (e.g., nature and extent of teacher feedback provided to students), and course factors (e.g., teaching activities). Since the purpose of this study was to identify instructor behaviors that lead to positive student outcomes in online courses, it focused specifically on the latter two factors (pedagogic and course factors) as the ones that individual instructors can address within a single course.

Menchaca and Bekele (2008) defined pedagogic factors as the "*how of learning and instruction* [italics added]" (p. 237). In other words, pedagogic factors are those that

inform an instructor's philosophical stance towards teaching. Menchaca and Bekele defined course factors as instructional design elements (p. 237). In other words, course factors are the specific activities that teachers use to convey content or assess learning. This framework therefore suggests that pedagogic and course factors can be identified from the literature as those that would be likely to promote successful student outcomes in online courses. However, "outcomes" must also be defined.

Outcome measures have also been defined in so many ways that a guiding model is needed in order to identify the ones of interest. Menchaca and Bekele (2008) identified the following measures of success: learning outcomes, student satisfaction, higher learning, faculty satisfaction, sustainability, scalability, and retention (p. 236). As with the pedagogic and course factors discussed previously the ones of interest here were those related to and measurable within individual courses. There are three of these: learning outcomes, student satisfaction, and retention within the course.

The purpose of the study was therefore to identify instructor behaviors that lead to positive student outcomes in online courses. The guiding framework suggested that these behaviors would be pedagogic or course factors and that the outcomes of interest would be related to student learning, satisfaction, and retention.

Significance of the Study

Educational leaders are confronted with a number of challenges related to economic changes, often referred to as the new "knowledge economy" (Batson, 2010; Gordon, 2011). This same observation has been made by influential educational agencies, including the Harvard Graduate School of Education (Symonds et al., 2011) and the Lumina Foundation (2011). "Higher learning has taken on new importance in today's

knowledge society. To succeed in the contemporary workplace, today's students must prepare for jobs that are rapidly changing, use technologies that are still emerging and work with colleagues from (and often in) all parts of the globe" (Lumina Foundation, p. 1). It is not surprising that online courses have been viewed as a solution to these challenges, given that both technology and distance are important factors in this new economy.

In 2003, 49% of higher education academic leaders identified online education as critical to long-term strategic planning (Allen & Seaman, 2005, p. 11). By 2005 this percentage increased to 56% (Allen & Seaman, 2005) and by 2011 the percentage increased to 65% (Allen & Seaman, 2011). Despite this focus on managing online growth, administrators focus very little on the quality of online courses. "Policy decisions are based on the rationale that providing broader access to a secondary education may be of more importance than the concerns and perceptions regarding the pedagogical value of online learning" (Picciano & Seaman, 2010, p. 24).

Inconsistencies in the quality of online courses, and indeed of education as a whole, has not been ignored by the public in general or students in particular, as noted in such diverse sources as a 2005 PBS documentary (Tulinko et al., 2005), a large-scale meeting of university presidents (Association of Public and Land Grant Universities, 2010), and increasingly complex federal and state government regulations surrounding the use of technology for education (Nagel, 2010).

Educational leaders must address these concerns, and online instruction is increasingly being seen as an effective means of doing so (Christensen, Horn, Caldera et al., 2011). For this reason, it is imperative that educational leaders have available to them

research-based recommendations about how best to promote positive student outcomes in online environments. As Ruhe and Zumbo (2008) pointed out in a text on evaluating online programs, online courses require continuous improvement based on ongoing assessment (p. 7); they cannot be treated as static because the challenges facing educators are continually changing. "A more definitive understanding of the specific tools and factors [used in distance learning], perhaps identifying the best of the best that exist for certain contexts is still needed" (Menchaca & Bekele, 2008). Indeed, the need for identifying qualities of effective (and ineffective) online instruction has been the focus of more recent studies (see for instance USED, 2009).

Methodology

Using the theoretical framework adapted from Menchaca and Bekele (2008), a correlational study was designed in which pedagogical and course factors were identified within a number of online courses and used to predict student outcomes in those courses.

Predictive and Outcome Variables

A comprehensive review of the literature resulted in four pedagogic factors and seven course factors that multiple theoretical or empirical papers had identified as predicting student success in online courses, though one was ultimately deemed unsuitable for this study. These ten variables served as predictive variables in the study. Additionally, three student outcome factors were adapted from the literature. These three variables served as outcome variables in the study. Chapter Two describes the literature concerning these variables in detail, but they are summarized in the following paragraph.

Menchaca and Bekele (2008) defined pedagogic factors as the "*how* of learning and instruction [*italics added*]" (p. 237). A review of the literature identified four

pedagogic factors that multiple studies suggest affect student achievement.

1. Student directed learning (Alonso, Manrique, & Viñes, 2009; Chetchumlong, 2010; Hiltz & Turoff, 2005; Huuhtanen, Yla-Mella, Jerone, & Keiski, 2008; Kerawalla, Minocha, Kirkup, & Conole, 2008; Precel, Eshet-Alkalai, & Alberton, 2009; Virkus, 2008).
2. Interactivity (Bernard et al., 2009; Brown, 2004; Cameron, Morgan, Williams, & Kostelecky, 2009; Hill et al., 2011). A distinction has been drawn between student-student, student-content, and student-teacher interactions (Arbaugh, 2008; Garrison, Anderson, & Archer, 2000; Gokhale, 1995; Mager, Heulett, & Karvonen, 2011; Moore, 1989; Ozkan, 2010; Picciano, 2002; Rhode, 2009; Swan, 2003) and all three forms of interaction were investigated in this study.
3. Building student capacity (i.e., preparing students for the online environment; Andrade & Bunker, 2009; Brown, 2004; Hill et al., 2011; Huett, Moller, Foshay, & Coleman, 2008; Kitsantas & Chow, 2007).
4. Formative assessment (Chetchumlong, 2010; Hatzipanagos & Warburton, 2009; Hill et al., 2011; Macdonald, 2004; Oosterhof, Conrad, & Ely, 2008; Posner, 2011; Smith & Ragan, 2005; Staker, 2011).

Course factors refer to instructional design elements (Menchaca & Bekele, 2008, p. 237). A review of the literature identified seven course factors that multiple studies suggest affect student success.

1. Measurable learning objectives (Alonso et al. 2009; Oosterhof et al., 2008; Posner, 2011; Smith & Ragan, 2005).
2. Varied teaching activities (Battalio, 2009; Brown-Syed, Adkins, & Tsai, 2005;

Du, 2004; Gainor, Goins, & Miller, 2004; Hsieh & Dwyer, 2009; Jackson & Helms, 2008; Menchaca & Bekele, 2008; Moore, 2002; Neighmond, 2011; Pasher, Mcdaniel, Rohrer, & Bjork 2008; Smith & Ragan, 2005).

3. Varied sensory modalities (Davies & Quick, 2001; Gainor et al., 2004; Ice, Curtis, Phillips, & Wells 2007; Menchaca & Bekele, 2008).
4. Preprogrammed instructor communication (Delaney, Johnson, Johnson, & Treslan, 2010; Heiman, 2008).
5. Synchronous instruction (Bernard et al., 2009; Menchaca & Bekele, 2008; Offir, Lev, & Bezalel, 2007).
6. Immediacy (Delaney et al., 2010; Richardson & Swan, 2003; Swan, 2003).
7. Hybrid instruction (Gainor et al., 2004; Haas & Senjo, 2004; Menchaca & Bekele, 2008; Precel et al., 2009; USED, 2009). However, since the focus of this study is on purely online courses, this seventh factor was not investigated.

As with these pedagogic and course factors, the student outcomes of interest in this study are those related to and measurable within individual courses. There are three of these that have been frequently used in the literature and that this study focused upon.

1. Learning outcomes, defined as both course grades (Ruhe & Zumbo, 2009) and grades on individual assignments (Johnson, Aragon, Shaik, & Palma-Rivas, 2000; Picciano, 2002).
2. Retention in the course (Bernard et al., 2009; Xu & Jagers, 2011a).
3. Student satisfaction, measured by an end-of-course survey (Gozza-Cohen & May, 2011; Ruhe & Zumbo, 2009).

Chapter Two discusses all of these variables in the context of previous studies and

Chapter Three describes operational definitions of these variables for use in this study.

Research Questions

The predictive factors have been distinguished as pedagogic factors (ones related to how instructors teach; the overall philosophical stance that informs their teaching) and course factors (ones related to course design and specific instructor behaviors). However, no studies have assessed either relative importance of these factors or interactions between them. Therefore, this study treats them all as equally likely to promote positive student outcomes; defined as course and assignment grades, course retention, and satisfaction. With the factors and outcomes described previously in mind, the purpose of this study was to identify instructor behaviors that lead to positive student outcomes in online courses.

Four research questions follow from the above:

1. Which of the four pedagogic factors and six course factors predict student grades at the course level?
2. Which of the four pedagogic factors and six course factors predict student grades at the assignment level?
3. Which of the four pedagogic factors and six course factors predict student retention in the course?
4. Which of the four pedagogic factors and six course factors predict aggregate student satisfaction on end-of-course surveys?

Study Design

Once potential predictive and outcome variables were identified from the literature, a study to investigate the relationship between these variables was designed. A

correlational study was performed, using a random selection of Moodle courses from Haywood Community College (HCC), a small community college located in rural Western North Carolina. The author was in a unique position to conduct this research. The author was employed at HCC in both its Research and Institutional Effectiveness office and its Distance Learning office and consequently had access to granular level course and student data *and* instructor evaluation data that other institutions are unlikely to share due to confidentiality concerns. Chapter Three discusses steps that were taken to minimize problems that might result from the author's relationship with the institution.

The available courses in the sampling frame represented a diverse range of subjects and instructional methods (287 course sections from seven semesters; 68 instructors in 40 academic disciplines; over 7000 duplicated students, meaning that one student might be enrolled in multiple classes and is counted independently each time). The average number of students enrolled in a single course was 24.52. Courses were randomly selected from within broader categories of discipline and instructor so that maximum variability in instructor behaviors was achieved. These courses were assessed for the presence of the ten predictive variables listed previously. Outcome variables included the student success measures of grade (both course and assignment), retention, and satisfaction; all three of which were available from the HCC's data warehouse or the Moodle gradebook for selected courses.

Courses were analyzed for the presence of each of the ten predictive variables. Analyses were conducted at the level of the student, the assignment, and the course as a whole. This granular approach to data collection is more precise than studies that have addressed only course grade (e.g., Ruhe & Zumbo, 2009) or student self-ratings of

learning (e.g., Sher, 2004). Additionally, measures of student satisfaction and retention were also used in order to further investigate the relationship between the predictive variables and student outcomes. The researcher counted the occurrences of each of the predictive variables at the course level and, where relevant, the assignment level in 10 randomly selected online courses. Additionally, the number of times that each student experienced certain predictive variable, in reference to both specific assignments and the course as a whole, was recorded.

Pilot Assessment Procedure

Because the nature of this study required careful analysis of a number of courses in order to identify multiple occurrences of ten precisely defined variables, a pilot assessment procedure was conducted. The pilot was used to investigate the validity and reliability of the operational definitions of the predictive variables as used in the literature and to select the operational definitions best suited for use in the Moodle learning management system. After piloting three iterations of the definitions, three reviewers reached consensus on how to identify the presence of each of the predictive variables within Moodle courses. The outcome variables were found to be unambiguous and were pulled from HCC's data warehouse or the Moodle gradebook. The pilot procedure is described in detail in Chapter Three.

Analyses

Statistical analyses were performed in order to determine the extent to which each predictive variable (course and pedagogic factor) predicted each of the four outcome variables and to investigate interactions between them. Because of the nested nature of the data (students nested within courses), statistical analyses required use of both

multilevel analyses (hierarchical linear modeling, HLM) and traditional analyses.

Remaining Chapters

Chapter Two provides a comprehensive review of the literature concerning online instruction, including a description of its rapid growth and the impact of that growth on education in general. It summarizes the kinds of research that have been conducted concerning online learning, including research related to parity of online and face-to-face instruction. From that literature, the Menchaca and Bekele (2008) framework identified previously is adapted and research concerning each of the predictive and outcome variables is described. Finally, the purpose and research questions are developed, based on the literature review that precedes them.

Chapter Three provides a thorough description of the methodology of this study. The sampling procedure and population are described; the operational definitions of each variable are delineated, with reference to both the literature and the pilot assessment procedure; and the pilot procedure itself is fully described. Finally, procedures for data collection and analysis are discussed.

Chapter Four describes the results of the study, in terms of both descriptive and inferential findings. HLM procedures for three outcome variables are described and correlational findings for the fourth outcome variable are described. Analysis limitations are discussed.

Chapter Five discusses the study's findings as they relate to the literature and to the study's research questions. Recommendations for practice and for future research are proposed.

Summary

Distance learning is an increasingly common way of providing instruction (Beaubien, 2002; Farnsworth & Bevis, 2006; Sims et al., 2002; USED, 2009), and it may all but replace face-to-face instruction within the next decade (Christensen, Horn, Caldera et al., 2011; NCCCS, 2010). The rise of distance learning is occurring within a cultural context that has changed not only what students must learn (Batson, 2010; Lumina Foundation, 2011; Symonds et al., 2011), but also how higher education is perceived by students, by the public, and by decision makers (Tulinko et al., 2005). It is therefore necessary that educational leaders understand the factors that promote positive student outcomes in online courses so that they can ensure that courses taught at their institutions do in fact benefit the students who take them. This study provides information that will inform those who create distance learning policy or teach distance learning courses. In particular, it assesses the role of 10 factors that are suggested to improve student outcomes online, but that have not yet been studied at the level of the individual student.

CHAPTER TWO: LITERATURE REVIEW

The Web is what Marshall McLuhan would call a “hot” medium. Web content engages the learner visually, orally, and kinesthetically, and the Web allows a high degree of interaction of the learner with the environment, with other students, and with instructors. It permits both individual and group learning, in both synchronous and asynchronous modes. (Brown-Syed et al., 2005, p. 21)

Despite the "heat" of Web-based learning, and despite the fact that distance learning through technology has become an increasingly common means of instruction (Beaubien, 2002; Farnsworth & Bevis, 2006; Sims et al., 2002; USED, 2009), little research has been conducted on the ways that specific instructor behaviors affect student outcomes. This is an important topic, since according to a National Center for Education Statistics report (2008, p. 3), 94% of institutions surveyed report that they develop their own distance learning courses. At the same time, for most educational administrators, "policy decisions are based on the rationale that providing broader access... may be of more importance than the concerns and perceptions regarding the pedagogical value of online learning" (Picciano & Seaman, 2010, p. 24). Clearly, institutions, students, and faculty would benefit from research that identifies ways of making distance learning more effective. Establishing the relationship of instructor behaviors to student outcomes in online courses would add considerably to the discourse concerning effective distance learning instruction.

Thus, the purpose of this study was to identify instructor behaviors that lead to positive student outcomes in online courses. In this context, "student outcomes" refer to

measures of student learning, perseverance, and satisfaction. The exact nature of these "outcomes" is described in the Outcome Measures section of this chapter.

This chapter describes the rise of online distance learning and depicts this increase as a "disruptive innovation" (Christensen, Horn et al., 2011). It then relates this disruption to additional challenges facing higher education today and applies these multiple challenges to educational leaders. Online instruction can be considered both a contributor to these challenges and a potential solution to many of them. With this in mind, the chapter describes the nature of online distance learning at present and raises questions about how best to use it to address the challenges previously described. It then discusses research into online learning, addressing the question of parity between online and traditional face-to-face learning, and briefly describing additional research topics in online learning. One way to view the many research topics in this area is through a model of factors leading to student success (Menchaca & Bekele, 2008). After describing this model and adopting it as a theoretical framework, the chapter uses that framework as a lens through which to "sort" the literature dealing with online learning. Four pedagogical factors and seven course factors are identified as potential predictors of student success (though one is ultimately discarded). Four outcome measures are identified (though one is ultimately discarded). The chapter concludes with hypothesized relationships between the remaining 10 predictors and three outcomes.

The Rise of Online Distance Learning

Though "distance learning" has a history dating back to the 1700s and the use of correspondence courses, today the phrase usually indicates some use of computer technology and networked applications (Farnsworth & Bevis, 2006; Oosterhof et al.,

2008). For the purposes of this study, "distance learning" is defined as instruction conducted entirely online (with the exception of a student choosing to meet an instructor in person during office hours) and in which students and teacher interact with each other either synchronously or asynchronously. Thus, this study does not address traditional correspondence courses, courses by cassette, automated "learning on demand" courses, and teleconference courses; though all of these methods are sometimes referred to as "distance learning" (Hiltz & Turoff, 2005). Likewise, the term "face-to-face" is used throughout as a synonym for "traditional instruction," "classroom learning" and related terms. Similarly, the term "hybrid" is used throughout as a synonym for "blended;" both terms refer to courses that combine face-to-face and online elements but where online elements make up a significant part of the course activities. (In quotations, the original term is always used.)

Distance learning has become an increasingly common method of delivering educational content to students (Ruhe & Zumbo, 2009; USED, 2009). Since the 1990s, the number of online courses and programs available to students worldwide has increased at a steadily climbing rate (Beaubien, 2002; Farnsworth & Bevis, 2006). To put this rapid growth in perspective, the National Center for Education Statistics (2008) reported that during the 2006-2007 academic year, 66% of 2- and 4-year Title IV degree-granting institutions offered some form of distance learning courses. Reasons cited by these institutions for providing distance learning opportunities included "meeting student demand for flexible schedules, providing access to college students who otherwise would not have access, making more courses available, and seeking to increase student enrollment" (p. 3). Further, many institutions offer entire programs through online

courses. For instance, Githens, Crawford, and Sauer (2010) reported that 47.5% of community colleges offer at least one program primarily or entirely online.

Online learning growth is also shaping K-12 education. A recent *Harvard Education Letter* report describes a "new type of school [that combines] the best of traditional face-to-face instruction with the best of the cutting-edge online curriculum available to virtual schools" (Schulte, 2011, p. 2). Though this study focuses on higher education, the expectations of K-12 faculty, parents, and students cannot be underestimated. Students in this "new type of school" (p. 2) will enter higher education expecting the same kinds of Web-based experiences that they encountered in their earlier educational institutions (Christensen, Horn, Caldera et al., 2011).

The reasons for the rapid growth of distance learning are multifold. One "inexorable force [is] students" (Smith, 2009, p. 7). Baker (2009) surveyed over 4,000 community college students at 16 Maryland colleges and found that over 60% of them perceived increased technology use as a benefit to learning and that over 50% perceived increased technology use as a benefit to communication and collaboration activities. Likewise, a qualitative study by Lim, Dannels, and Watkins (2008) noted that online students were comfortable with the online environment (p. 227), appreciate the "convenience and flexibility" (p. 227) of online courses, like being able to participate in class when they are at their best, rather than when they are "physically, mentally, and emotionally drained after a long day at work" (p. 227), enjoy the fact that they do not have to deal with parking and traffic (p. 227), and prefer to work at their own pace (p. 227). Likewise, a smaller exploratory study of master's degree students found similar student opinions (Martinez, Liu, Watson, & Bichelmeyer, 2006). Increasing demand for

online courses and access to educational technology has resulted in online distance learning that draws upon multiple networked applications in order to meet a variety of student and institution needs.

The "virtual classroom," which now allows students to fully participate wherever and whenever there is a computer, has had a monumental impact on the distance learning format. Through the use of computer groupware, or more popularly, the Internet via the World Wide Web, students enroll in classes using a computer at home, at work, or somewhere on campus to access the course, rather than occupy a desk at a specific time. They register online, download course materials, gain access to video and audio resources, and communicate with both the instructor and other students in the class. (Farnsworth & Bevis, 2006, p. 9)

Regardless of the cause, online distance learning has become a large part of education at all levels. The implications of this rapid rise are manifold, however, and we do not yet understand their full implications.

Online Distance Learning as Disruptive Innovation

The impact of the "virtual classroom" (Farnsworth & Bevis, 2006, p. 9) cannot be overstated. For instance, the North Carolina Community College System estimates that 90% of its courses will be delivered online by 2018 (NCCCS, 2010). Indeed, Christensen, Horn, Caldera et al. (2011) equate the advent of online education to the same "disruptive innovation" that all but eliminated domestic steel production in the U.S. or that changed consumer computers from expensive, place-bound devices to relatively inexpensive mobile ones. "Disruptive innovations fundamentally transform a sector by replacing expensive, complicated, and inaccessible products or services with much less expensive,

simpler, and more convenient alternatives" (Christensen, Horn, & Johnson, 2011, p. 17). In other words, disruptive innovation leads to a substitution model in which a new way of doing something replaces an older way. In this context, online students, courses, and programs are rapidly replacing traditional face-to-face students, courses, and programs. For instance, Only 10% of postsecondary students took an online course in 2002, but 25% did so in 2008, 29% did so in 2009, and it is estimate that 50% will do so in 2014 (Christensen, Horn, Caldera et al., 2011, p. 31). If hybrid learning is the focus, rather than exclusively online courses, this disruptive learning may occur even more rapidly (Nagel, 2011). This impending disruption was noted as early as 2005. Hiltz and Turoff (2005) wrote that "online learning is a new social process that is beginning to act as a complete substitute for [other forms of] distance learning and the traditional face-to-face class. [It] will infiltrate the ordinary face-to-face class and radically change the nature of what is thought of as the typical college course" (p. 60). These trends suggest that the number of face-to-face courses will continue to decrease as online instruction becomes increasingly prominent and that adjustment on the part of educators will be required.

Several factors have led to this disruptive innovation. One such factor is the expectations of K-12 students as they move into college courses. A number of studies have demonstrated the rise of distance and hybrid learning in K-12 environments (Picciano & Seaman, 2007; Picciano & Seaman, 2008; Staker, 2011). As K-12 students in the next decade come to expect hybrid and online educational experiences, the pattern noted in higher education by Christensen, Horn, Caldera et al. (2011) and Hiltz and Turoff (2005) may be even further accelerated. Other researchers have made this same observation. For instance, HP/Intel (n.d.) predicted that students coming from a

background of connected mobile devices will require learning environments that differ from traditional face-to-face ones. Borreson and Salaway (2007) conducted a web-based survey of undergraduate students. Responses from over 27,000 respondents from over 100 institutions were compared with similar data from the preceding three years. As expected, students generally embraced the use of technology. Salient to this study, over half agreed or strongly agreed with the statement "Overall, instructors use [technology] well in my courses"; just 13.6% disagreed (Borreson & Salaway, p. 10). Similarly, students generally expressed favorable opinions of online learning experiences (Borreson & Salaway). As early as 2002, student demand for increased integration between internet technologies and education was noted (Levin & Arafeh, 2002).

Though the studies described in the preceding paragraph report student *opinions*, some case studies have supported these students' perceptions through achievement data or other more objective means. For instance, Fox (2010) describes Roanoke City Schools' use of technology to improve the education that it provides to students. Upon conclusion of the initiative, "the school's dropout rate fell from 27 percent in 2004 to 4 percent in 2009, and its college-going rate rose to 75 percent" (Fox, p. 14). Similarly, Roschelle et al. (2010) illustrated how the use of technological math applications resulted in improved student math skills in a wide variety of Texas middle school settings. These examples show how much Web-based instruction has become a part of education at all levels, often with positive results. However, because the rise of distance learning and related educational technologies has occurred quite rapidly, educational leaders may not know how best to manage the change or use these technologies. To complicate matters, these changes have not occurred in a vacuum.

Educational Leadership and Online Distance Learning

Educational leaders are confronted with a number of challenges related to economic changes, often referred to as the new "knowledge economy" (Batson, 2010; Gordon, 2011). This same observation has been made by influential educational agencies, including the Harvard Graduate School of Education (Symonds et al., 2011) and the Lumina Foundation (2011). "Higher learning has taken on new importance in today's knowledge society. To succeed in the contemporary workplace, today's students must prepare for jobs that are rapidly changing, use technologies that are still emerging and work with colleagues from (and often in) all parts of the globe" (Lumina Foundation, p. 1). It is not surprising that online courses have been viewed as a solution to these challenges, given that both technology and distance are important factors in this new economy. However, as the nature of education and economy has changed, so has that of students and the public at large.

It has not been easy for higher education leaders to manage these changes. "One of the primary challenges facing higher education is the seemingly unending spiral of expectations regarding changes in the ways colleges and universities must operate" (Ulrich, 2009, p. 10). The need to incorporate distance learning into this "spiral of expectations" has been noted in a number of surveys. In 2003, 49% of higher education academic leaders identified online education as critical to long-term strategic planning (Allen & Seaman, 2005, p. 11). By 2005 this percentage increased to 56% (Allen & Seaman, 2005) and by 2010 the percentage increased to 63% (Allen & Seaman, 2010). Despite this focus on *managing* online growth, administrators focus very little on the *quality* of online courses. "Policy decisions are based on the rationale that providing

broader access to a secondary education may be of more importance than the concerns and perceptions regarding the pedagogical value of online learning" (Picciano & Seaman, 2010, p. 24). A National Center for Education Statistics (2008) report identified four reasons cited by institutions that offer online learning opportunities. All of these reasons dealt with increasing student access; none of them dealt with quality education.

Inconsistencies in the quality of online courses, and indeed of education as a whole, has not been ignored by the public in general or students in particular, as a 2005 PBS documentary (Tulinko et al., 2005) dramatically illustrated. For instance, the documentary showed the disdain of many students for rigorous courses and for academic honesty. Likewise, it illustrated a rising belief that college education is of little value. Perhaps in response to such perceptions, educational leaders have sought solutions through various venues, including a large-scale meeting of university presidents (Association of Public and Land Grant Universities, 2010).

While distance learning growth has been fueled by the new knowledge economy, changes in student expectations, and public beliefs about education; a final factor must be noted. The federal government, and by extension many state governments, has created regulations surrounding the use of technology for education and funding structures to promote such use (Nagel, 2010). While such plans will inevitably evolve with changes in administrations, it is unlikely that governing bodies will cease to address the role of technology in education or to promote the use of such technology.

Educational leaders must address a number of concerns, including the rapidly evolving knowledge economy; changing student demographics, especially where technology expectations are concerned; public concerns about the value of education;

student perceptions about the rigor of college courses; and government oversight.

Christensen, Horn, Caldera et al. (2011) suggest that delivery of online courses as an instructional method is increasingly being seen as a solution to these problems. However, online courses are not a "magic bullet" for the challenges facing educational leaders.

As Hill et al. (2011) stated, "we need to build a deeper understanding of how to design effective [distance learning] environments to enhance and extend the learning process" (p. 91). For example, Xu and Jagers (2011b) found that "gatekeeper" courses (introductory math and English courses) are experienced differently by online students than are other courses and that success rates in these courses are lower than research that aggregates across courses regardless of course type would suggest. In a 2009 meta-analysis of instructional and media elements designed to facilitate online student learning, Bernard et al. summarized extant research with three major conclusions, the first of which is that "distance education can be much better and much worse than classroom instruction" (p. 1246). Distance learning is not a singular experience; it can be made highly effective or highly ineffective and educational leaders must be able to shape online courses such that they are effective ones. As alluded to above, Xu and Jagers demonstrated that administrators should treat "gatekeeper" courses differently from other online courses. Likewise, Ruhe and Zumbo (2008) pointed out in a text on evaluating online programs that online courses require continuous improvement based on ongoing assessment (p. 7); they cannot be treated as static because the challenges facing educators are continually changing.

Online courses may address many of the challenges confronting educational leaders. For this to occur, however, educational leaders will need to be familiar with

practices that promote positive student outcomes in online environments. "A more definitive understanding of the *specific* tools and factors, perhaps identifying the best of the best that exist for certain contexts is still needed" (Menchaca & Bekele, 2008).

Identifying such "best practices" through replicable correlational or experimental research would greatly assist educational leaders as they face the challenges described throughout this section. Unfortunately, relatively few studies have addressed this goal.

The Nature of Online Distance Learning

According to a National Center for Education Statistics report (2008, p. 3), 94% of institutions surveyed reported that they develop their own distance learning courses. These institutions would certainly benefit from research that identifies ways of making distance learning more effective. However, there are few studies that these institutions can draw upon. Perhaps as a result of not having access to such empirical research, much distance learning content attempts to replicate the lecture portion of traditional face-to-face courses. Typically, it also replicates traditional assessments (midterm, final, term papers, etc.). Often, distance learning courses merely copy traditional content into a text-based course, in many ways resembling the correspondence courses of the 1700s (Falvo, 2004; Mager, Tignor, & Hipps, 2008). Despite the ability of networked computers to provide a wide range of media and interactivity, many distance learning teachers continue to treat online instruction as a "translation" of traditional face-to-face lectures into a new media (Stevens-Long & Crowell, 2002, p. 153).

This independent study model remains the norm, even though the technological ability exists for students to communicate easily with other people worldwide and to locate and use educational resources that they are not specifically directed to by their

teachers. Many researchers have emphasized the need to move beyond traditional "lecture" approaches (e.g., Falvo, 2004; Palloff & Pratt, 2002). Gillani (2003) stated that the World Wide Web is a highly effective tool not only for content presentation, but also for building collaborative experiences, conducting research, building communities of learners, and facilitating constructivist models of teaching (pp. 9-10). This suggestion echoes students' own beliefs that increased technology use can enable more effective collaboration and learning (Baker, 2009).

A number of theoretical papers have emphasized the fact that the shift to partial or purely online instruction requires a new way of teaching. Hiltz and Turoff (2005), for instance, argued that online learning facilitates a move from objectivist to constructivist teaching and predicted that collaborative online learning (e.g., wikis) will eventually replace traditional course management systems such as Blackboard (p. 61). Other researchers (e.g., Hrastinski, 2006) have made similar predictions. Though these researchers have argued that learning will ultimately improve as a result of this shift, they offered no empirical studies of student learning to support this assertion.

However, these assertions have found support from unexpected quarters. Dougiamas (2010), the founder and lead developer of the Moodle learning management system, developed Moodle according to a constructivist model and sees constructivist teaching as the best way to educate students. For example, he intentionally designed Moodle to resemble common Web 2.0 tools (e.g., blogs and wikis) that give users (students) the ability to create and edit content that is shared amongst entire larger communities by default (Dougiamas, 2010). The increase in the use of Moodle and corresponding decrease in the use of traditional learning management systems echoes

Hiltz and Turoff's (2005) prediction that constructivist teaching would replace traditional course management systems such as Blackboard. For instance, in 2006, the North Carolina Community College System almost exclusively used Blackboard; 52 of its 58 colleges did so. However, by 2011 over half were using Moodle (B. Randal, personal communication, June, 2011). Admittedly, the lower cost of Moodle (an open source product) might be a factor, but a number of these colleges hosted Moodle themselves, which costs more than paying Blackboard, Inc. to host Blackboard (Randall, Sweetin, & Steinbeiser, 2009). To illustrate, a small college would pay approximately \$45,000 for Blackboard hosting, but the estimated cost for self-hosting a comparable Moodle product is \$130,000 (Campbell, 2005).

Though there is some evidence of a shift in online teaching methods, at least as indicated by colleges' choice of learning management platforms, it remains true that the vast majority of colleges develop online courses "in house" (National Center for Education Statistics, 2008, p. 3). What do educational leaders need to know in order to shape online course development? Put another way, what does research tell us about "best practices" in online teaching and learning?

Research on Online Distance Learning

Research into online teaching and learning has taken a variety of tracks, but few studies have attempted to directly associate instructor behaviors with student outcomes. Research into online teaching and learning falls into two broad areas: research addressing parity (or lack thereof) between online and face-to-face instruction and research investigating specific elements of online instruction (e.g., student satisfaction, models of program development, building community).

Parity of Online and Face-to-Face Instruction

Until recently, most research on distance learning focused on establishing it as at least equivalent to traditional face-to-face instruction. These studies have taken a variety of approaches and have defined "effectiveness" in several ways. A discussion of several such studies is illustrative of the nature of this research.

Donavant (2009) used pre- and posttest scores to measure learning by police officers who participated in professional development through either online or traditional methods. Donavant found no statistically significant difference in learning as measured by the pre- and posttest differences as a result of the course format (online or traditional). Likewise, Manochehri and Young (2006) used an end-of-semester knowledge-based comprehensive exam to measure learning by nearly 400 undergraduate college algebra students in online and traditional formats. Manochehri and Young also found no statistically significant difference in learning as a result of teaching method, though they did find that students generally preferred traditional teaching methods to online ones. Similarly, Johnson et al. (2000) used blind judges to evaluate course projects in a graduate level course. Students completed the same projects in either an online or face-to-face course. Analysis of judges' evaluations found no significant differences between the two courses and grade distributions in the two courses were statistically equivalent. Finally, Maki, Maki, Patterson, and Whittaker (2000) conducted a two-year quasi-experimental study of undergraduate students taking online or face-to-face courses in introductory psychology. They found better performance on exams in the students taking the online version of the course.

Some research suggests that students in online courses will outperform face-to-

face students due to the nature of the course. For instance, Reuter (2009) found that students taking online biology courses that required them to complete hands-on labs at home outperformed those who took the same courses in a face-to-face environment. These findings were not unexpected, since students in the online condition were required to perform every aspect of the lab work themselves, while those in the face-to-face condition worked with lab partners and therefore were exposed to less of the course content.

Conversely, some studies have found the opposite results. For instance, Lawrence and Singhanian (2004) compared undergraduate business students' performance in online and face-to-face statistics courses across several semesters and found that online students performed more poorly on multiple choice and problem solving tests. Xu and Jagers (2011a) conducted a 5-year longitudinal study of community college students and demonstrated significantly lower retention and graduation rates in those who took online courses. It should be noted that student learning itself was not measured in this study, though grades (as required for graduation) can be thought of as an indicator of learning. In a similar study, Xu and Jagers (2011b) assessed community college students' success in "gatekeeper" courses (introductory math and English courses) in the Virginia community college system. Here too they found lower success rates for online courses than for face-to-face ones. Gozza-Cohen and May (2011) compared student satisfaction and performance in three types of instruction (online, face-to-face, and hybrid) of education courses. Face-to-face and hybrid students performed better and expressed higher satisfaction than did online students. However, Gozza-Cohen and May did note that the large number of asynchronous discussions used in the online condition, but not in

the other two, may have been a cause of the difference. As Grandzol and Grandzol (2010) illustrated, too much interaction between students may actually hinder learning and satisfaction.

Despite several studies that have found online courses to compare poorly to face-to-face ones, the majority of studies seem to find no difference or a benefit for online or hybrid courses. Several meta-analyses support this conclusion. A 2009 meta-analysis conducted by the USED surveyed research literature from 1996-2008 and screened more than 1,000 studies to identify ones that "(a) contrasted an online to a face-to-face condition, (b) measured student learning outcomes, (c) used a rigorous research design, and (d) provided adequate information to calculate an effect size" (p. ix). Over 90 courses were ultimately utilized in the meta-analysis. One conclusion of the analysis was that "students who took all or part of their class online performed better, on average, than those taking the same course through traditional face-to-face instruction" (p. xiv). These findings were not unexpected. Though preceding the Web-based learning of today, in 1999 Russell analyzed 355 comparative studies and concluded that there was no significant difference between face-to-face and distance learning in terms of student achievement or satisfaction (cited in Bernard et al., 2009). Over a decade later, Swan (2003) conducted a comprehensive review of the literature available at the time and concluded that, in aggregate, there is no significant difference between online and face-to-face courses where student learning is concerned.

Survey research also suggests a general consensus on this question. "The proportion of academic leaders [who] that say online is 'at least as good' – the total of those who rate online as either the same or superior to face-to-face – continues to

increase over time... and now represents just under two-thirds of all respondents" (Allen & Seaman, 2010, p. 10).

Overall, the parity of online and face-to-face instruction has been accepted, and teachers, students, and institutions are increasingly willing to utilize online distance learning to deliver courses, programs and even entire college experiences (Ramaswami, 2009). While questions comparing face-to-face learning to online learning have been addressed, the paucity of literature on effective instructional strategies for online learning suggests that there is a need to develop a research agenda in this area.

Other Online Distance Learning Research Topics

Sims, Dobbs, and Hand (2002) noted the need to "identify critical online learning factors and influences" (p. 135) during the online course development stage. However, their paper was based on theoretical literature and did not concretely specify what these "factors" were. A quick glance at the literature finds that considerable research has been conducted in several areas of online education, but that the foci of these studies is remarkably varied, including diverse topics such as the use of audio versus visual content as synchronous organizers (Astleitner, 2002), organizational structure of community colleges as a predictor of technology use by instructors (Mars & Ginter, 2007), varied teaching methods including the use of humor and multimedia (Lyons, 2004), and the use of blended learning (Precel et al., 2009). By way of example, a single text concerning online learning (*Handbook of Online Learning*, Rudestam & Schoenholtz, 2002) address such diverse topics as presence, critical dialogue, inflammatory email, ethics, complexity, corporate learning strategies, computer-mediated instruction, knowledge communities, case method techniques, and virtual cafes; in addition to many of the topics mentioned

previously (e.g., the rise of distance learning). The research is so varied that a framework through which to organize it is required.

Theoretical Framework: A Predictive Model of Student Success in Online Courses

Menchaca and Bekele (2008) suggest a framework that can be used to organize the sundry distance learning research topics. Building upon a model suggested previously (Bekele, 2008, cited in Menchaca and Bekele), they conducted a multi-year qualitative study of the first five cohorts of students in an online master's program in educational technology. Seventy-two students and six faculty participated in a series of surveys, interviews, and focus groups (p. 238). Data were analyzed using the constant comparative analysis method (p. 231). Menchaca and Bekele identified five kinds of factors that affect outcomes (e.g., learning outcomes, satisfaction, retention) in online learning: human factors (e.g., motivation), leadership factors (e.g., faculty training), technology factors (e.g., internet connection speed), pedagogic factors (e.g., nature and extent of teacher feedback provided to students), and course factors (e.g., teaching activities). They asserted that all five factors interact in order to explain positive student outcomes. In Menchaca and Bekele's view, these factors must be viewed holistically and in combination with each other (p. 249). However, it is also true that there are some factors over which individual instructors have no control (e.g., students' internet connection speed). As this study focused on individual teacher behaviors, only certain factors were of interest. Similarly, while there are many measures of a successful online course or program (Menchaca & Bekele identify seven such factors), only some of these occur at the level of the course.

This study's framework, then, was a predictive one. It assumed that specific

actions taken by instructors (course factors and pedagogic factors) directly affect measurable student outcomes (learning, satisfaction, higher learning abilities, and retention). This is not to say that the other factors are unimportant or that the other outcomes are irrelevant; merely that they are not ones that can be studied at the level of an individual teacher or course.

Menchaca and Bekele (2008) described these factors and outcomes in very broad terms. Their specific examples came from a relatively non-representative sample (master's students in an online educational technology program, p. 232). Thus, though the factor categories that they identified are of use, the particular factors identified by their participants cannot be viewed as definitive. It is therefore necessary to describe findings by other researchers pertaining to these factor categories.

Factors Un-related to Instructor Behaviors

This study focused specifically on the latter two factors (pedagogic and course factors) as the ones that individual instructors can address within a single course. However, copious research has been conducted in all five areas identified by Menchaca and Bekele (2008) and a few pertinent examples of the first three factors are described in the remainder of this section. While such research is informative, it deals only with factors that individual instructors cannot control. Three examples follow.

In the area of *human factors* (e.g., motivation), several researchers have shown that personality factors predict achievement in online courses. Holder (2007) surveyed over 400 online students in order to investigate factors that predicted persistence in online courses. Though the most significant factors were environmental (e.g., employment), personality traits such as "hope" were also found to be statically significant predictors of

persistence. However, personality variables are complex, difficult to assess through quantitative means, and unlikely to be observable in every class. In many cases, they are beyond the ability of individual instructors to affect.

In the area of *leadership factors* (e.g., faculty training), Abel (2005) summarized an Alliance for Higher Education Competitiveness study of 21 institutions and identified 10 causes for the success or failure of online programs. The top four were ones related to college leadership and structure, rather than course or teacher elements: executive leadership and support, faculty and academic leadership commitment, student services, and technology infrastructure (p. 5). Other causes included marketing and financial resources (p. 5). Of the 10 causes listed, only two were those related to teacher behavior (course/instructional quality and "learn-as-you-go" attitude of flexible instruction, p. 5).

In the area of *technology factors* (e.g., internet connection speed), Jackson and Helms (2008) listed "technology problems" (e.g., being locked out of an online assignment) as one of the top two weaknesses of online courses as identified by students (p. 11). Again, these are factors over which individual faculty members have no control.

Factors Related to Instructor Behaviors: Pedagogic and Course Factors

Teachers have a great deal of control over pedagogic factors (e.g., nature and extent of teacher feedback provided to students) and course factors (e.g., teaching activities). It is these areas that this study focused upon. Menchaca and Bekele (2008) defined pedagogic factors as the "*how* of learning and instruction [italics added]" (p. 237) and course factors as instructional design elements (p. 237). However, they described these factors in very broad terms and their specific examples came from a relatively non-representative sample (p. 232). It is therefore necessary to describe findings by other

researchers pertaining to these factor categories. What are the pedagogic and course factors that have been studied? Which ones have been shown to affect student outcomes?

Pedagogic factors. Menchaca and Bekele (2008) defined pedagogic factors as the "*how* of learning and instruction [italics added]" (p. 237) and included the following examples: student focused, collaborative, problem based, interaction, communication, and use of multiple tools. Pedagogic factors are those that inform how teachers operate during the semester, as opposed to particular tools or assignments that they may use (course factors). Other researchers have suggested other pedagogic factors. The pedagogic factors that this study focused upon are those for which multiple studies have provided empirical and theoretical support. These pedagogic factors include constructivist teaching methods, interactivity, building student capacity, and formative assessment.

Student directed learning. Hiltz and Turoff (2005) argued that the advent of Web-based online learning environments would cause a shift from "objectivist, teacher centered pedagogy [to] constructivist, collaborative, student-centered pedagogy" (p. 60). Though they asserted that such a change would facilitate student learning, their paper was purely theoretical and provided no empirical evidence to support this assertion.

However, the increased use of platforms such as Moodle and the corresponding decrease in the use of traditional platforms such as Blackboard do support their prediction in that it suggests an increase in constructivist teaching (Dougiamas, 2010). Indeed, the most recent version of Blackboard (Blackboard 9) introduced a collaborative wiki tool (Blackboard, Inc., 2011), further supporting Hiltz and Turoff's (2005) assertion.

Huhtanen et al. (2008) proposed a redesigned distance learning course in

Engineering Studies based on a constructivist model, but their paper too was theoretical and not based on empirical evidence of student learning. However, some studies have looked for empirical support for the constructivist model. Alonso et al. (2009) randomly assigned students to three conditions in a Java language programming course (traditional, online with virtualized content, online with constructivist teaching). In both learning and satisfaction, traditional and constructivist conditions were similar, and both exceeded the virtualized condition. In a study of formative feedback, Chetchumlong (2010) concluded that the formative feedback condition increased student "achievement by providing a communicative approach to encourage an emphasis on self-study and the constructivist approach to learning" (p. 150). Though not a direct confirmation of constructivist methods, Chetchumlong's study does support the use of individual techniques, such as the formative feedback condition used in the study, that are based upon constructivist principles.

Precel et al. (2009) surveyed over 90 students concerning the course elements that they preferred and that resulted in self-assessed learning in a hybrid course. Students highly rated the constructivist tasks of the hybrid condition (p. 11). It has been noted that Web 2.0 technologies enable constructivist teaching (Ulrich, 2009; Virkus, 2008). At least one study has supported this assertion. In a study of course blogging behavior, Kerawalla et al. (2008) found that more constructivist blogging assignments resulted in the development of a strong collaborative ethic among many students. Though again not tied to learning per se, collaboration is certainly a desired outcome of student-to-student interaction.

It should be noted that "constructivist" was defined in varying ways in these

studies. Given the extensive body of literature on constructivism and its diverse definitions, this study limited its focus to one definition of constructivist teaching: activities in which students direct the learning activities in order to meet student-determined interests and needs. Thus, the predictive variable used in this study is referred to as "student directed learning," rather than "constructivist." However, "constructivist" and "constructivism" are used if such terms were used in the cited source.

Interactivity. Hill et al. (2011) analyzed one of the seminal works in the distance learning field (Moore's *Handbook of Distance Education*, 2007) and noted that "a common theme cutting across all the chapters is the importance of interaction in online learning" (p. 93). According to Bernard et al. (2009), the distance learning "literature is largely univocal about the importance of interaction.... This is because of the integral role that interaction between students, teachers, and content is presumed to play in all of formal education" (p. 1247). They further stated that research supporting the importance of interaction tends to focus on interaction that is symmetrical. Symmetrical interaction is that which is balanced between participants (p. 1248). For example, a prerecorded video from an instructor to students is *asymmetrical* and not considered to lend much to interactivity. On the other hand, email between teacher and student and/or student and student is *symmetrical*, in that communication flows both ways. In a discussion of the various courseware available to build interaction in online environments, Brown (2004) identified a wide variety of studies that demonstrate the role of interactivity in student learning. For instance, Cameron (2003, cited in Brown) created a simulation situation for both online and hybrid courses and compared student performance on problem solving tasks in simulations that were interactive (in the symmetrical sense) and those that were

not. Symmetrical simulations consistently resulted in better student performance.

Though interactivity seems to be important to learning, one might ask, "Interacting with what?" Bernard et al. (2009) performed a meta-analysis designed to address this question. In brief, they identified distance learning courses as facilitating interaction between student-student (SS), student-teacher (ST), or student-content (SC). Seventy-four studies were analyzed for effect size. The findings indicated that SS and SC effects were significantly larger than ST ones, but not significantly different from each other. It may be that ST activities are difficult to consistently implement (Bernard et al., p. 1259) or that assessment of student learning seldom focuses on what they learn from interacting with the teacher, as opposed to what they learn from interacting with the course content, since course outcomes are often standardized across instructors (see for instance the NCCCS *Education Catalog*, 2011, which describes the content of every course taught in that community college system). The same study also found that higher levels of interaction in all three areas predicted student achievement and that higher SC interactions produced the greatest achievement. It should be noted that Bernard et al. used a broader definition of "distance learning" than this study does, including two-way video and other non-Web-based instruction, but many of the studies included in their meta-analysis were online courses of the kind addressed here.

This distinction between student-student, student-teacher, and student-content interactions shows up in other models of interaction as well. Moore (1989) states that research typically focuses on three kinds of activity that affect student learning: interaction with content, interaction with instructors, and interaction among peers. Swan (2003) bridges Moore's theory with Garrison et al.'s (2000) community of inquiry (CoI)

theory. CoI theory posits social presence, cognitive presence, and teaching presence as central to learning (Swan, 2003). Swan suggests that these can be equated with student-student interaction, student-content interaction, and student-teacher interaction, respectively (Swan, p. 4); though Swan admits that the CoI model is more complex than this equation suggests, since "both teachers and students have social presence [and] in many online courses, both teachers and students teach" (p. 4).

CoI theory has received a lot of focus in recent years (Arbaugh, 2008; Swan, 2003; Swan, 2004; Swan et al., 2008), but few studies have tied it directly to student outcomes. For instance, Anderson, Rourke, Garrison, and Archer (2001) focused on validating a tool for assessing teacher presence in online courses, while Rourke, Anderson, Garrison, and Archer (2001) focused on assessing social presence, and while Swan et al. (2008) focused on validating a tool for assessing presence in general. However, one study found that CoI theory was a strong predictor of perceived student learning and satisfaction. Arbaugh (2008) assessed learning in more than 50 online MBA courses using surveys that assessed students' perceptions of CoI elements and of perceived learning and satisfaction with the course. Though the findings offer some validation for the CoI model as a tool for promoting positive student outcomes, it must be noted that Arbaugh's study addressed only *perceived* learning and not an objective measure of achievement.

Though different models treat "interactivity" in varying ways and many of these models have received only very limited empirical study, taken as a whole, "interactivity" as defined by students interacting with each other and with the course content seems to be a strong predictor of course success. Interactivity with the teacher seems to play a less

significant role. Several less broad studies and position papers have supported this conclusion. For instance, Rhode (2009) assessed student perception of various types of interaction in a self-paced online course. Students found interactions with the content more valuable than interactions with the teacher. "Participants hailed the blogging and social bookmarking activities as integral to the quality of the overall learning experience" (p. 13). However, Rhode's study used student responses to Likert type questions; not measures of learning.

Gokhale (1995) and Ozkan (2010) have asserted that group projects (student-student interaction) produce benefits beyond mastering course content. They identify skills such as critical thinking, collaboration, and responsibility (Gokhale, p. 1). However, these papers often relied on studies of face-to-face classes or purely theoretical evidence. Ones that have empirically addressed group activities online find somewhat mixed results.

Cameron et al. (2009) explored the effects of particular social tasks on student perceptions of community in an online course. Generally, students saw group work as a means to a grade rather than a means to community building, and even when they did note the importance of community they still focused on individual, rather than group, efforts. The disconnect between student behavior and beliefs in practice and the benefits of student-student interaction in theory is echoed in a study by Picciano (2002). Picciano evaluated students' perceived participation in online courses and performance in the online courses as a function of their participation in asynchronous discussion boards. (Such participation is largely of a student-student nature.) Students who *perceived* themselves as posting high quality and high quantity posts also *perceived* themselves as

performing well in the online course. However, the relationship between *actual* posting behavior and *actual* course performance was more mixed. There was no significant relationship between actual posting behavior and exam scores, but there was a significant relationship between posting behavior and performance on written assignments, with high quality/high quantity posting predicting better performance on written assignments (p. 12). Picciano did not investigate the reasons for these results, but one interpretation is that a certain level of student-student interaction facilitates learning of objective material and that more interactivity does not necessarily produce higher levels of learning. Sher (2004) investigated the same topic, and found that both student-student and student-teacher interactions predicted student learning and student satisfaction (pp. 102-103). However, Sher measured these outcomes using self-reports of student learning and of satisfaction (p. 43), rather than objective measures of same.

Interactivity is considered a pedagogic factor in Menchaca and Bekele's (2008) model because *how* instructors facilitate student interaction with course content determines the quantity and quality of interactions. For example, Mager, Heulett et al. (2011) found that instructors who require discussion board posting (increasing the amount of student-student interaction) significantly decreased the likelihood that students would earn a C or less in the course. However, too much interaction may be a problem. Grandzol and Grandzol (2010) measured student and faculty time spent in interaction in over 350 undergraduate business courses. Increased levels of interaction, as measured by time spent, decreased course completion rates. These findings are contrary to the some of the interaction research discussed previously (e.g., Gokhale, 1995; Ozkan, 2010), but Grandzol and Grandzol also noted a positive correlation between enrollment size and

participation. They suggested that there may be a point of diminishing return for student participation, so that if participation increases with the number of students in the course, at some level students will cease to see a return on their time investment and become more likely to withdraw. Grandzol and Grandzol's study also supports the findings of Bernard et al. (2009), in that they found no significant relationship between faculty participation and retention. In other words, student-student interaction produced an effect (albeit a detrimental one), but student-teacher interaction did not.

Building student capacity. Another common focus is on the skills required of effective online learners and on ways to promote these abilities early in the semester. Based on experience with international online programs, Fox and Donohue (2006) recommend several steps that instructors can take to prepare students for online learning, including providing "pre-course information; reflective assessments; technology skill building specific to learning online; 'How to learn online' courses, resources and tutorials in context; and 'just in time' strategies to develop the skills progressively as they are needed" (p. 32). Empirical studies have supported these recommendations. For instance, Andrade and Bunker (2009) demonstrated the importance of early development of autonomy and self-regulation in online learners in order to promote effective learning. Drawing from the literature, Hill et al. (2011) suggested that "learner autonomy is chief among the distance design considerations" (p. 97) and proceeded to recommend varying amounts of structure based on learner experience levels as a way of building autonomy (pp. 97-98).

Huett et al. (2008) suggested that since students may not initially possess qualities such as autonomy and metacognition, online instructors can facilitate the development of

these abilities through the use of "supervision, simpler instructions, and a more extensive reinforcement system" (p. 64). Other studies have suggested that certain uses of technology can facilitate this development. Brown (2004), for instance, concludes a review of studies focused on building student interaction with the following: "The research... show[s] technology as a scaffold for continuous learning by making the teacher and class members accessible.... Students use technology functions to... ask questions they would be reluctant to ask [and] clarify thinking" (p. 38). Kitsantas and Chow (2007) found similar results. They showed that students' willingness to seek help is greater in online courses than in face-to-face ones. However, they also showed that there are steps that instructors can take to further increase this behavior. For instance, electronic and asynchronous opportunities to seek help greatly increased students' likelihood to do so and decreased the perceived threat to self-esteem (p. 393). Other factors that promoted help seeking behavior were "convenience (anytime), flexibility (not constrained to office hours), and [timeliness]" (p. 393). While help seeking behavior has not been directly associated with positive student outcomes, it is likely that those who seek help from instructors are more likely to ultimately succeed in a course or activity. Promoting students' willingness to seek teacher assistance can be considered a form of capacity building.

Formative assessment. Perhaps because formative assessment is a widely accepted means of facilitating learning in all venues (Crisp, 2007; Kelly, n.d.; Smith & Ragan, 2005), little research has directly addressed formative assessment in online environments. However, Hill et al. (2011) analyzed Moore's *Handbook of Distance Education* (2007) and noted that formative assessment was a major theme throughout the

text (p. 94). Hatzipanagos and Warburton (2009) conducted research on formative assessment in distance learning, but focused on social media tools suitable for providing such assessment in online environments, rather than on the effects of formative assessment. Likewise, Staker (2011) identified several educators who describe the value of formative assessment in online K-12 environments, but did not include empirical studies that connected such assessment to positive student outcomes.

One study has evaluated formative assessment's role in student learning in an online environment. Chetchumlong (2010) compared student satisfaction with and performance on English courses in two conditions, Web-based formative assessment and paper-and-pencil formative assessment. The former produced significantly higher satisfaction rates, but only modest differences in performance. It should be noted that Chetchumlong's subjects were *not* students in online courses and that *both* conditions used some form of formative assessment, so the effectiveness of an added online tool may demonstrate the value of a specific kind of formative online assessment or it may demonstrate the value of novelty.

Posner (2011) compared a proficiency model of formative assessment to traditional assessment in an introductory statistics course. The groups that were allowed to resubmit assignments for improved grades (those in the formative assessment group) did demonstrate a more positive attitude towards statistics on a survey of course experience than did the those in the traditional assessment group (p. 11). Despite this more positive experience, there was no significant difference in learning, as measured by either a standardized measure (Comprehensive Assessment of Outcomes in Statistics) or by the final exam (p. 11). However, within the formative assessment group, those

students who did choose to resubmit the most often did perform significantly better on both objective measures (p. 11). Again, it must be noted that this study used face-to-face, rather than online, courses. However, it does suggest that students who choose to receive formative assessment are likely to perform better than those who choose not to.

Finally, Macdonald (2004) suggested that the nature of online learning enables a more ready use of formative assessment and that such assessment, assuming that it is well aligned with course objectives, will promote student learning. Though Macdonald's paper was theoretical and did not provide empirical evidence to support its position, it is supported by other recommendations. For instance, Oosterhof et al. (2008) provide recommendations for using instructor, self, and peer means of formative assessment.

Course factors. Menchaca and Bekele (2008) identified course factors as instructional design elements (p. 237) and included the following examples: course organization, relevance to student need, clear goals and expectations, flexibility and "other quality elements" (p. 237). In other words, course factors refer to the actions taken by instructors during the design or teaching of a course, rather than overall pedagogical stances that inform their teaching (though the two are doubtless related). Other researchers have suggested additional course factors. The course factors that this study focuses upon are those for which multiple studies have provided empirical or theoretical support. These course factors include measurable learning objectives, varied teaching activities, varied sensory modalities, preprogrammed instructor communication, synchronous communication, immediacy, and hybrid instruction; though the last will ultimately not be a focus of this study.

Measurable learning objectives. In an evaluation of a modified constructivist

model for course development, Alonso et al. (2009) randomly assigned students to three conditions in a Java language programming course (traditional, online with virtualized content, online with constructivist teaching). The constructivist condition was the focus of the study and was developed using learning objectives. "A learning objective is the specific knowledge that the learner has to acquire about a concept or skill and the tasks to be performed" (Alonso et al., p. 58). In both learning and satisfaction, traditional and constructivist conditions were similar, and both exceeded virtualized. Since the constructivist condition utilized learning objectives, but the other conditions did not, Alonso et al. demonstrated that courses built around specific and measurable learning objectives resulted in higher grades and satisfaction for students in well-designed online courses than in traditional face-to-face courses. In Posner's (2011) comparison of formative assessment to traditional assessment in an introductory statistics course, suggestions based on student feedback were recommended. One was that learning objectives be refined, such that the evaluation of attainment of these skills be more concrete (e.g., grading rubrics were suggested; p. 12). These results are not surprising, as learning objectives are at the heart of instructional design principles (see for instance Oosterhof et al., 2008; Smith & Ragan, 2005).

Varied teaching activities. A number of researchers have noted that the rapid growth of online instruction means that various academic disciplines must find the best ways to provide online instruction to their particular groups of students (e.g., Gainor et al., 2004; Jackson & Helms, 2008). The underlying assumption of such research is that there are stable learner differences (learning styles) that affect the efficacy of online delivery. Smith and Ragan (2005) pointed out that these stable learner differences must

be taken into account when designing effective instruction. They state that instructional design must address student differences such as learning styles. However, "learning style" is a highly nebulous term and, perhaps consequently, the findings from studies on this topic are quite varied.

Several studies have found no effect from learning styles on learning. Brown-Syed et al. (2005) surveyed 108 Library and Information Science students at two universities in order to determine the relationship between learning styles and preference for online versus face-to-face instruction. The study used Felder and Solomon's Index of Learning Styles (ILS), which measures students on four scales: sensory versus intuitive, visual versus verbal, active versus reflective, and sequential versus global (Brown-Syed et al., p. 16). The researchers found no statistically significant differences in learning styles between students who preferred online over face-to-face instruction. (It is worth noting that learning styles were largely consistent within this sample of ILS students, with students strongly preferring sensory and visual learning styles and moderately preferring sequential and active learning styles.) Likewise, Hsieh and Dwyer (2009) used locus of control (internal versus external) as a learning style variable and assessed student satisfaction and improvement under three reading remediation strategies. Subjects were 169 undergraduate students who were randomly assigned to one of four online reading treatments. An insignificant interaction between learning style and treatment type was found, but treatment alone did significantly predict student achievement.

On the other hand, some research has found that learning style does predict outcome, especially when other factors are addressed as well. Du (2004) assessed learning style, student satisfaction with online courses, and computer competency in 237

Library and Information Science students. Learning styles were measured using Kolb's Learning-Style Inventory (LSI), which identifies learning style preferences as assimilating, accommodating, converging, or diverging. Though learning style alone was not a statistically significant predictor of student satisfaction, a statistically significant relationship was found between student satisfaction and course type when computer competence was also a factor. "When the subjects differ with regard to computer competency, there is a difference among learning styles with respect to student satisfaction level" (p. 60). This relationship was not a straightforward one. Subjects who prefer to gain information from concrete experience (accommodating and diverging styles) experienced a higher level of student satisfaction with online courses as their computer competency increased. On the other hand, subjects who prefer to gain information from abstract conceptualization (converging and assimilating styles) were *less* satisfied with online courses as their computer competency increased. It is important to note that no statistically significant difference in student learning was found; statistical differences were found only in regard to student satisfaction.

Other research has supported the importance of learning styles as well. Shu (2005) showed that learning style actually predicted the pattern that learners used in browsing for online information. Similarly, Battalio (2009) used a different conceptualization of "learning style," but found that these styles predicted student grades in online courses. Likewise, in summarizing several studies, Moore (2002) concluded that learning style predicts not only student success, but also student satisfaction, collaborative behaviors, and comfort with particular learning experiences. Clearly, then, learning style is a relevant factor in distance learning course design. Unfortunately, the

term "learning style" is poorly operationalized. In a comprehensive review of the topic, Pasher et al. (2008) noted that over 71 different learning style schemes have been suggested over the years.

Significantly, Pasher et al. (2008) also concluded that there were *no* rigorous studies demonstrating that better learning occurs when teaching to particular learning styles. How can we reconcile these findings with the studies mentioned previously (e.g., Du, 2004; Moore, 2002)? In an interview, psychologist and brain researcher Willingham agreed with Pasher et al.'s findings, but also indicated that the research also points to ways that human learning is common across all learners, providing the example that "Mixing things up is something we know is scientifically supported as something that boosts attention" (Neighmond, 2011). Thus, one reason for findings that support the effectiveness of teaching to various learning styles might be the novelty of varying teaching activities. This is not a new idea. Instructors must either ensure that their teaching strategies can accommodate various types of learners or must make sure that they use several different techniques (Smith & Ragan, 2005, p. 65). Indeed, Menchaca and Bekele's (2008) analysis of longitudinal qualitative data collected from online students noted "how important it was to integrate tools that appealed to multiple learning styles. That is, the *interaction* of tool and strategy was significant" (p. 240). While teaching to particular learning styles has not been shown to be an effective technique, varying teaching techniques has been.

Varied sensory modalities. Perhaps related to varied teaching activities, research has also focused on the benefits of including different sensory modalities (e.g., visual and auditory). While no studies could be found that showed varied sensory modalities as

predicting increased student learning, many were found that predicted both student satisfaction and *perceived* student learning.

Ice et al. (2007) surveyed students who received either voice or text feedback on assignments. Those receiving voice feedback rated themselves as more involved, better able to understand the feedback, and better able to understand the content. However, it should be noted that this study relied on *perception* of learning by students; not upon objective measures. In a case study of a hybrid doctoral program, Davies and Quick (2001) illustrated the importance of using varied sensory modalities to reach students and to build a sense of community between students. In both a discussion of the literature and a detailed description of one student's experiences, Davies and Quick provided a compelling case for the value of varied sensory modalities where student engagement is concerned. However, their research did not address evidence of student learning. Similarly, Gainor et al. (2004) describe the development of a hybrid program in geriatric education and also built a strong case for the importance of varying modalities. However, they too did not provide evidence for student learning as a result of such variety. Likewise, Menchaca and Bekele's (2008) analysis of longitudinal qualitative data collected from online students noted students' preference for a combination of online activities that combine visual and auditory modalities (p. 242).

Preprogrammed instructor communication. Several studies have documented the benefit of instructional design that incorporates frequent communication between instructor and student, typically in the form of email. These are *not* emails in response to student queries, but rather ones that occur according to a designated schedule, either based on date or current course activities/content. For instance, in a study conducted by

Heiman (2008), standardized emails were sent to 229 undergraduate social science students every two weeks. Students who received the emails reported higher levels of perceived academic support and were more satisfied with their courses; though it should be noted that this study did not address student learning. Most research on this topic has used less experimental approaches, however, but still found higher levels of perceived learning and satisfaction in students who received frequent contact from instructors (e.g., Delaney et al., 2010).

Synchronous communication. It has been theorized that synchronous communication will promote positive student outcomes in online learning (Bernard et al., 2009). Bernard et al. conducted a meta-analysis of over 70 studies related to student interaction. One variable they assessed was synchronous vs. asynchronous activities. Contrary to their hypothesis, there was no difference between synchronous, asynchronous, and mixed conditions. However, other studies have suggested a value to synchronous communication.

Offir et al. (2008) found that synchronous learning was a predictor of learning in students with high cognitive ability, but not with those of low cognitive ability. This is to be expected, in light of the interactivity findings discussed earlier. That is, students with high cognitive skills are able to be more participatory and spontaneous yet remain content-focused in a synchronous environment, while those with low cognitive abilities will likely be unable to do so. Student satisfaction also seems related to the inclusion of some synchronous activity. Menchaca and Bekele's (2008) analysis of longitudinal qualitative data demonstrated students' preference for a combination of synchronous and asynchronous communications (p. 242).

Immediacy. Though primarily based on face-to-face research, the role of immediacy in education also supports the importance of synchronous communication, but can also be developed in ways that do not require synchronous communications. Students who are immediately aware of instructors' verbal and nonverbal feedback express greater satisfaction with courses and greater self-reported learning (Richardson & Swan, 2003). In online environments, such immediacy takes on a more social, rather than proximal, form. Behaviors such as "giving praise, soliciting viewpoints, humor, [and] self-disclosure" (Swan, 2003, p. 11) are all perceived by students as increasing the sense of immediacy between teacher and student and this should be true whether or not these behaviors occur online or face-to-face. These behaviors "can lessen the psychological distance between teachers and their students, leading... to greater learning" (Swan, p. 11).

In a large-scale mixed methods study of college student experiences in face-to-face and online courses, Delaney et al. (2010) found support for the importance of immediacy behaviors by instructors. The promptness of instructor feedback and response to questions was identified as very important by respondents, especially in online courses (pp. 50-52). The same study highlighted other behaviors that Swan (2003) identified as indicating immediacy in online environments: constructive feedback (pp. 46-47), attentiveness (p. 47), and humor (pp. 55-57). Though anecdotal, these same observations have been made by practicing online instructors as well (see for instance Donohue, 2007).

Hybrid instruction. One fairly common finding is that students who take hybrid courses (ones combining face-to-face and online instruction) perform better than those taking either purely online or purely face-to-face courses. Several studies have supported

this conclusion (e.g., Gainor et al., 2004). Other studies have suggested that faculty also prefer hybrid instruction to purely online instruction (Haas & Senjo, 2004). A recent meta-analysis (USED, 2009) determined that hybrid instruction is one of the few variables that consistently predicts positive student outcomes in online, face-to-face, and hybrid environments. Menchaca and Bekele's (2008) analysis of longitudinal qualitative data found that students wanted the inclusion of face to face meetings in order to "optimize" their online learning experiences (p. 242). As Precel et al. (2009) stated, "as of today, [hybrid learning] is considered the most effective model for online learning" (p. 3).

Though it is perhaps the best established factor that predicts positive student outcomes in distance learning, hybrid instruction was not a focus of this study. Individual instructors seldom have the leeway to require a face-to-face element in their online courses, and thus seldom have the ability to engage in hybrid instruction unless the college's administration has already identified and advertised the course as a hybrid one.

Outcome Measures

Studies of distance learning have defined outcomes in numerous ways. Menchaca and Bekele (2008) included sustainability and scalability as measures of a program's success (p. 236). Other measures include number of students and revenue generated (Abel, 2005, p. 75), adherence to the college's mission (Abel, pp. 75-76), putting an entire program online (Abel, p. 76), faculty acceptance of online instruction/new technologies (Batson, 2011), positive faculty evaluations (Kelly, 2010), cost efficiency (Bledsoe, 2008), geographical reach (Ruhe & Zumbo, 2009, p. 226), and procedural elements such as the efficiency of the registration process (Martinez et al., 2006, p. 271). These examples were chosen because they illustrate measures that do not represent student

outcomes. However, this study focused on positive *student* outcomes.

Student outcomes in distance learning have been defined in a variety of ways. For example, Rhode (2009) and Sher (2004) used self-reports of learning; Picciano (2002) measured outcomes using grades on written assignments; and Alonso et al. (2009) used a combination of course grades and student satisfaction. In a text on evaluation of distance learning, Ruhe and Zumbo (2009) provide two examples of full-scale evaluations of online programs. The measures used to determine "success" include learner satisfaction, completion rates, and student perception of learning (pp. 227-228), as well as non-student factors such as geographical reach (p. 226).

As with "best practices," outcome measures have been defined in so many diverse ways that Menchaca and Bekele's (2008) framework is useful in order to identify the ones of interest. They identified the following measures of success: learning outcomes, student satisfaction, higher learning, faculty satisfaction, sustainability, scalability, and retention (p. 236). As with the pedagogic and course factors discussed in the preceding sections, the ones of interest here are those related to individual courses. Specifically, four student outcome measures are of interest: learning outcomes (as measured by course and assignment grades), student satisfaction, retention within the course, and higher learning; though the last was ultimately not a focus of this study.

Learning outcomes. Learning outcomes may be defined in a number of ways, such as student learning from course assignments (Johnson et al., 2000; Picciano, 2002), standardized measures (Posner, 2011), end of course exams (Manochehri & Young, 2006; Posner, 2011), pre- and post-test measures (Donavant, 2009), individual assignment grades and regular exams (Lawrence & Singhania, 2004; Maki et al., 2000),

and overall course grades/pass rates (Ruhe & Zumbo, 2009). Though not a perfect measure of learning, grades are often used as an indicator of learning. For instance, Johnson et al. (2000) and Picciano (2002) used assignment grades to indicate learning and Ruhe and Zumbo (2009) used course grades. Both of these measures were used in this study, as described fully in the next chapter.

Student satisfaction. Student satisfaction has usually been measured with some end-of-course survey (see for instance Gozza-Cohen & May, 2011; Ruhe & Zumbo, 2009) or interview/ focus group (see for instance Menchaca & Bekele, 2008). Likert-type questions are also often used for research purposes (see for instance Sher, 2004). In some cases, satisfaction data are collected by institutions as part of ongoing assessment processes (see for instance Ruhe & Zumbo, 2009, p. 216). For this study, a broad measure of satisfaction from ongoing assessment processes was used, as described fully in the next chapter.

Retention. In educational research in general, the focus is typically on retention in a program or in the college (see for instance Berge & Huang, 2004; Lotkowski, Robbins, & Noeth, 2004). However, some research has focused specifically on retention within a specific course (see for instance Park & Choi, 2009).

Because much distance learning research has focused on course-level data, whether or not students remain in a course is the common measure of retention used by researchers of distance learning (see for instance Bernard et al., 2009; Xu & Jagers, 2011a). However, some researchers have extrapolated these measures to an entire program or even to the college as a whole. Menchaca and Bekele (2008) addressed students' likelihood to remain in a purely online program while Xu and Jagers (2011a)

used participation in online courses to predict student graduation rates, even though the student may not have taken exclusively online courses. Since the focus of this study is on course-level outcomes, retention *in the course* will be the measure used, as described fully in the next chapter.

Higher learning. Menchaca and Bekele (2008) identify higher learning as a student outcome predicted by the instructor behaviors that this study focuses upon. However, higher learning has seldom been directly evaluated as to do so requires content experts able to evaluate relevant student assignments. Nevertheless, a few studies have addressed this level of learning with the use of content experts. For instance, Johnson et al. (2000) used blind judges to evaluate course projects in a graduate level course in which students completed the same projects in either an online or face-to-face course. However, the intent of the Johnson et al. study was to evaluate differences in student learning between the two conditions, rather than to assess the level of learning demonstrated by students. Level of learning has largely been ignored in studies of distance learning. Because of its requirement of content experts in any discipline being studied, it was also ignored in this study. Though desirable to evaluate the relationship between teacher behaviors in online courses and higher learning, it was impractical to include this measure in this study.

Purpose of the Study and Research Questions

The predictive factors described previously have been distinguished as pedagogic factors (ones related to how instructors teach; the overall philosophical stance that informs their teaching) and course factors (ones related to course design and specific instructor behaviors). However, no studies have assessed either relative importance of

these factors or interactions between them. Therefore, this study treated them all as equally likely to promote positive student outcomes; as defined as course and assignment grades, satisfaction, and course retention. The purpose of this study was to identify instructor behaviors that lead to positive student outcomes in online courses; we might call such behaviors "best practices." Specifically, the twelve online practices employed by teachers and four outcome measures (assignment and course grade, course retention, and student satisfaction) were studied.

Four research questions were investigated:

1. Which of the four pedagogic factors and six course factors predict student grades at the course level?
2. Which of the four pedagogic factors and six course factors predict student grades at the assignment level?
3. Which of the four pedagogic factors and six course factors predict student retention in the course?
4. Which of the four pedagogic factors and six course factors predict aggregate student satisfaction on end-of-course surveys?

A correlational study was performed, using a random selection of Moodle courses from Haywood Community College, in which repeated teachers or disciplines were replaced with a new random course in order to maximize variability. These courses represent a diverse range of subjects and instructional methods. They were drawn from 287 course sections offered over seven semesters and taught by 68 instructors in 40 academic disciplines. These courses contained over 7000 duplicated students, meaning that one student might be enrolled in multiple classes but was counted independently

each time. The average number of students enrolled in a single course was 24.52.

Courses were randomly selected from within broader categories of discipline and instructor so that maximum variability in instructor behaviors was achieved. These courses were assessed for the presence of the 10 "best practice" factors (recognizing three levels of interactivity, for a total of 12 factors). These factors served as predictor variables. Outcome variables included the student outcome measures of course and assignment grade, course retention, and satisfaction. The following chapter operationally defines these factors and outcomes, as well as describing the correlational procedure in full detail.

CHAPTER THREE: METHODOLOGY

The purpose of this study was to identify instructor behaviors that lead to positive student outcomes in online courses. Specifically, 10 predictive variables and four outcome variables were identified, as summarized in Table 1. Table 1 provides concise operational definitions of each variable, though the evolution of these definitions is described in the Operational Definitions and Pilot Assessment sections of this chapter. It also indicates the level of analysis (level-1 or level-2), which will be referred to in the Data Collection and Analysis section near the end of this chapter. Additionally, it introduces abbreviations for these variables that will be used throughout the remainder of this paper.

Table 1

Concise Operational Definitions by Level of Analysis

Variable	Operational definition
Level-1 Predictive variables	
Student-student interactivity (SSI)	Count of times that student-to-student communication within Moodle occurs
Student-teacher interactivity (STI)	Count of times that student-to-teacher or teacher-to-a specific student communication within Moodle occurs
Formative assessments (FOR)	Count of times that each individual student engages in an assignment which provides evaluative feedback but which either does not count for a grade, counts only for a complete/incomplete grade, or includes opportunities for re-submission for an improved grade
Synchronous communication (SC)	Count of a student's use of real-time communication with instructor or students as part of a given course activity

Immediacy (IMM)	Count of instructor's use of praise, solicitation of viewpoints, humor, or self-disclosure when that behavior is directed at individual students in response to those students' behaviors
Level-2 Predictive variables	
Student directed learning (SDL)	Count of times students actively acquire learning materials from non-course resources of their choice and choose how best to integrate those resources to accomplish a task that does not have a clearly defined objective
Student-content interactivity (SCI)	Count of times that all four of the following elements occur within single activity: (1) triggering event, (2) opportunity for exploration, (3) opportunity for integration, and (4) opportunity for resolution
Building student capacity (BSC)	Count of times that any information or activity is provided by the teacher with the expressed intention of preparing students for online learning
Number of formative assessments (NFOR)	Count of activities in which students receive evaluative feedback but which either do not count for a grade, count only for a complete/incomplete grade, or include opportunities for re-submission for an improved grade
Measurable learning objectives (MLO)	Count of clear descriptions of the skills/knowledge that students will acquire through a given exercise and a clear description of how they will demonstrate that acquisition for the purpose of summative assessment
Varied teaching activities (VTA)	Count of unique teaching activities (e.g., assigned textbook reading will count as one activity, assigned online reading will count as a second)
Varied sensory modalities (VSM)	Count of times that two or more means of communicating (visual, auditory, or tactile) occur within a single course activity
Preprogrammed instructor communication (PIC)	Count of email, message, or announcement communication by instructor to the entire class based on some pre-designated pattern that would be followed regardless of individual student behavior in the course

Level-1 Outcome variables	
Course grade (CGd)	5-point scale (4-0), translated from A, B, C, D, and F; from the college's data warehouse
Assignment grade (AGd)	0-100 grades, from the Moodle gradebook
Retention (RET)	Student earning a letter grade in the course (as opposed to a withdrawal)
Level-2 Outcome variables	
Student satisfaction (SS)	Average of student responses to four 5-point Likert-type questions on end-of-course evaluation, from the college's data warehouse

No studies have assessed either relative importance of these Table 1 variables or interactions between them. Therefore, this study treated them all as equally likely to predict course and assignment grades, retention, and satisfaction.

Four research questions follow from the above:

1. Which of the four pedagogic factors and six course factors predict student grades at the course level?
2. Which of the four pedagogic factors and six course factors predict student grades at the assignment level?
3. Which of the four pedagogic factors and six course factors predict student retention in the course?
4. Which of the four pedagogic factors and six course factors predict aggregate student satisfaction on end-of-course surveys?

The remainder of this chapter describes the research design of this study, including the population and sampling techniques used. It then describes the operational definitions of each of the variables, with reference to both the literature and a pilot assessment that was conducted. Following these definitions, the pilot assessment

procedure is described in detail. The chapter concludes with a description of the data analysis techniques.

Research Design

The study employed a correlational research design using archival data. The data were drawn from archived online courses at Haywood Community College (HCC), a small community college located in a rural region of western North Carolina. HCC was selected as the site for this study because of the author's unique relationship with the institution. The author was employed at HCC in both its Research and Institutional Effectiveness office and its Distance Learning office. Consequently, the author had access to granular level course and student data and instructor evaluation data that other institutions are unlikely to share due to confidentiality concerns. While this relationship might have produced potential ethical considerations, steps were taken to ensure the accuracy of the data (described in the description of the pilot study). Though the author was familiar with some of the online instructors whose courses were evaluated, the extensive pilot study and later "spot checking" using other reviewers resulted in unambiguous operational definitions, helping to ensure objectivity in the collection of data.

Additionally, in keeping with the suggestions of Creswell (2008, pp. 11-12), steps were taken to ensure student and teacher confidentiality. Student identities were masked and neither course names nor disciplines were reported except in the aggregate. Likewise, steps were taken to ensure respect for the research site (Creswell, p. 12). Permission to use this data for this study was obtained from HCC on the condition that the results were also shared with the institution as an abbreviated report.

The courses were all taught using the Moodle 1.9 learning management system (hereafter referred to as Moodle). The availability of courses goes back as far as HCC's use of Moodle; Summer Semester 2009. However, because HCC had a long history of using Blackboard prior to its transition to Moodle, the first year of Moodle courses (Summer 2009, Fall 2009, Spring 2010) was not used for this study, since it was assumed that both instructors and students experienced some confusion during the transition that might confound the outcome measures used in the study. Additionally, courses that were not offered on the traditional 16-week semester were eliminated in order to avoid possible confounds to the outcome measures. After eliminating short-semester courses and ones from the first year of HCC's use of Moodle, a sampling frame of 287 courses remained. The average number of students enrolled in these courses was 24.52. Sample selection is described in the Selection of Students and Population Demographics section of this chapter.

The sample courses were assessed in order to quantify the number of times that each of the predictive variables occurs within that course and, where relevant, within each graded assignment or by each student. Depending on the predictive variable, the level of analysis was the course (e.g., PIC), the assignment (e.g., MLO), or the student (e.g., SSI). This distinction and its implications are discussed in the Analysis section near the end of this chapter. For example, the number of times that preprogrammed instructor communication (PIC) occurred within each course was recorded. PIC is unrelated to particular course assignments, but might occur many times within a course. On the other hand, a number of variables are specific to assignments, such as measurable learning objectives (MLO). Thus, the number of assignments that incorporated MLO was

recorded within each course. The two examples above, PIC and MLO, indicate level-2 variables. That is, these are variables recorded at the course-level; all students in a course experience the same amount of PIC, MLO, and other level-2 variables. However, some variables depend on student behavior, rather than on instructor behavior or assignment design. These are level-1 variables; they are recorded at the level of the student and each student's experience with them is unique. Student-Student interaction (SSI) is one of these. The number of times that each student interacted with another student was recorded for each student in each course. These student-level variables were recorded using Moodle logs. Moodle records the number of times that each student engages in a given behavior, such as communicating with another student. The process by which each variable is quantified is described in detail in the Operational Definitions section of this chapter.

Data were recorded in a spreadsheet application (Microsoft Excel) for the purpose of recording data during the course assessment phase and later for data cleaning before importing it into statistical analysis software packages (SPSS 19 and HLM 6).

Sample and Population

Data were collected in two nested levels: students (level-1) nested within courses (level-2).

Selection of Courses

Courses were the second level of analysis (level-2). They were selected using random sampling. The sampling frame represents a diverse range of subjects and instructional methods. Courses were drawn from 287 course sections offered over seven semesters and taught by 68 instructors in 40 academic disciplines. Only curriculum

courses were used for this study, as continuing education courses typically do not assign grades. Where multiple course sections were combined in a single Moodle shell, the shell was treated as a single course section for the purpose of this study. These courses contained over 7000 duplicated students, meaning that one student might be enrolled in multiple classes but was counted independently each time. On average, 24.52 students were enrolled in each course section.

In order to select courses from the sampling frame, maximum variation sampling was employed. Random sampling by discipline and teacher without replacement was used. However, if the same teacher or discipline was selected as one that had already been chosen for the sample, it was replaced with a new randomly determined course. In one case, a course with only two students was selected and it was replaced by the next randomly generated course. The 30 courses that were ultimately selected came from the following 30 disciplines: Accounting, Art, Biology, Business Management, Communications, Computer Information Systems, Critical Thinking, Drama, Early Childhood Education, Economics, English, Entrepreneurship, Environmental Science, Fish and Wildlife, Forestry, Geospatial Technology, Healthcare Business Informatics, History, Horticulture, Journalism, Marketing, Medical Assisting, Music, Office Systems Technology, Personal Health and Wellness, Psychology, Reading, Religion, Sociology, and Spreadsheet. In total, there were 581 students, the range of students per course was 6 - 59, $M = 19.40$, $SD = 11.29$. From these 30 courses, 880 graded assignments were evaluated for the presence of the predictive variables. Evaluation of behaviors made by these 581 students produced data for five level-1 (student-level) variables: SSI, STI, FOR, SC, and IMM. Evaluation of course activities and design in these 30 courses

produced data for four level-2 (course-level) variables: BSC, VTA, VSM, and PIC. Evaluation of these 880 graded assignments produced data for four level-2 (course-level) variables: SDL, SCI, NFOR, and MLO. Admittedly, there was some overlap. For instance, some graded assignments required SSI and some non-graded activities were formative in nature.

Because of the nested nature of the data (students nested within courses), 30 courses were selected. Thirty level-2 observations, in this study courses, are deemed sufficient for conducting multilevel modeling (Maas & Hox, 2005), though this standard is not universally accepted. For instance, Tabachnick and Fidell (2007) drew on research concerning compliance with analytical assumptions and power differences in first- versus second-level effects to conclude that "the number of groups [should be] twenty or larger" (p. 788). However, Tabachnick and Fidell's recommendation is grounded in theory while Maas and Hox' recommendation is grounded in empirical testing. For this reason, 30 level-2 groups were used to ensure the study was adequately powered. However, being at the cusp of the recommended number of level-2 groups (per Maas & Hox) may mean that type II error was inflated. In other words, some level-2 variables may have been found not to be statistically significant, even though they were predictors of outcome variables. This inflated type II error likelihood is a limitation of this study.

The number of students in each course was less of an issue. A number of researchers have demonstrated that the number of level-1 observations within each group has little impact on statistical power (Jong, Moerbeek, & Van Der Leeden, 2010; Tabachnick & Fidell, 2007). Nevertheless, as mentioned previously, when a course with only two students was randomly selected for inclusion in this study, it was replaced with

another randomly selected course.

Selection of Students and Population Demographics

The subjects of this study were the students who participated in the Moodle courses that constitute the archival data used for this study. Students were the first level of analysis (level-1). Demographic factors are considered human factors, according to Menchaca and Bekele's (2008) framework. Since this study focuses only on course and pedagogic factors, demographic factors are beyond the scope of this study. However, the student population at HCC is quite homogeneous, so it is likely that the subjects resembled the college's overall population. That population is described on the college's Web site (HCC, 2011d): Mean age of 28.6; 60.1% female and 39.9% male; 89.8% white, 5.3% black, 4.9% other; 68.2% residing in Haywood County, NC; and 47.3% employed full-time, 17.9% employed part-time, 34.8% unemployed. The Web site does not provide other statistics that might be of interest, such as standard deviation.

Operational Definitions

Operational definitions were chosen such that they met two criteria. First, that they were objective, unambiguous measures. Second, that they could be determined from within Moodle, without recourse to external applications. An extensive pilot assessment was conducted in order to ensure that the definitions met these criteria. The literature that informed the operational definitions used during the various phases of the pilot assessment is described. Then, the pilot assessment itself is described. Both the literature and pilot ultimately resulted in the definitions seen in Table 1, presented at the outset of this chapter.

Counts of Predictive Variables

Each predictive variable was counted. Thus, each predictive variable was a continuous measure. As discussed in Chapter Four, the fact that each variable's measure was a "true count" determined that it was uncentered (or compared to a reference group of 0) when used in hierarchical linear modeling. Each predictive variable's operational definition is developed in the sections immediately following and each time that definition was met was simply tallied (in relation to students, assignments, or courses) during the data collection phase of this study.

Student Directed Learning

Student directed learning SDL is an aspect of constructivist teaching methods, which have been operationally defined in numerous ways. Alonso et al. (2009) considered activities to be constructivist if they required students to work together and where knowledge construction and reflection were necessary (p. 58). However, this definition requires group activities. Chetchumlong (2010) defined the term more broadly, as "when students construct their own understanding in meaningful ways. [When students] have to change their roles to be [more than] recipients of knowledge" (p. 27).

Thus, adapting from Chetchumlong (2010), for the purpose of this study the predictive variable of **student directed learning** was operationally defined as occurring when students, either individually or in groups, actively acquire learning materials from non-course locations and resources of their choice and when they choose how best to integrate those resources to accomplish some task which does not have a clearly defined objective. For the purpose of clarification, this distinguishes between a traditional research paper, which is largely teacher-initiated (usually limiting acceptable sources and

requiring that the "objective" meet very specific and narrow criteria), and a project in which the student(s) define the objective and may approach that objective in a multitude of ways. In the evaluation of each course used in this study, a *count* of SDL activities was performed.

Interactivity

Interactivity has been defined in numerous ways. One particularly direct definition comes from Sher (2004). "Interaction is defined as mediated communication between student and instructor, or between two or more students, which discusses some aspect of course content, assignment or student progress in the course" (p. 9). Such a definition works for both student-student interaction (SSI) and student-teacher interaction (STI), but student-content interaction (SCI) is more problematic to define. Sher, for instance, defines SCI as "the method by which students obtain information from the course materials. As a result of learner-content interaction, students achieve intellectual growth or change in perspectives" (p. 9).

Swan (2003) equates SCI with cognitive presence in the community of inquiry (CoI) model (Garrison et al., 2000; Swan, p. 4). Thus, the CoI model may suggest an operational definition for SCI. However, "cognitive presence may be the least researched and understood of the three presences" (Swan et al., 2008, p. 3). Though not clearly defined as a measurable concept, in its earliest conception cognitive presence included four phases: triggering event, exploration, integration, and resolution (Garrison et al., 2000, p. 4).

Adapting from Sher (2004), for the purpose of this study the predictive variable of **student-student interaction** (SSI) was operationally defined as student-to-student

communication within Moodle (e.g., within a discussion board). (In order to avoid counting each SSI twice, when associating SSI with a student it was associated only with the *sending* student.) Likewise, the predictive variable of **student-teacher interaction** (STI) was operationally defined as communication from student-to-teacher or teacher to *a particular student or subset of students* within Moodle. Note that communication from teacher to the entire class does not qualify, but is addressed in the operational definition of "preprogrammed instructor communication." It is recognized that because external-to-Moodle activities (e.g., email) may be missed, the count of SSI and STI may in fact be higher than that determined here, but not lower.

Adapting from Garrison et al. (2000), for the purpose of this study the predictive variable of **student-content interaction** (SCI) was operationally defined as any activity presented within Moodle that includes a triggering event (some story, question, or other stimulus that produces a sense of puzzlement), an opportunity for exploration of the trigger (students acquire or exchange information), an opportunity for integration (students connect the ideas and information acquired in phase two), and an opportunity for resolution (students apply the phase three integration to some new situation or stimulus; p. 4). All four elements must be present.

In the evaluation of each course used in this study, a *count* of all three forms of interaction was performed.

Building Student Capacity

Building student capacity (BSC) has been defined in numerous ways. Generally, it takes two approaches, a focus on specific skill building exercises or information (see for instance Fox & Donohue, 2006) or a focus on scaffolding (see for instance Brown,

2004). Since it may be impossible to distinguish scaffolding designed to build understanding of increasingly complex content from scaffolding designed to build student capacity, Fox and Donohue's approach is the easiest one to unambiguously measure.

Thus, adapting from Fox and Donohue (2006), for the purpose of this study the predictive variable of **building student capacity** was operationally defined as any information or activity provided by the teacher with the expressed intention of preparing students for online learning. Two examples suggested by Fox and Donahue are technology skill building exercises and tutorials about how to navigate the course. In the evaluation of each course used in this study, a *count* of BSC activities was performed.

Formative Assessment

Formative assessment has a widely accepted definition (see for instance Ruhe & Zumbo, 2009; Smith & Ragan, 2005). Formative assessment is assessment with the purpose of providing students with information that will help them to improve learning and performance (Ruhe & Zumbo; Smith & Ragan), rather than with the purpose of providing a grade or measuring achievement. Though this definition is relatively simple, it is complicated by two factors. First, teachers often note that students frequently do not take advantage of purely formative opportunities and that some grade, even if simply complete/incomplete, may be required (Mager, Tignor et al., 2008). Second, it could be argued that unit or chapter tests serve as "formative assessment" for final examinations, though this is not the typical definition of formative assessment.

Thus, in keeping with the more traditional meaning of formative assessment, for the purpose of this study the predictive variable of **formative assessment** was operationally defined as activities in which students receive evaluative feedback but

which either do not count for a grade, count only for a complete/incomplete grade, or include opportunities for re-submission for an improved grade (this last alternative suggested by the research of Posner, 2011).

In the evaluation of each course used in this study, a *count* of formative assessment activities was performed. Because the courses analyzed for this study are offered using Moodle, a rich source of data was available. Moodle tracks the number of times that each student visits a given Web page or participates in a given Moodle activity. Thus, as well as measuring the number of formative activities available in each course, the number of times that each student participated in these formative activities was also recorded. Consequently, formative assessment is coded in two ways: FOR (student-level) and NFOR (course-level).

Measurable Learning Objectives

As with formative assessment, measurable learning objectives (MLOs) have a fairly consistent definition in the literature (see for instance Alonso et al., 2009; Smith & Ragan, 2005). "A learning objective is the specific knowledge that the learner has to acquire about a concept or skill and the tasks to be performed" (Alonso et al., p. 58). While some definitions include more detailed specifics, such as the situation under which the knowledge or skills must be demonstrated and the criterion level which performance must meet (Smith & Ragan), the less detailed definition is used most frequently in the literature. However, Posner (2011) found that students benefited not only from a clear description of what they were to learn, but also from a clear description of what it means to demonstrate that learning.

Thus, in keeping with student needs as determined by Posner (2011), for the

purpose of this study the predictive variable of **measurable learning objectives** was operationally defined as a clear description presented within the Moodle course of the skills or knowledge that students will acquire through a given exercise and a clear description of how they will demonstrate that acquisition for the purpose of summative assessment (evaluation for a grade). In the evaluation of each course used in this study, a *count* of times that unique MLOs were articulated by the instructor was performed.

Varied Teaching Activities

As discussed in Chapter Two, at least 70 different definitions of "learning styles" exist in the literature (Pasher et al., 2008), and by extension at least 70 different teaching styles designed to address them. However, and also as discussed previously, the interest here is on *variety* of teaching style, rather than on any particular style type. Thus, a simple count of the *kind of activity assigned to students by teachers* was performed. For instance, assigned textbook reading counted as one activity, assigned online reading counted as a second, assigned research counted as a third, assigned participation in a discussion forum as a fourth, assigned watching of a vodcast as a fifth, etc. While this definition is quite arbitrary, it is the only approach that is sure to capture *all* of the 70+ potential teaching activities that instructors may engage in.

Thus, for the purpose of this study the predictive variable of **varied teaching activity** was operationally defined as the number of different activities assigned to students by teachers. In the evaluation of each course used in this study, a *count* of unique teaching activities was performed. However, because of the limitations of hierarchical linear modeling discussed in the Data Collection and Analysis section that concludes this chapter, assignments were coded as VTA or not, though total VTA occurrences for each

course were used.

Varied Sensory Modalities

Though related to teaching activities, modalities refer only to the *medium* through which students receive or send communication. Typically, research has focused on visual, auditory, and tactile (see for instance Gainor et al., 2004; Ice et al., 2007). Though the courses analyzed for this study are purely online, students may nevertheless be required to engage in visual activities (e.g., watching a film), auditory ones (e.g., listening to a podcast, engaging in online chat), and tactile ones (e.g., conducting lab work or field work). Each course activity was therefore identified as exhibiting one or more of these three modalities (visual, auditory, tactile). However, the focus here was on the use of *varied* sensory modalities. Therefore, the presence of two or more modalities in a single activity indicated varied sensory modalities (VSM). Other modalities (e.g., taste, smell) were not used by online instructors, but would have been counted for this purpose had they been evident. Though text reading is typically visual, it was not counted towards visual modality as it is assumed that every course uses text reading (online or textbook) for most or all of its activities.

Thus, for the purpose of this study the predictive variable of **varied sensory modalities** was operationally defined as the use of two or more means of communicating (visual, auditory, or tactile) within a single course activity. In the evaluation of each course used in this study, a *count* of VSM occurrences was performed.

Preprogrammed Instructor Communication

To distinguish preprogrammed instructor communication (PIC) from immediacy or student-teacher interaction, identifying PIC followed the technique used by Heiman

(2008), in which formulaic and non-individualized email communication was sent to students according to a designated calendar. The qualities of PIC are that it occurs as a result of some designated-in-advance date or event (e.g., "2 weeks into the semester" or "when the midterm project is assigned"), rather than based on student performance, questions, or communications.

Thus, for the purpose of this study the predictive variable of **preprogrammed instructor communication** was operationally defined as email, message, or announcement communication by the instructor to the entire class based on some pre-designated pattern that would be followed regardless of individual student behavior or activity in the course. These communications did not include content that was constantly available to students, such as the syllabus or a course content Web page. Incidents were of PIC were counted, recognizing that because external-to-Moodle activities (e.g., email) may be missed, the count may in fact be higher than that determined here, but not lower.

Synchronous Communication

Synchronous communication (SC) refers to real-time communication. It is clearly present when communication occurs in real time and absent when it does not. However, SC may be optional or required. Since this study addressed *student-level* performance, participation in SC, whether voluntary or required, was counted for each student and activity within the course.

Thus, for the purpose of this study the predictive variable of **synchronous communication** was operationally defined as a student's use of real-time communication with instructor or students as part of a given course activity (akin to a phone call or chat, versus non-real-time communication such as email or discussion boards). However, it is

recognized that because external-to-Moodle activities (e.g., phone) may be missed, some SC activities may not have been counted.

Immediacy

Immediacy (IMM) is easy to measure in face-to-face environments (see for instance Richardson & Swan, 2003), but much more difficult to measure in online ones. Swan (2003) identifies behaviors that indicate immediacy but that do not necessarily have to occur in a face-to-face environment: "giving praise, soliciting viewpoints, humor, [and] self-disclosure" (p. 11). Delaney et al. (2010) found support for this definition from students in online courses. Since immediacy has typically been studied in face-to-face environments, these four behaviors remain the only ones suitable for a measure of immediacy in an online setting. However, since these behaviors are designed to "lessen the psychological distance between teachers and their students" (Swan, p. 11), those that are pro-forma or pre-programmed will not be included. For instance, most discussion board prompts "solicit viewpoints" (Swan, p. 11), but solicitation that is aimed at the entire class will not be counted for this purpose. Only that which is directed at a particular student or subset of students (as when an instructor assigns group work and therefore interacts with that student group) will be counted.

Thus, for the purpose of this study the predictive variable of **immediacy** was operationally defined as the instructor's use of praise, solicitation of viewpoints, humor, or self-disclosure when that behavior is directed at individual students or groups in response to those students' behaviors. Such behaviors directed at the entire class will not be counted. Likewise, such behavior when it is not in response to a given student's (or subgroup of students') specific behavior will not be counted. However, it is recognized

that because external-to-Moodle activities (e.g., email, face-to-face) may be missed, some IMM behaviors may not have been counted.

Learning Outcomes

While it is ideal to measure learning, a rough substitute is to measure grades. For instance, Johnson et al. (2000) and Picciano (2002) used assignment grades to indicate learning and Ruhe and Zumbo (2009) used course grades. Both of these sources of information were available for the purposes of this study. Course grade served as one measure of learning outcomes, taken from Haywood Community College's data warehouse and converted to a 5-point scale (A - F, 4 - 0). Individual assignment grades were available from the Moodle gradebook and were recorded using a 0-100 scale.

Thus, for the purpose of this study **learning outcome** was operationally defined in two ways. The outcome measure of **course grade** (CGd) uses a 5-point scale, translated from A, B, C, D, and F, as taken from the college's data warehouse. The outcome measure of **assignment grade** (AGd) used 0-100 grades, taken from the Moodle gradebook. Both of these measures are ordinal, as we do not assume that a person who earned an F learned nothing or that a person who earned 100 knows twice as much as a person who earned 50.

Student Satisfaction

Student satisfaction (SS) has often been measured with an end-of-course survey (see for instance Gozza-Cohen & May, 2011; Ruhe & Zumbo, 2009) and that is the approach taken here. Because archival data were used for this study, student satisfaction could only be measured at the course-level, rather than the student-level. Haywood Community College conducts end-of-semester evaluations of every course. While many

of the survey items are indicative of effective teaching (e.g., "I had access to course syllabus, either electronically or printed copy;" HCC, 2011c, p. 1), there are four questions that are so general that averaging them allowed for a measure of SS. These items (HCC, 2011c, pp. 1-2) are:

- My knowledge of this area of study has increased.
- Overall, this course was excellent.
- Overall, the instructor was an excellent teacher.
- I would recommend this course to other students.

Thus, for the purpose of this study the outcome measure of **student satisfaction** was operationally defined as the average student agreement on four 5-point Likert-type questions.

This was an ordinal variable. As Creswell (2008) explains, though Likert-type scales are sometimes treated as interval data, this can only be done when procedures such as comparing several like questions and establishing normality are conducted (p. 176), and data to conduct such procedures were not available in this situation.

Retention

Whether or not students remain in a course is the common measure of retention (RET) used by researchers of distance learning (see for instance Bernard et al., 2009; Xu & Jagers, 2011a), and that is the approach taken here. Simply put, students who completed the course and earned a letter grade (A, B, C, D, or F) were considered to have been "retained." Those who withdraw from the course were considered to have been "not retained." (Note that the college distinguishes between drops and withdraws for reasons of financial aid [HCC, 2011a], but this distinction is not relevant here since only students

who remain in a course past the college's census date were included, and these students can only withdraw, rather than drop, from a course [HCC].)

Thus, for the purpose of this study the outcome measure of **retention** was defined as a student earning a letter grade in a course and **lack of retention** was defined as a withdrawal from the course. This was a binary variable.

Pilot Assessment

As illustrated in the preceding discussion, the operational definitions used in this study (Table 1) were informed by the literature. However, they were also tested and revised through the use of a pilot assessment.

Phase 1 - Selection of Operational Definitions from the Literature

Not every source reviewed for this study provided operational definitions of the variables referred to in the source, but for the ones that did the author sought operational definitions that met two criteria. First, that they were objective, unambiguous measures. Second, that they could be determined from within Moodle, without recourse to external applications (e.g., the college's email server). The first criterion is important because the goal of this study is to establish instructor behaviors that predict positive student outcomes, and behaviors are by definition observable and measurable. The second criterion is important because it is not possible to comprehensively acquire evidence from external sources (for instance, if a student visits an instructor's office for assistance, no record is available), and to use evidence from only *some* such sources could potentially confound the results of this study.

With these two criteria in mind, operational definitions from the literature were assessed and ones that fit these criteria were selected for use in phase 2 of the pilot

assessment.

Phase 2 - Operational Definition Testing by the Author

The author used one of the online courses that he teaches for operational definition piloting. Given Creswell's (2008, p. 12) recommendation that all possible measures be taken to ensure honest reporting of data, it was decided that the author should not identify "best practices" in the courses that he teaches for the purpose of this study. Thus, the author's own courses were not deemed suitable for this study and were not included in the 287 course sampling frame. However, they were considered suitable for the pilot phase, since in this phase the course was not comprehensively evaluated for the purpose of statistical analysis. Rather, the applicability of the operational definitions was tested.

The author assessed the first four content areas in the most recent course that he taught and attempted to apply the operational definitions chosen in phase 1 to the activities in the course. These original operational definitions are included in Appendix A.

As expected, the operational definitions of the outcome measures did not require revision, as these are unambiguous numbers drawn directly from the college's data warehouse or the course's Moodle gradebook (e.g., course grade). However, four of the predictive variables were too subjective.

Student directed learning was not defined clearly enough to definitely rule out particular activities. It originally used the phrase "modify their roles as learners," which, though drawn from the literature, was not concrete enough to be identified as clearly present or absent. More precise language was used in Phase 3, as illustrated in Appendix

B.

Building student capacity referred to scaffolding in order to build student cognitive ability, but the author found it too difficult to distinguish between scaffolding for the purpose of learning course content and scaffolding for the purpose of becoming a more effective online student. That element was removed from the definition, as illustrated in Appendix B.

Measurable learning objectives used the more comprehensive definition suggested by Smith and Ragan (2005). This definition seemed too restrictive, so the less comprehensive one suggested by Alonso et al. (2009) was used instead, albeit with the addition of an idea suggested by Posner (2011). The revised definition is seen in Appendix B.

Immediacy did not distinguish between communications aimed at the entire class and those aimed at specific students, though the concept of immediacy clearly refers only to the latter. The definition was adjusted to include this distinction (Appendix B).

After these four operational definitions were revised, usually after further review of the literature, they were then provided to another person with the same level of Moodle access as the author for further testing.

Phase 3 - Operational Definition Testing by Reviewer 1

Two of the author's co-workers were used as reviewers (hereafter referred to as Reviewers 1 and 2). These co-workers were selected because they were both employed in the Distance Learning office and therefore had access to all Moodle courses and the administrative skills needed to access all Moodle course content.

Reviewer 1 assessed the same four content areas that were used in phase 2, using

the revised operational definitions developed during phase 2. He was given very limited information about the nature of the study, but was provided the definitions created in phase 2 (Appendix B) and an Excel spreadsheet upon which to record data (Appendix C). In order to allow for further investigation into discrepancies between his ratings and the author's he was asked to identify where in the course each factor was located, though this step was not a part of the final data collection procedure.

Phase 4 - Consensus

Consensus was very consistent between the author and Reviewer 1. Only two operational definitions were interpreted differently: **varied teaching activities** and **preprogrammed instructor communication**. The author and Reviewer 1 met to discuss all items, in order to ensure that they had interpreted the agreed upon ones the same and to determine new language in order to reach agreement on the remaining two items.

In the first case, the phrase "kinds of activity required of students by teachers" had led to confusion. Reviewer 1 interpreted "required" differently from the author and had not counted any items that were not directly graded. For instance, textbook reading, reading of the syllabus, or watching a video were not identified as meeting this criteria. Language was changed to replace "required" with "assigned" and both parties agreed that this change removed any ambiguity. Phase 5 of the pilot supported this decision.

The second item that resulted in confusion was **preprogrammed instructor communication**. Reviewer 1 interpreted "communication by the instructor to the entire class based on some pre-designated pattern" to include "standing" communications, such as Web pages, the course syllabus, and other content made available at a particular course point or available throughout the semester. This is counter to the intent of

preprogrammed instructor communication as discussed in the literature (e.g., Delaney et al., 2010; Heiman, 2008). Therefore, the definition was changed in order to distinguish between course content—what the student is to learn—and communication designed solely to facilitate that learning. To clarify further, and given the limited ways in which such communication might occur within Moodle, three formats were specified: messages, emails, and announcements. Phase 5 of the pilot supported these decisions.

Phase 5 - Operational Definition Testing by Reviewer 2

Finally, another of the author's co-workers (Reviewer 2) agreed to assess the same four content areas that were used in phases 2 and 3, using the operational definitions that resulted from Phase 4. Again, he was given very limited information about the nature of the study, but was provided the definitions created in phase 4 (Appendix D) and an Excel spreadsheet upon which to record data. In order to allow for further investigation into discrepancies between his ratings and the author's he was asked to identify where in the course each factor was located, though this was not a part of the final data collection procedure.

As anticipated, no discrepancies between Reviewer 2's data and the revised data from phase 4 were noted. The operational definitions were judged as suitable for use in this study.

Phase 6 - Ongoing Validation of Operational Definitions ("Spot-checking")

In order to minimize any concerns about the complexity of the operational definitions or the author's role at the institution under investigation, Reviewer 1 agreed to "spot-check" randomly selected content areas during the course of the study. Phase 6 of the pilot assessment revealed overall consistency in the use of the operational definitions

delineated previously.

Five courses were randomly selected from the 30 courses that were used for this study. In each of these courses, a content area (a block of content within the Moodle site for that course) was randomly selected for assessment by Reviewer 1. The first three courses were assessed by Reviewer 1 and then discussed with the author. Reviewer 1 then assessed the latter two courses. These assessments occurred concurrently with the author's assessment of all elements of all 30 courses selected for this study.

Where the first three courses were concerned, there were several items of disagreement. In all cases but one, discussion between Reviewer 1 and the author resulted in Reviewer 1 agreeing that the item in question did not in fact meet the operational definition. Reviewer 1 admitted that he had relied on his memory from the earlier pilot phases, rather than reviewing the operational definitions, and that this was likely the source of the initial disagreement. For instance, Reviewer 1 included instructor provided tips about course content (e.g., hints about an exam) as a building student capacity (BSC) item, even though BSC's operational definition refers only to preparation for online learning. Notably, most areas of disagreement did not reoccur in the latter two courses.

The notable exception was in two items of student directed learning (SDL), once in the first three courses and once in the latter two. Specifically, even though the instructor of a course had directed students to specific Web sites for use in an otherwise student directed project, Reviewer 1 identified the project as SDL. This is contrary to the operational definition, which includes the requirement that students "actively acquire learning materials from non-course locations and resources *of their choice* [italics added]." Despite discussing this discrepancy following the first three courses, the same

erroneous coding occurred once in the latter two.

In total, there were 13 items of disagreement in the first three courses, but only 1 in the latter two that followed discussion between Reviewer 1 and the author. Because each "item" was a dichotomous decision (the predictive variable judged to be present or absent), agreement on either presence or absence was counted against disagreement of either presence or absence in order to determine interrater reliability. The original data from both Reviewer 1 and the author for this phase of the pilot assessment are provided in Appendix E and the interrater reliability data are provided in Table 2. Interrater reliability in the latter two courses exceeded 90% in all cases and was deemed sufficient to validate the use of the operational definitions used for this study.

Table 2

Pilot Assessment Phase 6 Interrater Reliability

Variable	Interrater Reliability for Courses 1-3	Interrater Reliability for Courses 4-5
Level-1 Variables		
Student-student interactivity (SSI)	100.00	100.00
Student-teacher interactivity (STI)	100.00	100.00
Synchronous communication (SC)	100.00	100.00
Immediacy (IMM)	100.00	100.00
Formative assessment (FOR)	93.33	100.00
Level-2 Variables		
Measurable learning objectives (MLO)	100.00	100.00
Varied teaching activities (VTA)	100.00	100.00
Varied sensory modalities (VSM)	100.00	100.00
Student directed learning (SDL)	93.33	93.75
Preprogrammed instructor communication (PIC)	93.33	100.00
Building student capacity (BSC)	80.00	100.00
Student-content interactivity (SCI)	40.00	100.00

Data Collection and Analysis

All variables were measured through analysis of the courses selected for use in this study or through use of HCC's data warehouse or the Moodle gradebook for the course in question. Data were recorded in a spreadsheet application (Microsoft Excel) for the purpose of recording data during the course assessment phase and for verifying the accuracy of the data (data cleaning) after data collection was complete. Excel is capable of generating basic descriptive statistics, which aided in the cleaning process.

However, the nature of the data in this study required particular statistical

procedures. Some of the data collected for this study did not meet the assumption of independence required for many statistical procedures (e.g., ANOVA).

The assumption of independence means that subjects' responses are not correlated with each other. ...When people are clustered within naturally occurring organizational units (e.g., schools, classrooms...), the responses of people from the same cluster are likely to exhibit some degree of relatedness with each other, given that they were sampled from the same organizational unit. (McCoach, 2010, p. 123)

The nature of this study produced hierarchical relationships within the data. Specifically, students were nested within courses.

For example, imagine that this study found a statistical difference between the course grades (CGDs) of students in Course 1 students in Course 2. Perhaps these differences are due to some of the variables described previously (e.g., SSI or IMM). Since students are exposed to these variables differentially, we might think that this difference can be attributed to some combination of these variables. However, students in Course 1 share an environment that students in Course 2 do not, and vice versa. These environments may explain the CGd difference, instead of the variables that each student experienced. The environmental differences might be other variables described earlier (e.g., BSC or VTA) or they might be variables that this study does not focus upon, such as difficulty of course assignments or late policies of the teachers.

Traditional analyses (e.g., ordinary least squares regression) ignore the nested nature of data and treat all variables as independent. However, the nested nature of these subjects means that variables are not independent. When this fact is ignored, type I error

is inflated (Luke, 2004). When hierarchical relationships within the data are present, "hierarchical linear modeling allows researchers to adjust for and model this nonindependence" (McCoach, p. 123). Hierarchical linear modeling (HLM) is a special type of regression analysis that accounts for nesting of the data and therefore provides a more rigorous measure of statistical significance when nested data are present.

It is useful to think of the predictive variables in terms of level-1 (student) and level-2 (course) variables. Variables that vary between students (e.g., SSI or IMM) are level-1 (student) variables. Variables that only vary between courses (e.g., BSC or VTA) are level-2 (course) variables. Note that formative assessment was measured for both student (FOR) and course (NFOR). Table 1, presented at the outset of this chapter, is organized by level-1 and level-2 variables, as well as by predictive and outcome variables.

Summary

This chapter describes the methodology used for this study, including the development of the operational definitions, the sampling process, and the pilot study. Data were collected by analyzing 30 online courses that were taught using Moodle. In each course, every instance of the predictive variables was counted. The next chapter describes the findings of the study and the analysis of the data.

CHAPTER FOUR: RESULTS

This chapter presents the findings from this research study. In particular, it presents descriptive statistics and exemplars of each of the predictive variables and descriptive statistics for each of the outcome variables. It then describes inferential findings, including those that refer to interactions of the predictive variables. It concludes by discussing analytical limitations of the study.

Descriptive Results

Thirty courses, taught by 30 different instructors in 30 different disciplines, were randomly selected for analysis, as described in Chapter Three. Assignments, student activities, and courses were coded for the presence of each predictive variable. Often, multiple codes applied to the same activity. Outcome variables were retrieved from the college's data warehouse or from the Moodle gradebook. Predictive variables that were unique to each student are considered level-1 variables. Predictive variables that were common within a single course are referred to as level-2 variables. These terms will be used in the Inferential Results section of this chapter and the descriptive results below are organized into level-1 and level-2 variables.

Table 3 provides general descriptive statistics for level-1 (student-level) predictive and outcome variables. Table 4 provides general descriptive statistics for level-2 (course-level) predictive and outcome variables, including aggregates of level-1 outcome variables so that course-level outcomes for these measures are described as well.

Table 3

Descriptive Statistics for Level-1 Predictive and Outcome Variables

Level-1 variable	Number of students	Number of courses	Number of times variable occurred	Range	M ¹	Mdn ²	SD
Predictive variable							
STI	393	30	3,752	0 - 100	6.46	3.00	10.26
SSI	314	20	8,833	0 - 228	15.20	1.00	28.38
FOR	206	14	3,758	0 - 188	6.51	0.00	23.40
IMM	85	21	222	0 - 14	0.38	0.00	1.25
SC	7	1	15	0 - 4	0.03	0.00	0.26
Outcome variable ³							
AGd (out of 100)	581	30	581	0.10 - 99.55	60.41	73.06	31.53
CGd (5-point scale, ignores withdrawals)	440	30	440	0 - 4	2.96	3.00	1.29
RET (binary variable)	440	30	440	0 - 1	0.76	1.00	0.43

¹Because of a large number of outliers, mean is not an accurate descriptive measure. However, it will be referred to in the Inferential Results section later in this chapter, so it is included here.

²Median is reported because these variables were not normally distributed.

³Also presented as a level-2 (course-level) variable in Table 4.

Note. This table includes *all* courses, not just the ones that contained the predictive variable. Descriptive statistics for *only* courses containing these variables are included in the remainder of the Descriptive Results section.

Table 4

Descriptive Statistics for Level-2 Predictive and Outcome Variables

Level-2 variable	Number of courses	Number of times variable occurred	Range	M ¹	Mdn ²	SD
Predictive variable						
VTA	30	714	4 - 73	22.80	18.50	18.44
PIC	28	507	0 - 79	16.90	13.50	19.31
SCI	17	72	0 - 14	2.40	1.00	3.46
VSM	17	165	0 - 32	5.50	1.00	8.07
NFOR	14	155	0 - 67	5.17	0.00	12.68
BSC	11	13	0 - 2	0.43	0.00	0.63
SDL	9	22	0 - 6	0.73	0.00	1.53
MLO	7	68	0 - 19	2.27	0.00	5.49
Outcome variable						
Average AGd for each course (out of 100)	30	30	42.88 - 97.47	63.40	62.01	14.35
Average CGd for each course (5-point scale, ignores withdrawals)	30	30	1.69 - 3.75	2.96	3.19	0.61
RET (% retained in each course)	30	30	33.00 - 100.00	75.73	76.39	18.10
SS (5-point scale)	30	30	2.25 - 5.00	4.27	4.28	0.68
SSdif (SS minus average SS that semester)	30	30	-0.64 - +2.20	0.14	0.13	0.69

¹Because of a large number of outliers, mean is not an accurate descriptive measure. However, it will be referred to in the Inferential Results section later in this chapter, so it is included here.

²Median is reported because these variables were not normally distributed.

Note. This table includes *all* courses, not just the ones that contained the predictive variable. Descriptive statistics for *only* courses containing these variables are included in the remainder of the Descriptive Results section.

Predictive and Outcome Variables

For each predictive variable, an exemplar or description of how coding was conducted is provided and descriptive statistics are used to summarize the data. Exemplars are italicized, indicating verbatim quotations. It is important to note that a single activity might be coded in multiple ways (e.g., an activity that involves student directed learning, student-student interaction, and immediacy). For each outcome variable, descriptive statistics are provided. Later in this chapter, inferential statistics are presented that allow for interpretation of the data.

Student directed learning. Student directed learning (SDL) is a level-2 (course-level) variable because it describes a quality of an assignment. All students who experience such an assignment experience the same level of SDL. An example of an activity coded as student directed learning is provided, with identifying information masked.

Select three (3) current issues in XXX and describe them in detail.

- How did they become an issue and what does the research tell us to do with the issues?

- Use the book and other resources (Internet) to research your three selections.

- This assignment should be at least 400 to 500 words in length.

- Please use your best writing skills with this assignment.

Nine courses included activities that were coded as SDL. There were 22 SDL activities in these nine courses and the number of SDL activities ranged from 1 to 6, $Mdn = 2.00$, $SD = 1.88$. The percent of SDL activities to all graded activities in these courses ranged from 2.33 to 45.45, $Mdn = 6.67$, $SD = 0.15$.

Student-student interaction. Student-student interaction (SSI) is a level-1 (student-level) variable because each student engages in a different amount of SSI. SSI was coded whenever a student communicated specifically with another student. For instance, a post to a discussion board in reply to the instructor's prompt was *not* considered SSI, while a reply to another student's post was. A typical example of SSI activity, with identifying information masked, is:

Hi XXX,

YYY has posted her RAJ a little early, which is fine. You will need to read her entry and respond to it. I have already read hers since hers was the first one posted in our class to get an idea of what your groups RAJ's will look like.

There is a small chance that it should have been posted under "Research a Journal-checkpoint C" link on your small group forum instead of as a post all by itself. Not sure what the Professor will say about that. He has gone over how to post things many times in his materials. May want to look back over how to post a RAJ before you post yours. Once they are posted you cannot remove or delete them.

Good Luck!

ZZZ

Twenty courses included activities that were coded as SSI. A large amount of SSI behavior was not related to graded assignments. For instance, students often "chatted" in the discussion board about other courses or asked questions of each other about course policies, though these interactions were not connected to a graded activity. In total, 314 students exhibited 8,833 SSI behaviors. The number of SSI activities that these 314

students participated in ranged from 1 to 228, $Mdn = 14.00$, $SD = 33.58$.

Student-teacher interaction. Student-teacher interaction (STI) is a level-1 (student-level) variable because each student experiences in a different amount of STI. STI was coded whenever a student communicated specifically with the instructor or when the instructor communicated specifically with a particular student. For instance, an instructor message to the entire class was *not* coded as STI, while an instructor message to an individual student was coded as STI. A sample STI message from an instructor to a student is provided, with identifying information masked.

Dear XXX,

Thanks for your Research a Journal Project. Your learning style application per auditory and intrapersonal in your own life was most interesting. I really agreed with you that having this information about yourself is "advances" us. Nicely done.

The following criteria were presented as a framework for your Research A Journal (RAJ) Project:

- *Research your chosen Journal further. Let your personal interest in this Journal guide your research process. Answer this question, "What additional information, about this cool topic, will be really interesting to my small group?"*
- *Cite at least three outside sources to support your research. Please list these sources. Our textbook can be one of these sources. (I would Google the topic.)*
- *Your project needs to have a two page minimum, using a twelve point font*

and double spaced.

- *Moodle placement – go to the MoodleTab called MY SMALL GROUP, then click on your Small Group.*
- *From the list of Threads, select your assigned Research a Journal A, B, or C and attach a digital copy of your project.*

You did an excellent job of completing these requirements. Please go to peach page one and record your RAJ project by placing a 6 on the Point Tally line. You have “reaped what you’ve sown”, pretty work. Also, your small group will now be able to respond with their Small Group Dialogue.

Remember that you will be doing a Small Group Dialogue for this checkpoint as well. You can respond to your own paper with a “counter point”. If you had a group member do the same checkpoint, you may respond to their RAJ. Or if you’d like you can go to another Small Group Forum and respond to a RAJ from their Group.

In addition to the above response to your work, you can check your grade in Moodle. You’ll see that I’ve recorded your 6 points under the RAJ column.

Thanks again for your work in this course, and pretty work,

Instructor YYY

Every course included activities that were coded as STI. A large amount of STI behavior was not related to graded assignments. For example, instructors often sent messages to students praising them for overall progress or reminding them of course requirements. Likewise, students often messaged instructors to ask about late policies and similar issues. Such interactions were not connected to specific graded activities. In total,

393 students experienced 3,752 STI behaviors. The number of STI activities that these 393 students participated in ranged from 1 to 100, $Mdn = 6.00$, $SD = 11.23$.

Student-content interaction. Student-content interaction (SCI) is a level-2 (course-level) variable because it describes a quality of an assignment. All students who experience such an assignment experience the same level of SCI. An example of an activity coded as SCI is provided.

CHAPTER ONE WEB SEARCH

Conduct a web search on the following topic and report your findings. In the conclusion, write your thoughts about this topic and how it may apply to you.

Please cite your source. Do not use Wikipedia.

- 1. Steps to career planning*
- 2. How negativity affects workplace success*
- 3. Improving self-image*

Seventeen courses included activities that were coded as SCI. Within these 17 courses, there were a total of 72 examples of SCI. The number of SCI activities in these 17 courses ranged from 1 to 14, $Mdn = 3.00$, $SD = 3.67$. The percent of SCI activities to all graded activities in these 17 courses ranged from 2.38 to 71.43, $Mdn = 10.00$, $SD = 0.20$.

Building student capacity. Building student capacity (BSC) is a level-2 (course-level) variable because it describes a quality of a course activity. All students who engage in that activity experience the same level of BSC. It was discovered that the course template used by all online instructors at Haywood Community College (HCC) includes one BSC item. Since that item is required in every course and was the same for every

course, it was not counted. It should be noted that there are no other required elements in the HCC course template. The single BSC item (used to track student attendance for the census date) and availability dates for the course itself are the only default items in the template.

An example of one of the remaining BSC items is a syllabus quiz that included items about succeeding in online courses as well as items specific to the course syllabus. One such item: "If you are having problems, and cannot submit your work on time, explain briefly what you should do." The correct answer includes how to contact the Distance Learning Help Desk and alternate computer lab resources.

Eleven courses included activities that were coded as BSC. Within these 11 courses, there were 13 total examples of BSC. The number of BSC activities in these 11 courses ranged from 1 to 2, $Mdn = 1.00$, $SD = 0.40$. The percent of BSC activities to all graded activities in these 11 courses ranged from 1.33 to 5.00, $Mdn = 3.13$, $SD = 0.01$.

Formative assessment. A typical example of formative assessment was a quiz that students were allowed to take multiple times for an improved score. Two formative measures were recorded. FOR was a level-1 variable and indicated the number of formative behaviors made by each student (e.g., the number of times a student took the same quiz). NFOR was a level-2 variable and indicated the number of formative activities in a course (e.g., a count of formative quizzes available to students).

Fourteen courses included activities that were coded as NFOR. Within these 14 courses, 155 examples of NFOR activities occurred. The number of NFOR activities in these 14 courses ranged from 1 to 67, $Mdn = 8.50$, $SD = 16.97$. The percent of FOR activities to all graded activities in these 14 courses ranged from 2.94 to 117.54 (100 is

exceeded because students participated in a number of non-graded formative activities), $Mdn = 17.84$, $SD = 0.32$.

The number of FOR behaviors made by each student was also recorded. The number of FOR activities made by these students ranged from 1 to 188, $Mdn = 8.00$, $SD = 36.47$.

Measurable learning objectives. Measurable learning objective (MLO) is a level-2 (course-level) variable because it describes a quality of an assignment. All students who experience such an assignment experience the same level of MLO. An example of a measurable learning objective (MLO) was an activity that included a lengthy set of instructions that included not only the rationale for the assignment, but specific instructions for each part of the assignment, including suggested resources. A separate downloadable file included the rubric that the instructor would use to grade the assignment; that rubric included elements that were to be included in the project and requirements such as "Project is neat and uniformly presented, meets outlined requirements, has correct use of spelling, grammar, and punctuation."

Seven courses included activities that were coded as MLO. Within these seven courses, there were 68 examples of MLO. The number of MLO activities in these seven courses ranged from 1 to 19, $Mdn = 7.00$, $SD = 7.83$. The percent of MLO activities to all graded activities in these seven courses ranged from 3.33 to 80.00, $Mdn = 26.09$, $SD = 0.27$.

Varied teaching activities. Varied teaching activities (VTA) is a level-2 (course-level) variable because it describes a quality of a course activity. All students who engage in the activity experience the same level of VTA. Occurrences of varied teaching

activities were counted for each activity in the courses. Where only one activity was identified (e.g., "Read Chapter Five"), VTA was not recorded. Where two or more activities were identified, VTA was recorded. For instance, one assignment required students to type a paper that reflected upon textbook reading and a YouTube video, and was thus coded as VTA since three specific activities were required.

All 30 courses included activities that were coded as VTA. Within the 30 courses, there were 714 occurrences of VTA. The number of VTA activities per course ranged from 4 to 73, $Mdn = 18.5$, $SD = 18.44$. The percent of VTA activities to all graded course activities ranged from 18.75 to 164.29 (100 is exceeded because students participated in a number of non-graded activities, many of which were coded as VTA), $Mdn = 85.14$, $SD = 0.29$.

Varied sensory modalities. Varied sensory modalities (VSM) is a level-2 (course-level) variable because it describes a quality of a course activity. All students who engage in the activity experience the same level of VSM. Occurrences of varied sensory modalities were counted for each activity in the courses. Where only one modality was identified (e.g., listen to a podcast), VSM was not coded. Where two or more modalities were identified (e.g., watch a YouTube video), VSM was coded.

Seventeen courses included events that were coded as VSM. Within these 17 courses, there were 165 occurrences of VSM. The number of VSM activities in these 17 courses ranged from 1 to 32, $Mdn = 7.00$, $SD = 8.64$. The percent of VSM activities to all graded course activities in these 17 courses ranged from 1.75 to 135.71 (100 is exceeded because students participated in a number of non-graded activities, many of which were coded as VSM), $Mdn = 33.33$, $SD = 0.39$.

Preprogrammed instructor communication. Preprogrammed instructor communication (PIC) is a level-2 (course-level) variable because it describes an instructor's behavior. All students in the course experience the same level of PIC. Occurrences of preprogrammed instructor communication were counted. Most PIC activities were not associated with a particular assignment, but instead dealt with general course progress or recommendations. For instance, a typical example of a PIC occurrence was a Moodle announcement programmed to be visible to students at specific points during the course. Two examples, from two different courses, are:

(1) *You will need to have Microsoft Office for this class. Since it is required you can purchase it through the College and those with financial aid can use it to pay for the software.*

(2) *Hope you all are well. Please let me know if anyone is having trouble, I really want you all to succeed!*

Here is a quick reminder about our upcoming week.

First we have EXAM II on March 11-12. This exam will cover chapters 17-20.

However, for the week of March 11-18 we will have NO QUIZ (in observance of spring break). Hopefully, this will give you a little break, and make the exam a little less stressful. It would be great if you went ahead and read Chapter Two1, though.

*We will resume our regular schedule on the week of March 18-25:
Chapter Two2.*

Please let me know if you have questions.

Twenty-eight courses included events that were coded as PIC. Within these 28 courses, there were 507 occurrences of PIC. The number of PIC activities in these 28 courses ranged from 1 to 79, $Mdn = 14.00$, $SD = 19.44$. The percent of PIC activities to all graded course activities in these 28 courses ranged from 3.75 to 658.33 (100 is exceeded because PIC activities are typically ungraded), $Mdn = 44.62$, $SD = 1.31$.

Synchronous communication. Synchronous communication (SC) is a level-1 (student-level) variable because each student engages in a different amount of SC. Only one course provided opportunities for synchronous communication. It used the Moodle chat tool on several occasions to provide "virtual study sessions," but did not require that students participate in these chats. Seven students elected to participate in these SC activities. There were 15 chat interactions engaged in by these seven students. The number of times that a student participated in a synchronous chat ranged from 1 to 4, $Mdn = 2.50$, $SD = 1.21$.

Immediacy. Immediacy (IMM) is a level-1 (student-level) variable because each student experiences a different amount of IMM. Occurrences of immediacy (IMM) were counted. Some IMM activities were related to specific student assignments while others were not. Common IMM examples include praise in response to a student's behavior (often non-specific to an assignment), such as:

I also appreciate having you--and many others like you--who have attempted to do the best work possible. You have accepted my constructive criticism and used it to your advantage. For that, and having you in this class, I thank you.

XXX

Another common IMM activity was the instructor's self-disclosure of interests,

limitations, or history, such as:

I have studied and continue to study semantics. This may be helpful to you, but I am not 100% sure of symbols either--nor are other scholars. Humans are symbolic animals, so we continue to explore the nature of our symbolic actions. Symbols are any human invention or object to which we assign meaning. As suggested this can be an object which we "artifact" that we assign meaning--a "pet rock." We do not make the rock, of course, but we "text" it with meaning. Environments are symbolic as well. Normally when we think of symbols we think of the verbal conventions that we construct such as the word "tweaque." These are abstract and our tools which represent our reality but is not reality itself. As our conventions, these do not carry meaning. We assign meaning to them.

I hope this helps.

XXX

Twenty-one courses included activities that were coded as IMM. A large amount of IMM behavior was not related to graded assignments. For example, students often messaged instructors in order to ask for clarification about a concept in the reading. This often prompted self-disclosure on the part of the instructor, as in the second example above. Within these 21 courses, 85 students experienced 222 IMM behaviors. The number of IMM activities that those students participated in ranged from 1 to 14, $Mdn = 2.00$, $SD = 2.22$.

Course grade. Course grades (CGd) was a level-1 (student-level) variable, since each student earned his or her own CGD. Course grades were based on a 5 point scale (A - F, 4 - 0). The average course grade (excluding W and WF grades) ranged from 1.69 -

3.75, $M = 2.96$, $SD = 0.61$. In the courses assessed for this study, 29.09% of students earned A's, 22.89% earned B's, 11.70% earned C's, 3.96% earned D's, 8.09% earned F's, and 24.27% earned a W or WF (withdrawal).

Assignment grade. Assignment grades (AGd) was a level-1 (student-level) variable, since each student earned his or her own AGd on each assignment. However, it is illustrative to average all AGd values within each course in order to describe courses according to this variable. Both AGd and this "level-2 interpretation" of AGd are described in this section. Though each student's grade on each assignment was recorded, the limitations of HLM analysis (discussed in the Inferential Results section) required that a single assignment grade (AGd) variable be created for each student. As it was not always clear how individual assignments were weighted, a simple average was used. Where AGd is used in this study, it refers to the average of all individually graded assignments for each student in each course, even when a student was assigned a 0 (but not when the assignment was ungraded). It should be noted that because of factors like class participation and late penalties, AGd and course grade (CGd) were not identical, though as illustrated later (Table 6), they are strongly positively correlated, $r = .81$, $n = 440$, $p < .001$, 2-tailed.

There were 880 graded assignments, though some were not associated with any predictive variable while others were associated with more than one. The number of graded assignments per course ranged from 8 to 80, $Mdn = 25.00$, $SD = 19.51$.

The average assignment grade per student (level-1), on a 100-point scale and excluding assignments that were not graded, ranged from 0.10 to 99.55, $Mdn = 73.06$, $SD = 31.53$.

The average assignment grade per course (level-2), on a 100-point scale and excluding assignments that were not graded, ranged from 42.88 to 97.47, $Mdn = 62.01$, $SD = 14.35$.

Retention. Retention (RET) of students in a course was simply the number who received a grade, as opposed to a W or WF. RET is a level-1 (student-level) variable, since each student earns or does not earn a course grade individually. Of the 581 students in this study, 440 were retained (75.73%).

The percent retained per course ranged from 33.00 to 100.00, $Mdn = 76.369$, $SD = 18.19$.

Student satisfaction. Student satisfaction (SS) was calculated by averaging four 5-point Likert-type scales, as described in Chapter Three. Because SS was only available on a per course basis, it was a level-2 variable. The average student satisfaction rating per course ranged from 2.25 to 5.00, $Mdn = 4.28$, $SD = 0.68$.

However, response rates were very low (ranging from 1 to 10 respondents, depending on the course). Response rates were comparable, however, to other courses assessed that semester. For comparison purposes, average course evaluations for each of the 30 courses was compared to the average for *all* courses evaluated that semester by subtracting the average for all courses from the average for each particular course (this variable is abbreviated SSdif). Thus, positive SSdif value indicate a higher than average student satisfaction, while a negative value indicate the opposite. The average SSdif per course ranged from -0.64 to +2.20, $Mdn = +0.13$, $SD = 0.69$.

Comparisons between Courses

Of the 30 courses analyzed, some used several of the predictive variables and

some used very few. Table 5 provides a binary indicator of the presence or absence of each predictive variable *at least once* per course. The number of predictive variables present at least once in a course ranged from 3 to 11, $M = 6.8$, $SD = 1.86$, mode = 6. Figure 1 provides a graphic illustration of the variability across courses.

Table 5

Presence of Each Predictive Variable at Least Once per Course

Course	Level-1					Level-2						
	SSI	STI	FOR	SC	IMM	SDL	SCI	BSC	MLO	VTA	VSM	PIC
1	X	X	X		X	X	X		X	X	X	X
2	X	X			X	X				X	X	X
3	X	X	X		X	X		X		X	X	X
4	X	X			X	X				X	X	X
5	X	X			X					X		
6	X	X			X		X			X		X
7	X	X					X			X		X
8		X	X		X					X		X
9	X	X	X			X	X	X		X	X	X
10		X								X	X	X
11		X	X				X			X	X	X
12		X	X				X		X	X		
13		X	X		X		X		X	X	X	X
14	X	X	X				X	X	X	X		X
15	X	X			X	X	X			X	X	X
16		X	X		X			X		X	X	X
17	X	X	X							X	X	X
18	X	X			X					X	X	X
19	X	X	X		X		X			X	X	X
20		X			X		X	X		X		X
21	X	X			X	X	X			X	X	X
22		X								X		X
23		X			X		X			X	X	X
24	X	X			X	X	X			X		X
25	X	X	X		X		X	X	X	X	X	X
26	X	X	X	X	X		X	X	X	X	X	X
27		X			X			X	X	X		X
28	X	X			X			X		X		X
29	X	X	X		X			X		X		X
30	X					X	X	X		X		X

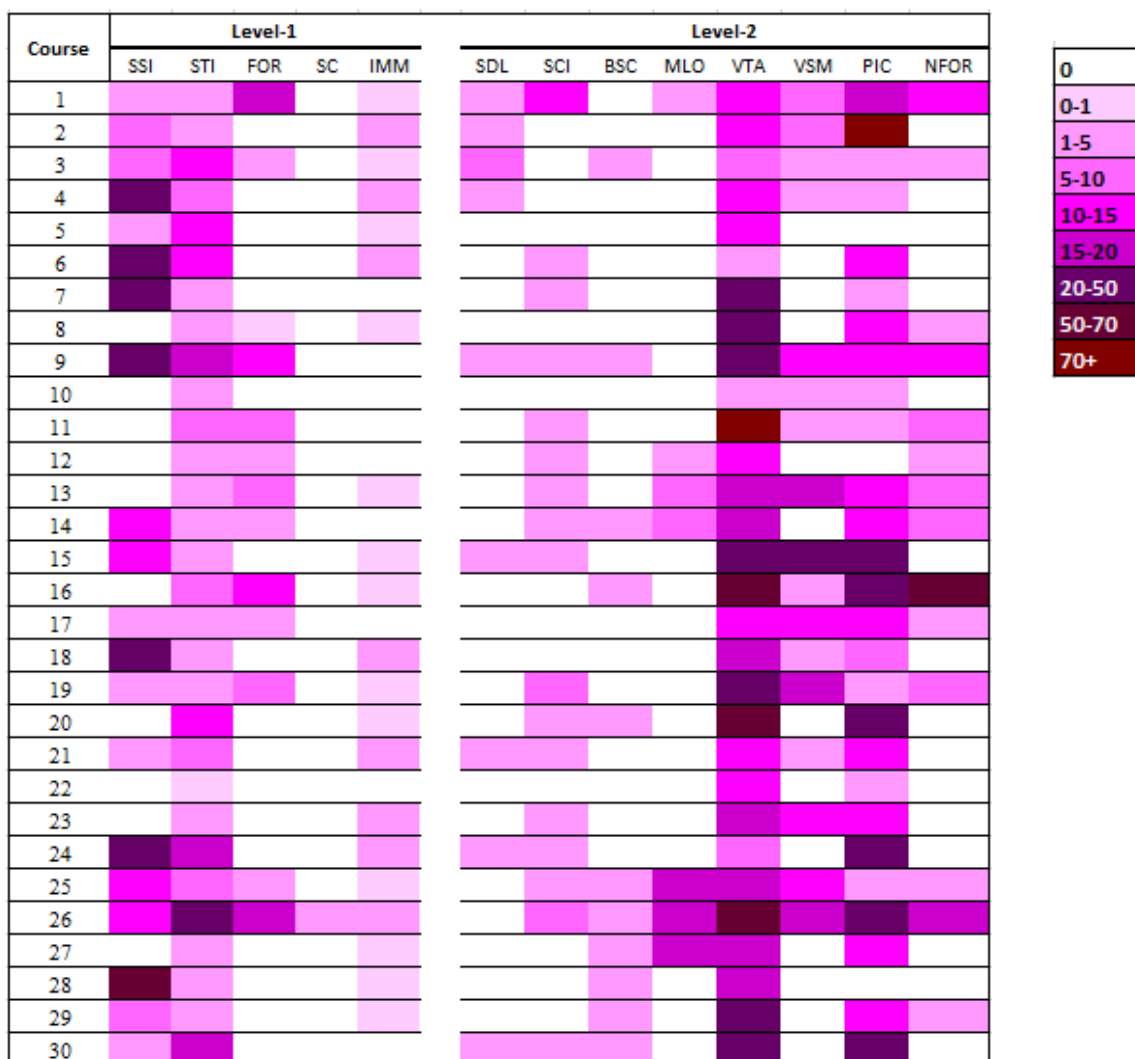


Figure 1. Variability in the presence of each predictive variable across courses. Level-1 variables show the average number of occurrences per student within course. Level-2 variables show the total number of occurrences within course.

Intercorrelations Among Predictive and Outcome Variables

Intercorrelations among predictive variables and among outcome variables will be important in the Inferential Results section that follows and in Chapter Five. This is because intercorrelations among predictive variables were used in developing the final HLM models used to predict CGd, AGd, and RET, as described in the Inferential Results section later in this chapter. Thus, before addressing inferential findings, it is important to clearly describe the intercorrelations among variables. There were statistically significant correlations between some of the predictive variables (Table 6).

Table 6

Intercorrelations for All Predictive Variables

	1	2	3	4	5	6	7	8	9	10	11	12
1. SSI												
2. PIC	.435***											
3. BSC	.184**	.353***										
4. STI	.116**	-.064	.160***									
5. IMM	.078	.014	-.159***	.359***								
6. SDL	-.022	-.039	-.015	.210***	.169***							
7. VSM	-.048	-.046	-.205***	.104*	.116**	.007						
8. FOR	-.049	.005	.332***	.131**	-.034	-.042	.026					
9. VTA	-.090*	.01	.257***	.144***	-.064	-.273***	.149***	.399***				
10. NFOR	-.106*	-.041	.402***	.131**	-.071	-.039	.064	.801***	.533***			
11. SCI	-.134***	-.151***	-.253***	.128**	.061	.164***	.254***	.099*	.120**	.202***		
12. MLO	-.138***	-.092*	.285***	.085*	.021	-.181***	.032	-.029	.044	-.018	.046	

* $p < .05$ (2-tailed). ** $p < .01$ (2-tailed). *** $p < .001$ (2-tailed).

Some statistically significant intercorrelations were in expected directions. For instance, the number of formative activities engaged in by students (FOR) and the number of formative opportunities in courses (NFOR) were significantly and positively correlated, $r = .80$, $N = 581$, $p < .001$, 2-tailed (Table 6). On the other hand, other correlations were not expected because they are conceptually unrelated, such as the number of preprogrammed instructor communications (PIC) and the number of student-student interactions (SSI), $r = .44$, $N = 581$, $p < .001$, 2-tailed (Table 6).

There were also statistically significant intercorrelations among outcome variables (Table 7). Specifically, AGd was significantly correlated with each other outcome variable, but the other variables were not significantly correlated with each other. In other words, AGd (average course grade) was significantly correlated with CGd (course grade), with RET (retention), and with SS (student satisfaction); all significant at the .001 level (2-tailed). These were the *only* statistically significant correlations among outcome variables.

Table 7

Intercorrelations for All Outcome Variables

Measure	1	2	3	4
1. CGd				
2. AGd	.81*			
3. SS	.03	.15*		
4. RET	.00	.74*	.07	

* $p < .001$ (2-tailed).

Inferential Results

Hierarchical linear modeling (HLM) was used to analyze the data collected during the course of this study. HLM is a type of regression analysis that is specifically designed to analyze nested data (Luke, 2004, McCoach, 2010). HLM is used when environmental influences may influence outcomes. When students are nested within courses, observations cannot be fully independent since they occur within online courses where individual variation between students is somewhat explained by course differences. When data are nested, an assumption of ordinary least squares regression is violated: that individual observations are independent of each other (Luke). HLM analysis accounts for the nested nature of the data and allows for analysis of individual observations that are nested within larger groups (Luke). HLM analyses were performed using HLM 6 software (SSI Inc., n.d.a). HLM analyses and findings for three outcome variables (CGd, AGd, and RET) are described in the Research Question 1-3 sections that follow. The fourth outcome variable (SS) was not suitable for HLM and is discussed in the Research Question 4 section.

Research Question 1: Course Grade

Research question 1 was "Which of the four pedagogic factors and six course factors predict student grades at the assignment level?" In order to answer this question, course grade (CGd) was analyzed using HLM. Four models were ultimately required in order to draw conclusions about the variables that predict course grade. The evolution of these models is described fully in the Model sections that follow. Table 8 presents statistical findings for all four models.

Table 8

Fixed Effects Estimates for Models of the Predictors of Course Grade

Parameter	Model 1	Model 2	Model 3	Model 4
Intercept	2.91*** (0.11)	2.35*** (0.18)	2.41*** (0.18)	2.61*** (0.21)
Level-1 (student)				
SSI		0.02*** (0.00)	0.02*** (0.00)	0.02*** (0.00)
STI		0.01* (0.00)		
FOR		0.01*** (0.00)	0.01*** (0.00)	0.01*** (0.00)
IMM		0.14* (0.06)	0.17** (0.05)	0.16** (0.05)
Level-2 (course)				
SDL				
SCI				
BSC				-0.47* (0.34)
NFOR				
MLO				
VTA				
VSM				
PIC				
	Variance Components			
Within-course	1.40	1.10	1.05	1.05
Between (intercept)	0.27	0.83	0.90	0.88
-2 log likelihood	1435.10	1389.19	1385.53	1381.69

Note. For fixed effect portion of the table, standard errors are in parentheses.

* $p < .05$. ** $p < .01$. *** $p < .001$.

Model 1: Mean course grade. The first model was the null model (or one-way ANOVA). This model provides an estimate of overall grand mean CGd across all students. In its reduced statistical form, the null model is expressed as:

$$\text{CGd} = y_{0c} + u_0 + r \quad (1)$$

where CGd represents course grade for a given student in a particular class, y_{0c} represents overall grand mean course grade, u_0 represents residual prediction error

associated with level-2 variables, and r represents residual prediction error associated with level-1 variables. Throughout the study, robust standard errors are reported in all cases. In Model 1, mean course grade, or "intercept" in Table 8, is 2.91.

Model 1 is the null model; it assumes that there is no predictive value to adding either level-1 or level-2 variables. There is only one fixed effect estimate (the intercept, 2.91), which is interpreted as the average value of the outcome variable CGd across all students. In other words, the average course grade is 2.91, or a C+. (This differs slightly from Table 4's 2.96 due to rounding errors. Table 4 uses the mean for each course to calculate the average, but Table 8 averages individual students' grades.)

The null model may also be used to provide a statistical (as opposed to theoretical) rationale for the use of HLM for analyzing data. This is because the null model can be used to calculate the intraclass correlation coefficient (ICC), which identifies the proportion of variance occurring within each level in a nested data set (Raudenbush & Bryk, 2002). The results generated by Equation 1 allow for ICC calculation using the following equation:

$$\rho = \frac{\tau_{00}}{\sigma^2 + \tau_{00}} \quad (2)$$

where ρ represents the ICC, τ_{00} represents the variance of the level-2 residuals u_0 , and σ^2 represents the variance of the level-1 residuals r . For Model 1, ICC is calculated as 16.03%. According to Raudenbush and Bryk, when the proportion of variance attributable to grouping is greater than approximately 10%, HLM analysis is appropriate.

The variance of CGd within courses is 1.40 and between courses is 0.27 (Table 8), so a larger portion of variability in CGd lies within courses. The ICC was 16.03%, indicating that 16% of the total variation in CGd is accounted for by differences across

courses, as opposed to within specific courses. It is important to note that this does not mean that this 16% is due solely to the level-2 variables used in this study (or that the 84% remaining is due solely to the level-1 ones). It is possible, and indeed likely, that some of the variance is explained by variables that this study did not address.

Regardless, the ICC findings for CGd also provide a statistical rationale for the use of HLM, which lends support to the theoretical rationale discussed at the outset of the Inferential Results section.

Model 2: Mean course grade with level-1 controls. The second model was an alternative model that tested whether inclusion of level-1 variables improved the predictive value of Model 1. Model 1 provided an estimate of overall grand mean CGd across all students. The second model planned to add five level-1 (student) variables to Equation 1 at level-1. In other words, it tested an alternative model that included all level-1 predictors, assuming that doing so would increase our ability to predict CGd. Essentially, this model would provide an estimate of CGd across all 30 courses controlling for the influence of five level-1 variables. However, synchronous communication (SC) was only present in one course and therefore was unsuitable for inclusion as a predictor in the HLM analysis. It was deleted from all models in this Research Question section and in the Research Question 2 and 3 sections, leaving only four level-1 variables for inclusion. No level-2 variables were included in Model 2. In its reduced statistical form, Model 2 is expressed as:

$$\text{CGd} = y_{0C} + y_{1C} * \text{SSI} + y_{2C} * \text{STI} + y_{3C} * \text{FOR} + y_{4C} * \text{IMM} + U_0 + r \quad (3)$$

where $y_{1C} * \text{SSI}$ represents the effect of student-student interaction on course grade,

$y_{2C} * \text{STI}$ represents the effect of student-teacher interaction on course grade, $y_{3C} * \text{FOR}$

represents the effect of formative behaviors on course grade, and $y_{4C} * IMM$ represents the effect of immediacy on course grade. Because all of these variables were true counts, they were entered into Equation 3 uncentered.¹

Model 2 tested an alternative model that included four level-1 predictors. One measure of how well a model fits the data is the extent to which it explains variance within and between groups. While true R^2 cannot be obtained in HLM, there are at least two ways of calculating pseudo R^2 (Division of Statistics and Scientific Computation, College of Natural Sciences, n.d.). In this study, pseudo R^2 was calculated using this formula: $(unrestricted\ error - restricted\ error) / unrestricted\ error$, where unrestricted error refers to the null model and restricted error refers to the alternative model (Kreft & de Leeuw, 1998; Singer, 1998; both cited in Division). In Table 8, these terms are indicated by the "Within-course" and the "Between (intercept)" variance components. This formula may be applied to either within group variance or between group variance, though when random intercepts are present, level-1 variance may be larger in the alternative model than in the null model, meaning that a negative value is calculated and pseudo R^2 cannot be determined (Division). Using the formula above, pseudo R^2 within groups was .21 and between groups was a negative value. In other words, Model 2 explained 21% of the variance in CGd within courses (between students). However, the amount of variance in CGd between courses explained by Model 2 could not be determined.

Model 2 introduced four level-1 predictors. The intercept (2.35) is no longer a mean, but instead represents the predicted CGd if all predictor values are 0. In other

¹ "Uncentered" means that 0 is used as the reference group, as opposed to using the mean of the group or of all groups.

words, if no SSI, STI, FOR, or IMM occurred in the course, the average CGd is predicted to be 2.35 (letter grade C, on a 4-point scale). The effects estimates provide a prediction of the relative contribution of each of the level-1 variables. For instance, for each instance of student-student interaction, we expect CGd to increase by .02. While all four variables are statistically significant predictors, IMM is predicted to have the strongest relationship with CGd (Table 8).

In summary, by including level-1 variables, Model 2 is a more effective predictive model than Model 1, which did not include them. Additionally, all four level-1 variables were statistically significant contributors to this improved model. However, when level-2 variables were introduced (Model 4), it became clear that Model 2 could be adjusted to better fit the data. Model 3 describes these adjustments.

Model 3: Mean course grade with final level-1 controls. The third model was an alternative model that tested whether a reduced number of level-1 variables improved the predictive value of Model 2. Model 3 added three level-1 (student) variables to Equation 1 at level-1: SSI, FOR, and IMM. Model 3 simply adjusted Model 2 based on correlations and model fit statistics. The resulting Model 3 resembled Model 2, except that the STI variable had been removed. These changes resulted from the development of Model 4, explained in greater detail below. Where Model 3 is concerned, however, there were two primary reasons for these adjustments. The first is statistical and the second is conceptual.

Statistically, a model with smaller deviance than Model 2 was sought. "For each model, a deviance statistic, equal to $-2 \ln L$ for that model, is computed. The deviance can be regarded as a measure of lack of fit between model and data" (SSI Inc., n.d.b). Smaller

deviance indicates a better fit to the data than larger deviance. Thus, each subsequent model that did not result in a lower deviance was rejected.

Two level-1 variables were strongly correlated (Table 6): student-teacher interaction (STI) and immediacy (IMM), $r = .36$, $N = 581$, $p < .001$, 2 tailed. Conceptually, IMM cannot occur without STI, since immediacy requires student-teacher interaction. However, the reverse is not true, since many examples of STI do not contain IMM. Because IMM was associated with fewer students, allowing for a higher likelihood that it could be linked to CGd, it was included; while STI was removed from the model. Deviance decreased as a result, from 1389.19 in Model 2 to 1385.53 in Model 3 (Table 8), indicating that this reduction to three level-1 variables resulted in a model with a better fit to the data.

In its reduced statistical form, Model 3 is expressed as:

$$\text{CGd} = y_{0C} + y_{1C} * \text{SSI} + y_{2C} * \text{FOR} + y_{3C} * \text{IMM} + u_0 + r \quad (4)$$

Model 3 tested an alternative model that included only 3 level-1 predictors (SSI, FOR, and IMM). As described in the Model 1 section, one measure of how well a model fits the data is the extent to which it explains variance within and between groups. Pseudo R^2 was calculated for Model 3. Within groups variance explained was .25 and between groups was a negative value. In other words, Model 3 explained 25% of the variance in CGd within courses (between students), which was an improvement over Model 2's 21%. However, the amount of variance in CGd between courses explained by Model 3 could not be determined.

Put another way, including only three level-1 variables did improve the predictive value of the model (Table 8). In Model 3, the intercept (2.41) is no longer a mean, but

instead represents the predicted CGd if all predictor values are 0. In other words, if no SSI, FOR, or IMM occurred in the course, the average CGd is predicted to be 2.41. The effects estimates provide a prediction of the relative contribution of each of the level-1 variables. For instance, for each instance of student-student interaction, we expect CGd to increase by .02. While all three variables are statistically significant predictors, IMM is predicted to have the strongest relationship with CGd (Table 8).

In summary, Model 3 found that including these three level-1 variables results in a more effective predictive model than Model 1, which did not include them, or model 2, which included an additional variable. However, the predictive value of the model can be improved by adding level-2 variables.

It is important to note that Model 3 was determined during the process of developing Model 4. Though the elements of Model 3 are presented before those of Model 4, it was actually developed simultaneously. In essence, Model 3 represents the "final" model (Model 4), but without any level-2 variables.

Model 4: Mean level-1 and level-2 effects on mean course grade. The fourth model was an alternative model that tested whether inclusion of level-2 variables improved the predictive value of Model 3. Model 4 included both level-1 and level-2 variables that contributed significantly to the prediction of course grade. Several models were tested in order to identify the most parsimonious model that explained course grade variability as a result of level-1 and level-2 variables. Variables were included or excluded based on conceptual factors, based on correlations between variables, and based on the level of deviance indicated when HLM modeling was conducted.

As a starting point, all level-1 variables and all level-2 variables were included.

Deviance was higher than Model 2, so the model required revision. As a starting point for revising the model, correlations between variables were examined.

There was a strong correlation between formative activities at the student level (level-1, FOR) and formative opportunities at the course level (level-2, NFOR); $r(579) = .80, p < .001$, 2-tailed. Because FOR is a more precise measure (associated with specific students, rather than with a course as a whole where individual students may or may not have experienced the variable), it was included; while NFOR was removed from the model. Deviance decreased as a result, indicating that this change improved the model's fit to the data.

Two level-1 variables were strongly correlated: student-teacher interaction (STI) and immediacy (IMM), $r = .36, N = 581, p < .001$, 2-tailed. As described in Model 3, above, IMM cannot occur without STI. Thus, IMM was retained in the model and STI was removed. Deviance decreased as a result, again indicating that this change improved the model's fit to the data.

Finally, the HLM model that incorporated these two changes (removing NFOR and removing STI) suggested a third change. Significance levels from that HLM analysis suggested that only one level-2 variable (building student capacity, BSC) was likely to result in statistical significance. All other level-2 variables were removed from the model. The resulting model did indeed result in the lowest deviance of those tested, indicating that including all three changes resulted in the best fit to the data. In its reduced statistical form, the final Model 4 is expressed as:

$$CGd = y_{0C} + y_{01} * BSC + y_{1C} * SSI + y_{2C} * FOR + y_{3C} * IMM + u_0 + r \quad (5)$$

where $y_{01} * BSC$ represents the effect of building student capacity on course grade and

the remaining elements have been identified previously.

Model 4 tested an alternative model that included three level-1 predictors and one level-2 predictor. Pseudo R^2 was calculated for Model 4. Within groups variance explained was .25 and between groups was a negative value. In other words, Model 4 explained 25% of the variance in CGd within courses (between students), which was the same variance explained as for Model 3. The amount of variance in CGd between courses explained by Model 4 could not be determined. Even though Model 4 did not explain more variance than Model 3, it was still a better fit to the data for other reasons. Based on deviance, Model 4 provided the best fit between model and data, or the most accurate prediction of CGd using the variables measured in this study. Deviance in Model 3 was 1385.53, but in Model 4 it had decreased to 1381.69.

In Model 4, three level-1 variables and one level-2 variable are present. If all of these predictor values are 0, we would predict average CGd to be 2.61 (the intercept). Since all four variables are statistically significant, we can also estimate their relative predictive value for CGd. For each instance of SSI, CGd is expected to increase by .02; for each instance of FOR, by .01; for each instance of IMM, by .16; and for each instance of BSC, it is expected to decrease by .47. BSC is predicted to have the strongest relationship with CGd, though it predicts a decrease in course grade.

These numbers may appear quite small, but it is important to remember that CGd is on a scale of 0-4. For example, Model 4 suggests that a student who engages in average SSI behavior will experience an increased course grade of nearly 1/3 of a letter grade. That is, the SSI intercept is .02 (Table 8) and average SSI is 15.20 (Table 3). Multiplying these numbers results in .30, and since we are referring to a letter scale where 0 = F and 4

= A, .33 is one-third of a letter grade.

Other variables are less dramatically predictive of CGd, but in Model 4 are statistically significant nonetheless. For instance, Model 4 suggests that a student who engages in average FOR behavior will experience an increased course grade of nearly 1/10 of a letter grade. That is, the FOR intercept is .01 (Table 8) and average FOR is 6.51 (Table 3). Multiplying these numbers results in .07, and since we are referring to a letter scale where 0 = F and 4 = A, .07 is nearly one-tenth of a letter grade.

Similarly, Model 4 suggests that a student who engages in average IMM behavior will experience an increased course grade of .06 of a letter grade. That is, the IMM intercept is .16 (Table 8) and average IMM is 0.38 (Table 3). Multiplying these numbers results in .06.

More dramatically, Model 4 suggests that a student who experiences average BSC levels will experience a decreased course grade of one-fifth of a letter grade. That is, the BSC intercept is -0.47 (Table 8) and average BSC is 0.43 (Table 4). Multiplying these numbers results in -.20, and since we are referring to a letter scale where 0 = F and 4 = A, .20 is one-fifth of a letter grade and the negative value indicates an expected decrease in grade.

In summary, CGd is best explained by the level-1 variables SSI, FOR, and IMM and the level-2 variable BSC. All variables were statistically significant at the .05 level or lower. While SSI, FOR, and IMM were positively predictive of CGd, BSC was negatively predictive of CGd.

Research Question 2: Assignment Grade

Research question 2 was "Which of the four pedagogic factors and six course

factors predict student grades at the assignment level?" In order to answer this question, assignment grade (AGd) was analyzed using HLM. Three models were ultimately required in order to draw conclusions about variables that predict assignment grade. The evolution of these models is described in the Model sections that follow. Table 9 presents statistical findings for all three models.²

Model 1: Mean assignment grade. The first model was the null model (or one-way ANOVA). This model provided an estimate of overall grand mean AGd across all students. In its reduced statistical form, the null model is expressed as:

$$\text{AGd} = y_{0c} + u_0 + r \quad (6)$$

where AGd represents average assignment grade for a given student in a particular class, y_{0c} represents overall grand mean assignment grade, u_0 represents residual prediction error associated with level-2 variables, and r represents residual prediction error associated with level-1 variables. In Model 1, mean assignment grade, or intercept, is 61.90.

Model 1 is the null model; it assumes that there is no predictive value to adding either level-1 or level-2 variables. There is only one fixed effect estimate (the intercept, 61.90; Table 9), which is interpreted as the average value of the outcome variable AGd across all students. In other words, the average assignment grade is 61.90, or a D- on a 10-point scale. (This differs slightly from Table 3's 60.41 due to rounding errors.)

² The process of HLM modeling is described in less detail here than in the Research Question 1 section, since that section addressed HLM processes that are common to all HLM modeling.

Table 9

Fixed Effects Estimates for Models of the Predictors of Assignment Grade

Parameter	Model 1		Model 2		Model 3	
Intercept	61.90***	(2.15)	47.89***	(3.57)	49.05***	(5.36)
Level-1 (student)						
SSI			0.61***	(0.08)	0.59***	(0.07)
STI			0.39**	(0.15)	0.38**	(0.14)
FOR			0.45***	(0.13)	0.53***	(0.15)
IMM			2.82**	(1.04)	2.61*	(1.11)
Level-2 (course)						
SDL					-0.08	(1.87)
SCI					0.61	(0.56)
BSC					-5.71	(4.97)
NFOR					-1.24***	(0.32)
MLO					0.48	(0.37)
VTA					0.49*	(0.17)
VSM					0.04	(0.35)
PIC					-0.39*	(0.14)
Variance Components						
Within-course	909.47		623.66		24.97	
Between (intercept)	87.63		425.45		15.04	
-2 log likelihood	5632.43		5461.61		5430.12	

Note. For fixed effect portion of the table, standard errors are in parentheses.

* $p < .05$. ** $p < .01$. *** $p \leq .001$.

The ICC³ for AGd was only 8.79% and did not meet the 10% threshold suggested by Raudenbush and Bryk (2002). However, there are theoretical and statistical reasons that HLM might nevertheless be an appropriate analytical method for this data. The theoretical reasons are those referred to previously and described by Luke (2004): HLM analysis accounts for the nested nature of the data and allows for analysis of individual observations that are nested within larger groups, and this data involves students who are

³ Intraclass correlation coefficient (ICC) is calculated from the null model and indicates the proportion of variance occurring within each level in a nested data set (Raudenbush & Bryk, 2002). It is described fully in the Research Question 1 section.

nested within courses.

Additionally, there are two statistical reasons for using HLM, despite ICC being less than 10%. First, as discussed previously, 30 courses is the minimum number of level-2 groups needed for a study of this type (Maas & Hox, 2005). It is possible that the low ICC is a result of using only the minimum number of level-2 observations, and not due to the data per se. A second statistical rationale comes from research that challenges Raudenbush and Bryk's (2002) assertion that ICC must exceed 10% if HLM is to be used. Roberts (2007, p. 3) argues that the 10% threshold is less important than the predictors included in the model being tested. Using a variety of data sets, Roberts demonstrated that "there may never be a time when it is acceptable to say that the only time that multilevel analysis is appropriate is when ICC is beyond some threshold" (p. 15).

Given that there are theoretical and statistical arguments for the use of HLM in this case, HLM analyses were used in order to investigate the role of level-1 and level-2 variables as predictors of assignment grade. This decision was confirmed by conducting an ordinary regression to test the same model that HLM ultimately resulted in (described later in this section). All variables identified as significant predictors using HLM (Table 9) were also identified using ordinary regression, though ordinary regression did identify an additional significant variable. MLO significantly predicted AGd, $\beta = .64$, $t(568) = 2.53$, $p < .05$. MLO also explained a significant proportion of variance in AGd, $R^2 = .28$, $F(12,568) = 18.16$, $p < .001$.

It is not surprising that ordinary regression resulted in a larger number of significant predictors than HLM did. When data are nested, yet the nested nature of the data is not taken into account, type I error is inflated. Had ordinary regression failed to

confirm the significance of the variables identified by HLM, then we would have cause for concern. However, the fact that it confirmed those variables and suggested an additional one only serves to support the decision to use HLM. Regardless of the estimation model, the predictors that resulted from Model 3 and identified in Table 9 were significant. However, the use of HLM adjusts standard errors to account for nesting, thus minimizing the likelihood of type I error and increasing the likelihood that the variables identified in the Model 3 are relevant to our understanding of AGd.

The variance of AGd within courses is 909.47 and between courses is 87.63 (Table 9), so a larger portion of variability in CGd lies within courses. The ICC is 8.79%, indicating that almost 9% of the total variation in AGd is accounted for by differences across courses, as opposed to within specific courses.

Model 2: Mean assignment grade with level-1 controls. The second model was an alternative model that tested whether inclusion of level-1 variables improved the predictive value of Model 1. Model 2 added four level-1 (student) variables to Equation 6 at level-1. Essentially, this model provides an estimate of AGd across all 30 courses controlling for the influence of level-1 variables. No level-2 variables were included in Model 2. In its reduced statistical form, Model 2 is expressed as:

$$AGd = y_{0C} + y_{1C} * SSI + y_{2C} * STI + y_{3C} * FOR + y_{4C} * IMM + U_0 + r \quad (7)$$

where $y_{1C} * SSI$ represents the effect of student-student interaction on assignment grade, $y_{2C} * STI$ represents the effect of student-teacher interaction on assignment grade, $y_{3C} * FOR$ represents the effect of student formative behaviors on assignment grade, and $y_{4C} * IMM$ represents the effect of immediacy on assignment grade. Because all of these variables were true counts, they were entered into Equation 7 uncentered.

Model 2 tested an alternative model that included four level-1 predictors. One measure of how well a model fits the data is the extent to which it explains variance within and between groups.⁴ Pseudo R^2 was calculated for Model 2. Within groups variance explained was .31 and between groups variance was a negative value. In other words, Model 2 explained 31% of the variance in AGd within courses (between students). However, the amount of variance in AGd between courses explained by Model 2 could not be determined.

Put another way, including level-1 variables did improve the predictive value of the model (Table 9). Model 2 introduced all level-1 predictors. The intercept (47.89) is no longer a mean, but instead represents the expected AGd if all predictor values are 0. In other words, if no SSI, STI, FOR, or IMM occurred in the course, the average AGd is expected to be 47.89 (an F). The effects estimates provide a prediction of the relative contribution of each of the level-1 variables. IMM is predicted to have the strongest relationship with AGd (for every instance of IMM, AGd is expected to increase by 2.61; over one-quarter of a letter grade).

In summary, Model 2 demonstrates that including level-1 variables results in a more effective predictive model than Model 1, which did not include them. Additionally, it found that all four level-1 variables were statistically significant contributors to this improved model. However, introducing level-2 variables to the model (Model 3) resulted in a model that was an even better fit to the data.

Model 3: Mean level-1 and level-2 effects on mean assignment grade. The third model was an alternative model that tested whether inclusion of level-2 variables

⁴ Pseudo R^2 as a means of calculating variance in HLM is explained in detail in Research Question 1's Model 2 section.

improved the predictive value of Model 2. Model 3 included both level-1 and level-2 variables that contributed significantly to the prediction of assignment grade. As a starting point, all level-1 variables and all level-2 variables were included. Deviance was lower than Model 2, so the inclusion of level-2 variables improved the model's fit to the data. Based on correlations between three level-2 variables (NFOR, VTA, and PIC), three exploratory models were tested, but in all cases deviance was higher than the "starting point" model, so these exploratory models were rejected. For the same reason as in the course grade modeling process described previously, a model without the level-1 variable of STI was also tested, but it too resulted in higher deviance. Thus, the ultimate Model 3 included all four level-1 variables and all 8 level-2 ones, as this model resulted in lower deviance than all other models tested.

In its reduced statistical form, Model 3 is expressed as:

$$\begin{aligned} \text{AGd} = & y_{0c} + y_{01} * \text{SDL} + y_{02} * \text{SCI} + y_{03} * \text{BSC} + y_{04} * \text{NFOR} + & (8) \\ & y_{05} * \text{MLO} + y_{06} * \text{VTA} + y_{07} * \text{VSM} + y_{08} * \text{PIC} + y_{1c} * \text{SSI} + y_{2c} * \text{STI} + \\ & y_{3c} * \text{FOR} + y_{4c} * \text{IMM} + u_0 + r \end{aligned}$$

where $y_{01} * \text{SDL}$ represents the effect of student directed learning on assignment grade, $y_{02} * \text{SCI}$ represents the effect of student-content interaction on assignment grade, $y_{03} * \text{BSC}$ represents the effect of building student capacity on assignment grade, $y_{04} * \text{NFOR}$ represents the effect of the number of formative assignments on assignment grade, $y_{05} * \text{MLO}$ represents the effect of measurable learning objectives on assignment grade, $y_{06} * \text{VTA}$ represents the effect of varied teaching activities on assignment grade, $y_{07} * \text{VSM}$ represents the effect of varied sensory modalities on assignment grade, and $y_{08} * \text{PIC}$ represents the effect of preprogrammed instructor communication on

assignment grade (other elements of Equation 8 were identified in Model 1).

Model 3 tested an alternative model that included four level-1 predictors and eight level-2 predictors. Pseudo R^2 was calculated for Model 3. Within groups variance explained was .97 and between groups variance was .83. In other words, Model 3 explained 97% of the variance in AGd within courses (between students) and 83% of the variance in AGd between courses. This is a dramatic improvement over Model 2, indicating a better fit to the data.

Based on deviance (-2 log likelihood), Model 3 provided the best fit between model and data, or the most accurate prediction of AGd using the variables measured in this study. In Model 3, four level-1 variables and eight level-2 variables are present. If all of these predictor values are 0, we would expect average AGd to be 49.05 (an F). Seven of these variables are statistically significant and we can estimate their relative predictive value for AGd. For each instance of SSI, AGd is expected to increase by .59; for each instance of STI, by .38; for each instance of FOR, by .53; for each instance of IMM, by 2.61; and for each instance of VTA, by .49. However, for each instance of NFOR, AGd is expected to *decrease* by 1.24 and for each instance of PIC it is expected to *decrease* by .39.

AGd was measured using a 100-point scale and we can use the numbers in the preceding paragraph to estimate average student assignment grade. For example, Model 3 suggests that a student who engages in average SSI behavior will experience an increased assignment grade of nearly a letter grade (assuming a 10-point grading scale). The SSI intercept is .59 (Table 9) and average SSI is 15.20 (Table 3). Multiplying these numbers results in 8.97, or 9 out of the 100-point scale.

Likewise, Model 3 suggests that a student who engages in average STI behavior will experience an increased assignment grade of nearly one-quarter of a letter grade (assuming a 10-point grading scale). The STI intercept is .38 (Table 9) and average STI is 6.46 (Table 3). Multiplying these numbers results in 2.45, or just under one-quarter of a letter grade on a 100-point scale.

Using similar calculations, we can determine that a student who engages in average IMM behavior will experience an increased assignment grade of one point, or .99 on a 100-point scale.

A student who engages in average FOR behavior will experience an increased assignment grade of one-third of a letter grade, or 3.45 on a 100-point scale. Conversely, a student who experiences average NFOR behavior will experience a *decreased* assignment grade of over half of a letter grade, or 6.41 on a 100-point scale. This seeming inconsistency is addressed in Chapter Five.

Finally, a student who experiences average PIC behavior will experience a decreased assignment grade of over half of a letter grade, or 6.59 on a 100-point scale.

In summary, AGd is best explained by the level-1 variables SSI, STI, FOR, and IMM and the level-2 variable SDL, SCI, BSC, NFOR, MLO, VTA, VSM, and PIC. However, only SSI, STI, FOR, IMM, VTA, NFOR, and PIC were statistically significant predictors (at the .05 level or lower). Of the statistically significant variables, all were positively correlated with AGd except for NFOR and PIC, which were negatively correlated. These results are discussed in Chapter Five.

Research Question 3: Retention

Research question 3 was "Which of the four pedagogic factors and six course

factors predict student retention in the course?" In order to answer this question, retention (RET) was analyzed using HLM. Three models were ultimately required in order to draw conclusions about the variables that predict retention. The evolution of these models is described in the Model sections that follow. Table 10 presents statistical findings for all three models.⁵

Table 10

Fixed Effects Estimates for Models of the Predictors of Retention

Parameter	Model 1		Model 2		Model 3	
Intercept	.77**	(0.03)	.64**	(0.04)	.65**	(0.04)
Level-1 (student)						
SSI			0.01**	(0.00)	0.01**	(0.00)
STI			0.00	(0.00)		
FOR			0.00	(0.00)	0.00	(0.00)
IMM			0.02	(0.02)	0.03*	(0.01)
Level-2 (course)						
SDL						
SCI						
BSC						
NFOR						
MLO						
VTA						
VSM						
PIC						
	Variance Components					
Within-course	0.17		0.14		0.14	
Between (intercept)	0.02		0.04		0.04	
-2 log likelihood	648.02		606.48		596.39	

Note. For fixed effect portion of the table, standard errors are in parentheses.

* $p \leq .05$. ** $p < .001$.

⁵ The process of HLM modeling is described in less detail here than in the Research Question 1 section, since that section addressed HLM processes that are common to all HLM modeling.

Model 1: Mean retention. The first model was the null model (or one-way ANOVA). This model provides an estimate of overall grand mean RET across all students. In its reduced statistical form, the null model is expressed as:

$$\text{RET} = y_{0c} + u_0 + r \quad (9)$$

where RET represents course retention for a given student in a particular class, y_{0c} represents overall grand mean retention, u_0 represents residual prediction error associated with level-2 variables, and r represents residual prediction error associated with level-1 variables. Model 1 is the null model. There are no level-1 or level-2 predictors. There is only one fixed effect estimate (the intercept, .77; Table 10), indicating that mean retention is .77. Note that RET is a "dummy" variable, where 0 = not retained and 1 = retained. Obviously, a student cannot be 77% retained, but this intercept should be interpreted to mean that each student has a 77% likelihood of being retained. Put another way, the average retention rate is 77% of students retained. (This differs slightly from Table 3's 75.73 due to rounding errors.)

The ICC⁶ for RET was 9.56% and this is very close to the 10% threshold suggested by Raudenbush and Bryk (2002). Additionally, as discussed in the Research Question 2 section, there are other reasons for the use of HLM when analyzing these data. Given that there are theoretical and statistical arguments for the use of HLM in this case, HLM analyses were used in order to investigate the role of level-1 and level-2 variables as predictors of retention. This decision was confirmed by conducting an ordinary regression to test the same model that HLM ultimately resulted in. All variables

⁶ Intraclass correlation coefficient (ICC) is calculated from the null model and indicates the proportion of variance occurring within each level in a nested data set (Raudenbush & Bryk, 2002). It is described fully in the Research Question 1 section.

identified as significant predictors using HLM (Table 10) were also identified using ordinary regression. Regardless of the estimation model, the predictors that resulted from the final model used to predict RET were significant.

The variance of RET within courses is 0.17 and between courses is 0.02 (Table 10), so a larger portion of variability in RET lies within courses. The ICC is 9.56%, indicating that almost 10% of the total variation in RET is accounted for by differences across courses, as opposed to within specific courses.

Model 2: Mean retention with level-1 controls. The second model was an alternative model that tested whether inclusion of level-1 variables improved the predictive value of Model 1. Model 2 added four level-1 (student) variables to Equation 9 at level-1. This model provides an estimate of RET across all 30 courses controlling for the influence of level-1 variables. No level-2 variables were included in Model 2. In its reduced statistical form, Model 2 is expressed as:

$$\text{RET} = y_{0C} + y_{1C} * \text{SSI} + y_{2C} * \text{STI} + y_{3C} * \text{FOR} + y_{4C} * \text{IMM} + U_0 + r \quad (10)$$

where $y_{1C} * \text{SSI}$ represents the effect of student-student interaction on retention, $y_{2C} * \text{STI}$ represents the effect of student-teacher interaction on retention, $y_{3C} * \text{FOR}$ represents the effect of formative behaviors on retention, and $y_{4C} * \text{IMM}$ represents the effect of immediacy on retention. Because all of these variables were true counts, they were entered into Equation 10 uncentered.

Model 2 tested an alternative model that included four level-1 predictors. One measure of how well a model fits the data is the extent to which it explains variance

within and between groups.⁷ Pseudo R^2 was calculated for Model 2. Within groups variance explained was .18 and between groups variance was a negative value. In other words, Model 2 explained 18% of the variance in RET within courses (between students). However, the amount of variance in RET between courses explained by Model 2 could not be determined.

Put another way, including level-1 variables did improve the predictive value of the model (Table 10). Model 2 introduced all level-1 predictors. The intercept (.64) is no longer a mean, but instead represents the expected RET if all predictor values are 0. In other words, if no SSI, STI, FOR, or IMM occurred in the course, the average RET is expected to be 64% (a 64% retention rate). The effects estimates provide a prediction of the relative contribution of each of the level-1 variables, but only one of these variables is statistically significant. Every instance of SSI is predicted to increase RET by 1%. Put another way, each time that a student interacts with another student, it is expected that that student's likelihood of completing the course increases by 1%.

In summary, including level-1 variables results in a model that is a better fit to the data than Model 1, which did not include them. However, only one level-1 variable (SSI) was a statistically significant contributor to this improved model. However, introducing level-2 variables to the model (Model 3) resulted in a model that was an even better fit to the data.

Model 3: Mean level-1 and level-2 effects on mean retention. The third model was an alternative model that tested whether inclusion of level-2 variables improved the predictive value of Model 2. Model 3 included only three level-1 variables. As with the

⁷ Pseudo R^2 as a means of calculating variance in HLM is explained in detail in Research Question 1's Model 2 section.

previous student outcomes, an attempt was made to identify the most parsimonious model to explain retention in terms of level-1 and level-2 variables. As a starting point, all level-1 variables and all level-2 variables were included. However, deviance increased considerably, indicating that these "test" models were poorer fits to the data than the existing Model 2. Further model testing determined that introducing *any* level-2 variable increased deviance, so only by revising which level-1 variables were included in the model could a more accurate model be developed. Ultimately, three level-1 variables were used: SSI, FOR, and IMM. STI was removed from the model because it was the least statically significant variable in the "starting point" model that included all variables. Additionally, as discussed in the Research Question 1 section, there are conceptual reasons to remove STI if IMM is included. This decision was supported by the lower deviance of the three-variable model that was ultimately used. Model 3's deviance was 596.39, compared to Model 2's deviance of 606.48 (Table 10). In its reduced statistical form, Model 3 is expressed as:

$$RET = y_{0C} + y_{1C} * SSI + y_{2C} * FOR + y_{3C} * IMM + u_0 + r \quad (11)$$

Model 3 tested an alternative model that included three level-1 predictors. Pseudo R^2 was calculated for Model 3. Within groups variance explained was .18 and between groups variance was a negative value. In other words, Model 3 explained 18% of the variance in AGd within courses (between students). However, the amount of variance in AGd between courses explained by Model 3 could not be determined. This is identical to the variance explained by Model 2. Nevertheless, there are other reasons that Model 3 is a better fit to the data than Model 2. Based on deviance (-2 log likelihood), Model 3 provided the best fit between model and data, or the most accurate prediction of RET

using the variables measured in this study.

In Model 3, three level-1 variables are present. If all of these predictor values are 0, we would expect average RET to be 65%. Two of these variables are statistically significant. For every instance of SSI, RET is expected to increase by 1%. For every instance of IMM, it is expected to increase by 3%.

For example, Model 3 suggests that a student who engages in average SSI behavior is 15% more likely to complete the course than a student who participates in no SSI behavior. The SSI intercept is .01 (Table 10) and average SSI is 15.20 (Table 3). Multiplying these numbers results in .15, or 15%. Likewise, a student who experiences average IMM behavior is slightly more than 1% more likely to be retained than a student who experiences no IMM behavior. The IMM intercept is .03 (Table 10) and average IMM is 0.38 (Table 3). Multiplying these numbers results in 1.14%.

In summary, RET is best explained by the level-1 variables SSI, FOR, and IMM, but only SSI and IMM were statistically significant predictors (at the .05 level or lower). Both statistically significant variables were positively correlated with RET.

Research Question 4: Student Satisfaction

Research question 4 was "Which of the four pedagogic factors and six course factors predict student retention in the course?" Unfortunately, the outcome variable of student satisfaction (SS) was not suitable for HLM analysis. SS was calculated as an average of student course evaluations for each of the 30 courses used in this study, so it was a level-2 (course) variable and thus did not involve nesting. Ideally, the variables predicting SS would be analyzed using multiple regression. However, 30 courses did not produce sufficient statistical power to conduct this analysis. The statistical program

G*Power 3 (Institut für Experimentelle Psychologie, n.d.) revealed only 0.52 statistical power based on the data collected for this study, and thus multiple regression was not an option. Even after combining variables that were closely correlated, power remained insufficient. G*Power 3 indicated that a minimum of 58 groups would be required in order to reach sufficient power for multiple regression. Unfortunately, time did not allow for the collection of such a large amount of additional data.

Analysis Limitations and Assumptions

As discussed in the Sample and Population section of Chapter Three and in the Inferential Results section in this chapter, the number of level-2 groups included in this study is the minimum number suggested by Maas and Hox (2005) for 2-level HLM: 30 groups. It is possible that this limited number of groups affected some of the analyses. For instance, as discussed in the Research Question 2 and 3 sections, low ICC in both AGd and RET may be due to the use of only 30 groups, rather than to the nature of the data per se. Being at the cusp of the recommended number of level-2 groups (per Maas & Hox) may mean that type II error was inflated. In other words, some level-2 variables may not have been found to be statistically significant, even though they were predictors of outcome variables.

A second potential limitation to analysis is the lower than 10% ICC found in both AGd and RET. As discussed in the Research Question 1 section, Raudenbush and Bryk (2002) assert that ICC must exceed 10% if HLM is to be used. However, this is a statistical rationale and there are theoretical reasons for using HLM to analyze these outcomes (Luke, 2004; Tabachnick & Fidell, 2007), as well as statistical ones that challenge Raudenbush and Bryk's assertion (Roberts, 2007). The possibility that HLM

was an inappropriate analytical tool for use with the AGd and RET outcome variables is a second potential limitation.

A third potential limitation is the inability to draw inferential conclusions about the SS outcome. Due to limitations in the data, only descriptive statistics are available for SS analysis. Because SS is a group-level (level-2) variable, it cannot be analyzed using HLM. Given the large number of predictive variables in this study, a much larger sample than that collected here is required if multiple regression is to be used to analyze the factors predicting SS.

A fourth limitation is suggested by some of the inconsistencies seen in the generation of HLM models for outcome variable RET. ICC was 9.56%, so even if adhering to Raudenbush and Bryk's (2002) threshold, this is very close to 10% and the other reasons for using HLM discussed previously apply as well. Nevertheless, none of the level-2 variables improved the model (based on deviance) that was generated using only level-1 variables. A possible reason for this is that there are other level-2 variables that were not measured that better predict RET. Failure to measure relevant level-2 variables is a limitation of this study's analyses.

Additionally, there was a threat to the validity of the research design that could not be fully addressed. Mortality was a threat to internal validity. All of the outcome measures (with the exception of assignment grades) assume that the student has completed the course in question. Unfortunately, factors unrelated to instructor behavior (e.g., student job change) might result in the student withdrawing from the course. The outcome variables of course grade and student satisfaction were not available for such students, and interpretations concerning the retention variable must also be made

cautiously.

Finally, there was a threat to instrument validity. The construct validity of the operational definitions used to measure the predictive variables is suspect. Though the "spot-checking" phase of the pilot study (phase 6) helped to ensure reliability of the measures, construct validity might be questioned. Choosing one or more definitions used by other researchers and adapting them for a Moodle environment was used to develop the measures, and thus the content validity of these measures should be sufficient. However, the independence of the 12 constructs being measured here could not be determined until data had been collected. As discussed in the Intercorrelations Among Predictive and Outcome Variables section, there were a number of measures that were closely correlated, drawing into question the construct validity of some of these constructs.

Summary

HLM was used to determine the predictive variables that are significantly correlated with three of the outcome variables (CGd, AGd, and RET). The fourth outcome variable (SS) was not suitable for HLM or for other inferential statistical analyses, but was described descriptively. Briefly, higher levels of SSI, FOR, and IMM predict higher levels of CGd, while higher levels of BSC predict lower levels of CGd; higher levels of SSI, STI, FOR, IMM, and VTA predict higher levels of AGd, while higher levels of NFOR and PIC predict lower levels of AGd; and higher levels of SSI and IMM predict higher levels of RET. These findings are discussed in terms of the research questions and the implications of this study in Chapter Five.

CHAPTER FIVE: DISCUSSION, IMPLICATIONS, AND LIMITATIONS

The purpose of this study was to identify instructor behaviors that lead to positive student outcomes in online courses. It was guided by four research questions:

1. Which of the four pedagogic factors and six course factors predict student grades at the course level?
2. Which of the four pedagogic factors and six course factors predict student grades at the assignment level?
3. Which of the four pedagogic factors and six course factors predict student retention in the course?
4. Which of the four pedagogic factors and six course factors predict aggregate student satisfaction on end-of-course surveys?

This chapter summarizes the answers to these four questions, drawing on the statistical analyses described in the preceding chapter. It then discusses each research question in detail, addressing both expected and unexpected findings and implications for practice. It ends with a discussion of the limitations of this study and recommendations for future research.

Answers to Research Questions

The purpose of this study was to identify instructor behaviors that lead to positive student outcomes in online courses. As discussed in Chapter Two, the majority of research on this topic has used anecdotal data or student self-ratings, or else has drawn from theoretical literature without including empirical evidence. Thus, it should not be surprising if some of the predictive variables used in this study do not in fact predict

successful student outcomes. Table 11 summarizes the answers to each of the four research questions. Figure 2 illustrates the statistical significance and direction of the variables that contributed to the HLM models used to answer research questions 1, 2, and 3.

Table 11

Answers to Research Questions

Research question	Answer to research question
1. Which of the four pedagogic factors and six course factors predict student grades at the course level?	Increased amounts of level-1 variables SSI, FOR, and IMM predicted higher course grades. Increased amounts of level-2 variable BSC predicted lower course grades.
2. Which of the four pedagogic factors and six course factors predict student grades at the assignment level?	Increased amounts of level-1 variables SSI, STI, FOR, and IMM and level-2 variable VTA predicted higher course grades. Increased amounts of level-2 variables NFOR and PIC predicted lower course grades.
3. Which of the four pedagogic factors and six course factors predict student retention in the course?	Increased amounts of level-1 variables SSI, FOR, and IMM predicted higher course grades.
4. Which of the four pedagogic factors and six course factors predict aggregate student satisfaction on end-of-course surveys?	Due to limitations in the data, research question 4 could not be answered.

Course Grade			Assignment Grade			Retention		
SSI		BSC	SSI	SDL	NFOR	SSI	FOR	
FOR			STI	SCI	PIC	IMM		
IMM			FOR	BSC				
			IMM	MLO				
			VTA					

Figure 2. Statistical significance and direction of the variables that contributed to the HLM models used to answer research questions 1, 2, and 3. Green cells indicate statistically significant variables that were positively correlated with the outcome variables (course grade, assignment grade, or retention). Yellow cells indicate variables that were not statistically significant but that did contribute to the model's fit to the data. Red cells indicate statistically significant variables that were negatively correlated with the outcome variables.

Research Question 1

Research question 1 asked, "Which of the four pedagogic factors and six course factors predict student grades at the course level?" Three level-1 variables were positive predictors of course grade (CGd) and one level-2 variable was a negative predictor. Increased amounts of student-student interaction (SSI), formative student behavior (FOR), and immediacy (IMM) predicted higher course grades. Increased amounts of building student capacity (BSC) predicted lower course grades. While the first three of these are not unexpected, the last one certainly is.

Most of the research on SSI has suggested a positive relationship between SSI and student achievement, of which CGd is one measure (see for instance Picciano, 2002; Sher, 2004). What is surprising, in the area of interaction, is that STI (student-teacher interaction) was not a significant predictor of CGd. Indeed, as described in Chapter Four, removing STI from the HLM model improved its fit to the data. However, some research has noted that STI plays less of a role in student learning than we might expect (see for

instance Bernard et al., 2009; Rhode, 2009).

It is logical that FOR was a significant predictor of CGd. By definition, FOR allows students to re-take an assignment for an improved grade, so it stands to reason that students who most frequently take advantage of these opportunities will in fact experience higher assignment grades, and thus higher course grades. It is noteworthy that FOR (student participation in formative behaviors) was a significant predictor of CGd, but that NFOR (number of formative assignments in a course) was not. However, this finding echoes that of Posner (2011), who found that being in a formative learning condition did not improve overall student performance, but that among students in that condition, those who made the most formative behaviors did perform better than those who made fewer formative behaviors. Posner's research did not use an online environment, so this study extends that research to a second learning environment.

IMM has received limited research in online environments, and consequently was defined quite narrowly for the purpose of this study (instances of instructor praise, solicitation of viewpoints, humor, or self-disclosure directed at individual students in response to those students' behaviors). The fact that it proved a significant predictor supports Swan (2003), who is most responsible for applying immediacy research to online, as opposed to face-to-face, environments. IMM has received considerable face-to-face research, but defining it for an online environment is challenging and this study is the first to test Swan's operational definition empirically.

The surprising finding in research question 1 is the negative relationship between BSC and CGd. Why might building student capacity activities actually decrease course grade? Table 6 and the other variables that BSC is highly correlated with suggest one

explanation for this unexpected finding. Both NFOR (number of formative assessments in a course) and PIC (preprogrammed instructor communication) are strongly positively correlated with BSC ($r = .40$, $N = 581$, $p < .001$, 2-tailed and $r = .35$, $n = 440$, $p < .001$, 2-tailed; respectively). These variables also seem related because of the research question 2 findings, discussed in next section of this paper. While BSC is significantly and negatively correlated with course grade, PIC and NFOR are significantly and negatively correlated with assignment grade.

What do these three variables (BSC, PIC, NFOR) have in common? All three provide students with resources that the instructor intends to help them, but that are not specific requirements of assignments. Both PIC and NFOR are discussed in detail in the Research Question 2 section. Here, the focus is on BSC.

BSC is designed to ensure that students are prepared for online courses. It may be assumed by instructors that the "how to take an online course" skills learned during a BSC assignment will be used later in the semester. However, in fact many students may not refer to those earlier BSC activities when they are struggling with online learning. They may instead expect additional support, which the instructor believes is unnecessary. Studies of BSC have found that it increases help-seeking behavior by students, but BSC has not in fact been linked to improved student grades (Brown, 2004; Kitsantas & Chow, 2007). However, we do not know how instructors respond to this increased help-seeking, particularly if they feel that the prior BSC activities (and the often associated PIC and NFOR ones) should be sufficient.

It is possible that instructors who provide a lot of BSC (or PIC or NFOR) opportunities to students feel less need to "help" them later. Indeed, BSC is negatively

correlated with one measure of instructor attention to individual student needs, immediacy ($r = -.16, N = 581, p < .001, 2\text{-tailed}$). If instructors assume that BSC (or PIC or NFOR) provide students with all the opportunities that they need to succeed in the course, then they may do less to monitor student success or take other action when students are not performing well. They may assume that students can refer to the earlier BSC activities or later PIC if they are "lost" or "struggling." Conversely, students may not take these proactive measures. More precisely, *some* students do not take these proactive measures, and this percentage is large enough to result in the negative relationship between BSC and CGd.

In summary, increasing student opportunities for SSI, increasing students likelihood to engage in FOR, and increasing instructor IMM behaviors appear to improve course grade. However, instructors who use BSC assignments to prepare students for online learning should not assume that this preparation is sufficient for students to succeed in the course. Additional measures are probably still required.

Research Question 2

Research question 2 asked, "Which of the four pedagogic factors and six course factors predict student grades at the assignment level?" Four level-1 variables and one level-2 variable were positive predictors of assignment grade (AGd) and two level-2 variables were negative predictors. Increased amounts of student-student interaction (SSI), student-teacher interaction (STI), formative student behavior (FOR), immediacy (IMM), and varied teaching activities (VTA) predicted higher assignment grades. Increased amounts of formative assignments (NFOR) and preprogrammed instructor communication (PIC) predicted lower assignment grades. As described in Chapter

Three, all of the predictive variables measured in this study are expected to be positively correlated with the outcome variables. Thus, the SSI, STI, FOR, IMM, and VTA findings are not surprising. However, the negative relationship between PIC and NFOR and AGd is unexpected.

Most of the research on SSI has suggested a positive relationship between SSI and student achievement, of which AGd is one measure (see for instance Picciano, 2002; Sher, 2004). The same is true of STI (Bernard et al., 2009) and IMM (Swan, 2004). Likewise, and as discussed previously, we expect that students who engage in formative behaviors (FOR) will achieve higher assignment grades (Posner, 2011).

As discussed in Chapter Three, there is a large amount of theoretical literature extolling the benefit of addressing various learning styles (e.g., Gainor et al., 2004; Jackson & Helms, 2008), but attempts to empirically test these theories have found little support for the theory that learning styles affect learning (e.g., Brown-Syed et al., 2005). One possible explanation for this discrepancy is that it is the *variety* of teaching behaviors, rather than a match between teaching behavior and learning style, that benefits students (Neighmond, 2011). This study tested this theory and did in fact find a strong positive relationship between varied teaching activities (VTA) and AGd. Providing students with different kinds of learning and assessment activities did predict higher grades on individual assignments.

The surprising finding in research question 2 is the negative relationship between NFOR and PIC and CGd. Why might the number of formative activities in a course and the number of preprogrammed instructor communications both be related to lower assignment grade? As discussed in the Research Question 1 section, this may be because

instructors who use these teaching methods are less likely to provide students with additional support and assistance.

To illustrate, PIC appears at various times throughout the course, often through different means (message, email, announcement) and often not specifically related to an assignment. For instance:

Hope you all are well. Please let me know if anyone is having trouble, I really want you all to succeed!

Here is a quick reminder about our upcoming week.

First we have EXAM II on March 11-12. This exam will cover chapters 17-20.

However, for the week of March 11-18 we will have NO QUIZ (in observance of spring break). Hopefully, this will give you a little break, and make the exam a little less stressful. It would be great if you went ahead and read Chapter Two1, though [sic].

We will resume our regular schedule on the week of March 18-25: Chapter Two2 [sic].

Please let me know if you have questions.

In this PIC quote, students need to take proactive measures to benefit from the communication. They need to visit the Moodle Announcements page, they need to plan for both spring break and an upcoming quiz, and they need to contact the instructor if they have questions. Nevertheless, the instructor may assume that the PIC message is sufficient, and take it for granted that students will take the necessary steps alluded to by the message.

Likewise, NFOR allows students to go beyond the minimum for an assignment, by retaking it one or more times. But merely providing this opportunity does not ensure that students will engage in formative assessment. This discrepancy is illustrated by the fact that FOR (student formative behaviors) was a positive predictor of assignment grade (AGd), but that NFOR (the number of formative activities in a course) was a negative predictor of AGd. Posner (2011) found similar results in a study of face-to-face courses: that formative opportunities did not improve student performance, but that among students who had these opportunities, those who engaged in them most frequently did perform better than those who engaged in them less frequently.

Imagine that an instructor makes assignments formative (students can repeat them in order to earn a higher score). Having given students this opportunity, the instructor provides few other supports and may in fact make the assignments especially difficult. After all, the instructor thinks, "students can still earn the grade they want if they just put sufficient effort into retaking the assignments." However, retaking the assignments is presented as "optional," so many students choose not to do it. This increased difficulty and lessening of assistance results in lower assignment grades, unless students do in fact take advantage of the "optional" formative opportunities. Thus, those who do take advantage of these opportunities (students with high FOR numbers) perform well on assignments, but since so many choose not to do this, courses with high NFOR numbers actually result in lower AGd averages.

In summary, increasing student opportunities for SSI, STI, IMM, and VTA appears to improve assignment grade. However, instructors who use PIC to remind students of what to do in order to succeed in the course should not assume that these

reminders are sufficient in and of themselves. Additional measures are probably still required. Even more complicated, providing formative opportunities to students only results in improved assignment grade if students take advantage of these opportunities. Instructors might wish to make them "mandatory," rather than "optional."

The reason that so many variables predicted AGd, but fewer predicted CGd, may have to do with the fact that AGd is less affected by factors that were not included in this study. Presumably, AGd represents a grade on a student's assignment, while CGd represents several other factors such as assignment weights, late penalties, and the grades earned on other assignments as well.

Research Question 3

Research question 3 asked, "Which of the four pedagogic factors and six course factors predict student retention in the course?" Three level-1 variables were positive predictors of retention: SSI, FOR, and IMM, though only two were statistically significant (SSI and IMM). The role of SSI, FOR, and IMM as predictors of student success has been described previously, so their role here is not surprising. What may be surprising is that no level-2 predictors proved to be significant predictors of retention. This might be due to the statistical limitations of only 30 level-2 groups, as discussed in Chapter Four. However, it may also be that students' decisions to stay in a course are largely a result of interactions with other students (SSI), teacher support (IMM), and proactive efforts on their own parts (FOR); rather than the nature of course assignments and activities (level-2 variables such as SDL and VTA).

In summary, increasing student opportunities for SSI and FOR and providing teacher-to-student communication that reflects IMM appears to improve student

retention. These three elements have something in common. SSI requires students to actively engage with other students. FOR requires students to actively engage with the course material. IMM occurs when an instructor actively engages with an individual student in a personally meaningful way. In essence, this is the strongest support for the CoI (community of inquiry; Moore, 1989; Garrison et al., 2000) theory discussed in Chapter Two that has come out of this study. CoI theory posits that student-student, student-content, and student-teacher interactions should be at the center of online learning (Swan, 2003). At least where students' decisions to remain in an online course are concerned, this does seem to be the case.

Research Question 4

Research question 4 asked, "Which of the four pedagogic factors and six course factors predict aggregate student satisfaction on end-of-course surveys?" Unfortunately, due to limitations in the data, this question could not be answered. However, we can acknowledge that student satisfaction (SS) varies widely across courses. On a 5-point scale, SS ranged from 2.25 to 5.00 in the course sample used in this study. Clearly, there are elements to instructor behavior and course design that result in dramatic differences in SS and this topic is worthy of further research using a more suitable research design.

Relationships between the Research Question Answers

There are some commonalities between the answers to the three research questions that could be answered. Most obviously, both student-student interaction (SSI) and immediacy (IMM) were significant predictors of all three level-1 measures of successful students: course grade (CGd), assignment grade (AGd), and retention (RET).

A second commonality was the role of formative assessment, though the

relationship between formative assessment and student success is complex. The number of formative behaviors made by students (FOR) is positively correlated with course grade, assignment grade, and retention. On the other hand, the number of formative activities in a course (NFOR) is negatively correlated with assignment grade and was not a statistically significant predictor of CGd or RET. It seems that merely providing formative opportunities to students does not predict student success; but that when these opportunities are available, students who engage in them are more likely to be successful.

A third commonality is that instructor behaviors that ask students to perform additional work that is not a requirement of specific assignments seem to be negatively correlated with several measures of student success. Building student capacity (BSC) is negatively correlated with course grade and NFOR and preprogrammed instructor communication (PIC) are negatively correlated with assignment grade. This may suggest that students are unlikely to take advantage of supportive opportunities if not required to do so.

A less obvious pattern is that level-1 variables are consistently positively correlated with measures of student success (or are not statistically significant predictors). Level-2 variables, on the other hand, are sometimes positive predictors and sometimes negative ones, depending on the outcome variable in questions. This might be a result of the fact that using only 30 groups (courses) limits the likelihood of finding statistical significance, as discussed in Chapter Four. However, it might also be due to the fact that level-1 variables are the result of student actions while level-2 variables are the result of instructor actions. As has been suggested by several researchers, teacher behaviors are less important than are student behaviors in predicting student success (see for instance

Bernard et al., 2009; Rhode, 2009). Nevertheless, these findings do suggest some steps that instructors can take to increase the likelihood of student success in online courses.

Recommendations for Educators and Educational Administrators

A number of researchers have noted that many online courses do not take advantage of the networked nature of online communication or the interactive capacity of Web 2.0 tools and instead resemble traditional correspondence courses (Falvo, 2004; Mager, Tignor et al., 2008). In fact, many online instructors simply try to replicate face-to-face courses online when they teach online (Stevens-Long & Crowell, 2002). The findings of this study suggest that the tools inherent in online instruction can be used to increase the likelihood that students will complete courses (RET), earn high assignment grades (AGd), and earn high course grades (CGd).

Specifically, instructors should provide copious opportunities for student-student interaction (SSI). In the final models for all three research questions, SSI was found to be a significant predictor of positive student outcomes. While many instructors make SSI opportunities available, requiring such activity is recommended. For example, by requiring a specific number of discussion board posts to other students within each unit of the course. Instructors often require a specific number of discussion board posts, but this study suggests that the focus should be on responding to other students, rather than merely generating one's own posts.

Similarly, requiring students to engage in formative behavior (FOR) is also recommended. Many instructors make formative opportunities (NFOR) available, but doing so has mixed results. It is suspected that instructors who do this assume that students will take advantage of these opportunities, and therefore provide fewer supports

in other areas. Thus, it is recommended that NFOR opportunities be available *and* that FOR behaviors be required. For example, rather than *allowing* students to re-take a quiz up to three times for a better grade, students could be *required* to re-take it if they do not meet a given threshold (e.g., 80%).

Related, instructors who engage in building student capacity (BSC) and preprogrammed instructor communication (PIC) should be cautioned that such scaffolding is insufficient alone to promote student success. As the intercorrelations described in Chapter Four suggest, it appears that instructors who provide students with activities designed to build online course-taking skills (BSC) and regular reminders about what to do and how to do it (PIC) are those who engage in less direct facilitation (such as immediacy, IMM). While instructors may hope that providing copious supports to students will result in students using these supports, this does not appear to be the case. However, there are steps that instructors can take to ensure that BSC and PIC are not merely presented to students without instructor-student interaction.

It is recommended that instructors who rely on BSC or PIC find ways to ensure that students use these supportive activities and communications. It is not sufficient to merely make them available. For instance, instructors can ensure that PIC is read and acted upon if it includes a graded activity. Embedding a "click here to earn 1 point towards class participation" link in each PIC may ensure that students read the communications. Even more detailed requirements could be used to ensure that students act upon these communications. For instance, "submit a screenshot showing that you have watched the YouTube video used to introduce Unit 4."

Similarly, requiring that students engage in BSC activities for a grade can ensure

that students become familiar with online course-taking strategies. However, merely making these available at the beginning of the semester does not ensure that students will retain this knowledge when they need it later. It is recommended that instructors embed BSC activities at each point in the course where new online course-taking skills might be required. For instance, when students first engage in groupwork, a required (graded) BSC activity could be used that walks students through the various groupwork tools that the LMS makes available (e.g., file sharing, wiki) and external resources that groups might choose to use (e.g., Facebook, Google Docs).

The two remaining recommendations require a change in how instructors approach course design and interactions with students. Variety of teaching activities (VTA) predicts higher assignment grade. While instructors may be most comfortable using activities that they are familiar with (e.g., individual students complete a test based on textbook and online reading), it is recommended that they seek places in their courses where other activities and assessments can be used (e.g., field work resulting in a learning journal, group projects resulting in a shared grade on a multi-part project). In other words, it is recommended that instructors step outside of their comfort zones and try new ways of exposing students to course content and of assessing student learning.

Finally, because immediacy (IMM) predicted course grade, assignment grade, and retention, it is recommended that instructors interact with students on a personal basis whenever possible. Given the time constraints that instructors are often faced with, it is tempting to send a general email to the class or post an announcement on the course home page that addresses a number of concerns that students have individually expressed (perhaps via email or discussion board). However, it is recommended that instructors

instead reply to these students individually whenever possible. In addition, instructors should not hesitate to ask students for their opinions, to self-disclose, to use humor, or to praise students. These four elements are examples of IMM, and it seems that students who experience higher levels of IMM from their instructors also experience better course outcomes, as measured by grade and retention. Admittedly, teaching with this kind of personal immediacy may be a more difficult adaptation for instructors than simply changing the requirements surrounding formative assessment or adding a step to PIC, but this study suggests that there is immense value to such immediacy.

As Picciano and Seaman (2010, p. 24) noted, educational administrators generally make decisions regarding online instruction based on the need to provide broader access to students, rather than based on pedagogical concerns. However, if we assume that administrators do want students to complete courses (RET), earn high assignment grades (AGd), and earn high course grades (CGd), then this study suggests that they should encourage the instructor behaviors described immediately above. This could be done in several ways. First, administrators could implement instructor training requirements before instructors are allowed to teach online. These training sessions could include techniques for effectively providing the experiences that are linked to student success, such as requiring FOR behaviors of students or sending IMM communications to students. Second, administrators could require regular professional development for online instructors that addresses ways of implementing the recommendations described above. Third, they could include SSI, FOR, IMM, STI, VTA, BSC, PIC, and NFOR elements in teacher evaluation and course evaluation procedures. For example, student evaluations of online courses could ask questions that assess these qualities. Likewise,

criteria for merit pay, tenure, promotion, etc. could include these elements if faculty teach online. Finally, administrators could encourage their institutional effectiveness departments to conduct studies such as this one. Such studies would allow for more precise recommendations targeted to their particular colleges' unique populations and would add to the growing body of literature concerning effective online instruction.

A final recommendation suggested by this study applies not only to teachers and administrators, but to macro-structures such as colleges and university systems. This study suggests that the model of online instruction in which a single instructor teaches several hundred students is unlikely to lead to high levels of student achievement. SSI and IMM are virtually precluded by such an approach. Conversely, easily automated techniques such as BSC and PIC are more likely to be used, though in this study suggests that such techniques are negatively correlated with measures of student achievement.

Limitations of the Study and Recommendations for Future Research

There are statistical and analytical limitations to this study and there are conceptual ones. The statistical/analytical ones were addressed in the previous chapter, but two of them suggest future research methodologies. Future research should use more than 30 level-2 groups, as this would decrease the chance of type II error. Indeed, increasing the number of groups to more than 58 would also ensure statistical power necessary to analyze the student satisfaction data using normal regression (based on analysis using G*Power 3, Institut für Experimentelle Psychologie, n.d.). Since student satisfaction is a level-2 variable, it is unsuitable for HLM; but a sufficiently large student would allow for other analytical methods.

There are also statistical limitations that cannot yet be addressed, but that future

research should consider. These deal with the construct validity of the 12 predictive variables, as discussed in Chapter Four. As additional research on these topics is conducted, it should be possible to more accurately determine whether or not these measures do in fact assess different constructs.

Along with the statistical limitations mentioned already, there are conceptual ones as well; and future research should attend to these too. First, this study does not use a true measure of "learning." Both course grade and assignment grade are only limited proxies for learning. Future research might be able to use true measures of learning, such as pre-post assessments.

A second limitation is the fact that 30 different courses were used. While this did allow for the maximum variety of instructor behaviors and course designs, it also means that Assignment 1 in Course 1 was not the same as Assignment 1 in Course 2. Thus, even though data were collected for every assignment, only by averaging assignment grade for each student could the variable be analyzed using HLM. Future research could use the same course taught by the same instructor over multiple semesters, allowing for each assignment to be analyzed independently, rather than only aggregate assignment grade being used. However, the variability in teacher behavior would be lost in such a study, so there is a trade-off.

A possible compromise would be to use the same teacher, course and assignments; but to instruct the teacher to use different techniques each semester (e.g., SDL one semester, high amounts of PIC another, etc.). In other words, an experimental design could be used in order to eliminate some of the limitations of this study. Unfortunately, adding an experimental design would decrease the generalizability of the

findings, since instructors in the "real world" are seldom prescribed teaching methods.

In the current study, there are minimal threats to external validity. Over 40% of HCC's students take online courses (HCC, 2011b), so it is likely that the sample of courses randomly selected for this study, particularly as neither discipline nor instructor was repeated within the sample, was representative of the sampling frame. One concern, however, is that the homogeneity of HCC's student population (e.g., 89.8% white; HCC, 2011d) means that the study's results may not be generalizable to other populations. Likewise, community college students as a whole may not learn in the same ways as other students. Certainly, since the mean age is 28.6 (HCC, 2011d), it would be a mistake to apply these findings to K-12 environments or even to traditional college age students without replication. Future research could replicate this study in other populations.

Similarly, these findings may not be generalizable to courses that are designed in very different ways from those included in this study. For instance, colleges with very structured courses that all instructors teach in similar ways may not exhibit the same qualities measured in this study. Likewise, MOOCs (massive open online courses) are vastly different from the courses used in this study. It would be a mistake to assume that the results of this study generalize to such different online environments.

Replicating this study in other community college, such as those that use highly prescribed courses or those with vastly different populations, would help to establish external validity beyond courses and populations similar to HCC's.

Another limitation is the inability to fully measure some variables. As discussed in Chapter Three, data sources were limited only to the college's data warehouse and the courses' Moodle site. Thus, a number of variables may have been underestimated. For

instance, student-teacher interaction (STI) that occurred in person, by phone, or by email external to Moodle was not counted. Consequently, the following predictive variables may have been underestimated for some students: SSI, STI, IMM, and PIC. It is difficult to imagine a situation where these "external" communications could be measured by the researcher, but if a true experiment was used, perhaps the teacher could record the number of external communications with each student and share those data with the researcher.

There are several directions for future research that do not directly address a limitation of this study, but that would add considerably to the discourse concerning effective distance learning. One such research direction is to perform content analysis of SSI, STI, and IMM communications. The interpretations of findings earlier in this chapter hypothesize some reasons for student behavior. For instance, that instructors who provide BSC, NFOR, and PIC may be less likely to be supportive in other ways. This study's focus was on *which* variables predict student success, but future research could delve more deeply into *why* these relationships exist by analyzing the content of student-student and student-teacher interactions. For example, SSI of instructors who provide high amounts of BSC, NFOR, and PIC may in fact be less supportive of students than SSI of other instructors. Content analysis of these communications could measure the frequency of supportive words and phrases in order to support or refute this theory.

A related research direction would be to assess the quality, rather than just the quantity, of predictive variables. This would be a form of content analysis as well, though its nature would differ here (no longer merely counting supportive statements, for example). For instance, in counting incidents of SSI, students who primarily posted "I

agree" were assigned the same "weight" as those who posted substantive responses to the instructor or to other students. The role of SSI in student achievement may be very different if quality, rather than quantity, is assessed. Indeed, several studies have found significant differences in the quality of student discussion board posts when outcome measures such as course grade are used to differentiate among students (e.g., Stacey & Rice, 2002; Mager, Heulett et al., 2011).

Similarly, STI in which the teacher emails individual students to tell them that they have fallen behind in a course are vastly different from communications in which the teacher also includes additional steps that the student can take to get caught up. While neither communication qualifies as immediacy (IMM), there is a qualitative difference between the two. Similarly, communications from student to teacher were also coded as STI, but these also vary considerably. For example, "Have you graded exam 2 yet?" is vastly different from a substantive discussion of the meaning of "language." Analyzing these kinds of interactions would allow for more precise coding than was used in this study.

Likewise, formative assessment with detailed feedback explaining why an answer received the score that it did was "weighted" the same as formative assessments in which "right" and "wrong" were simply indicated with a reference to the page of the textbook where the material was covered. This same "equal weighting" occurred in all counts used in this study. Therefore, evaluating the quality of the predictive variables might help to explain some of the findings of this study and to more precisely predict student outcomes.

A third direction for future research would be to integrate the research conducted in this study with the holistic model proposed by Menchaca and Bekele (2008). This

study addressed only course and pedagogic factors (factors controllable by individual instructors), but some of the results of the study (e.g., variables that contributed to HLM models but that were not statistically significant) may be explained by other elements of Menchaca and Bekele's model, such as human factors. For example, perhaps the non-statistically significant contributors to HLM models resulted from qualities of the students not measured in this study, such as prior experience with online courses, personality traits, or writing skills. Some of these traits could be measured and included in future models.

A final research direction worth pursuing would be to investigate "outlier" courses. For instance, as illustrated in Figure 1, there were only three courses in which more than 20 incidents of three or more predictive variables occurred. Likewise, there were several courses where any predictive variable occurred only five or fewer times. Grouping courses into "high," "medium," and "low" incident courses might yield further understanding into the role of the predictive variables in student outcomes.

Conclusions

The purpose of this study was to identify instructor behaviors that lead to positive student outcomes in online courses. "Positive student outcomes" was defined in four ways, and the variables that were significant predictors did vary depending on the outcome definition. Student-student interaction, formative student behaviors, and immediacy were positive predictors of course grade; while building student capacity was a negative predictor of course grade. Student-student interaction, student-teacher interaction, formative student behavior, immediacy, and varied teaching activities were positive predictors of assignment grade; while number of formative activities and

preprogrammed instructor communication were negative predictors of assignment grade. Student-student interaction, formative student behaviors, and immediacy were positive predictors of retention. The data did not allow for a conclusion concerning the predictors of student satisfaction, but it is clear that satisfaction varies greatly across courses.

This study was an exploratory one, since most of these variables had not been tested empirically in an online environment. It is important that future research build upon studies like this one. Distance learning is an increasingly common way of providing instruction (see for instance Beaubien, 2002; USED, 2009), and it may all but replace face-to-face instruction within the next decade (Christensen, Horn, Caldera et al., 2011; NCCCS, 2010). The rise of distance learning is occurring within a cultural context that has changed not only what students must learn (Batson, 2010; Lumina Foundation, 2011; Symonds et al., 2011), but also how higher education is perceived by students, by the public, and by decision makers (Tulinko et al., 2005). Simultaneously, decisions about how to develop online courses are often made with little attention paid to sound instructional practices. Partially, this is because little empirical research has been conducted on what elements make distance learning effective, but it is also because the primary focus of educational leaders has been on managing the growth of online education at the expense of pedagogical concerns (Picciano & Seaman, 2010, p. 24).

While only 10% of postsecondary students took an online course in 2002, it is estimated that 50% will do so in 2014 (Christensen, Horn, Caldera et al., 2011, p. 31). Clearly, this rapid growth of a new way of instruction requires close scrutiny. However, there are few empirical studies available to course developers that indicate which teaching techniques work best in an online environment, even though 94% of institutions

report that they develop their own distance learning courses (National Center for Education Statistics, 2008, p. 3).

The need for effective online instruction has never been greater. Distance and technology now dominate not only how students learn, but also what they need to learn. A Lumina Foundation (2011) report highlighted what students, educators, and educational leaders can expect in the immediate future. "Higher learning has taken on new importance in today's knowledge society. To succeed in the contemporary workplace, today's students must prepare for jobs that are rapidly changing, use technologies that are still emerging and work with colleagues from (and often in) all parts of the globe." Distance learning is increasingly being seen as a way of providing the skill needed in this "contemporary workplace." Given the rapid growth of distance learning, educational leaders need to understand the factors that promote positive student outcomes in online courses so that they can ensure that courses taught at their institutions do in fact benefit the students who take them. This study helped to establish the relationship of instructor behaviors to student outcomes in online courses. It is hoped that this study will add to the discourse concerning effective distance learning instruction.

REFERENCES

- Abel, R. (2005, November 3). Implementing best practices in online learning. *Educause Quarterly*, 28(3), 75-79.
- Allen, I. E., & Seaman, J. (2005). *Growing by degrees: Online education in the United States, 2005*. Sloan Consortium.
- Allen, I. E., & Seaman, J. (2010). *Class differences: Online education in the United States, 2010*. Babson Survey Research Group/Sloan Consortium.
- Allen, I. E., & Seaman, J. (2011). *Going the distance: Online education in the United States 2011*. Babson Survey Research Group/Pearson/Kaplan University/Inside Higher Ed/Sloan-C.
- Alonso, F., Manrique, D., & Viñes, J. (2009, August). A moderate constructivist e-learning instructional model evaluated on computer specialists. *Computers & Education*, 53(1), 57-65. doi:10.1016/j.compedu.2009.01.002
- Anderson, T., Rourke, L., Garrison, D. R., & Archer, W. (2001). Assessing teaching presence in a computer conferencing context. *Journal of Asynchronous Learning Networks*, 5(2), 1-17.
- Andrade, M., & Bunker, E. (2009, May). A model for self-regulated distance language learning. *Distance Education*, 30(1), 47-61. doi: 10.1080/01587910902845956
- Arbaugh, J. B. (2008). Does the Community of Inquiry framework predict outcomes in online MBA courses? *International Review of Research in Open and Distance Learning*, 9(2), 1-21.
- Association of Public and Land Grant Universities. (2010). Presidents' symposium on the future of higher ed [Streaming media]. Washington, DC: George Mason

- University. Streamed from Association of Public and Land Grant Universities website. Retrieved from <http://www.aplu.org/page.aspx?pid=1643>
- Astleitner, H. (2002, June). Teaching critical thinking online. *Journal of Instructional Psychology*, 29(2), 53-76. Retrieved from Academic Search Premier database.
- Baker, B. K. (2009). Technology that makes a difference. *EdTech: Focus on Higher Education*, 5(2), 9.
- Batson, T. (2010, March 3). Jobs for graduates in the knowledge economy. *Campus Technology*. Retrieved from <http://campustechnology.com/articles/2010/03/03/jobs-for-graduates-in-the-knowledge-economy.aspx>
- Batson, T. (2011, April 6). Faculty 'buy-in'- to what? *Campus Technology*. Retrieved from <http://campustechnology.com/articles/2011/04/06/faculty-buy-in-to-what.aspx>
- Battalio, J. (2009). Success in distance education: Do learning styles and multiple formats matter? *American Journal of Distance Education*, 23(2), 71-87.
- Beaubien, J. (2002). Harnessing the power of complexity in an online learning environment. In K. E. Rudestam & J. Schoenholtz-Read (Eds.), *Handbook of online learning: Innovations in higher education and corporate training* (pp. 221-236). Thousand Oaks, CA: Sage.
- Berge, Z. L., & Huang, Y.-P. (2004). A model for sustainable student retention: A holistic perspective on the student dropout problem with special attention to e-learning. *DEOSNEWS*, 13(5). Retrieved from http://www.ed.psu.edu/acsde/deos/deosnews/deosnews13_5.pdf
- Bernard, R. M., Abrami, P. C., Borokhovski, E., Wade, A. C., 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.

- Blackboard, Inc. (2011). Blackboard Learn Feature Showcase. Retrieved from <http://www.blackboard.com/Platforms/Learn/Products/Blackboard-Learn/Teaching-and-Learning/Feature-Showcase.aspx>
- Bledsoe, C. A. (2008). *Distance education financial expenditures in North Carolina community colleges* (Unpublished doctoral dissertation). University of North Carolina, Greensboro. Retrieved from <http://libres.uncg.edu/ir/uncg/listing.aspx?styp=ti&id=182>
- Borreson, C., & Salaway, G. (2007). *Key findings: The ECAR study of undergraduate students and information technology, 2007*. Washington, DC: Educause Center for Applied Research.
- Brown, K. (2004, September). Technology: Building interaction. *TechTrends: Linking Research & Practice to Improve Learning*, 48(5), 36-38. Retrieved from Academic Search Premier database.
- Brown-Syed, C., Adkins, D., & Tsai, H. (2005, January 1). LIS student learning styles and web-based instruction methods. *Journal of Library & Information Services in Distance Learning*, 2(1), 5-26. (ERIC Document Reproduction Service No. EJ841344)
- Cameron, B. A., Morgan, K., Williams, K. C., Kostelecky, K. L. (2009). Group projects: Student perceptions of the relationship between social tasks and a sense of community in online group work. *American Journal of Distance Education*, 23(1), 20-33.

- Campbell, P. (2005, October 13). Re: Blackboard ate WebCT - the real costs of changing to OSS? [Electronic mailing list message]. Retrieved from BLKBRD-L - A list for Blackboard course administrators and faculty.
- Chetchumlong, S. (2010). The effect of web-based formative assessment on students' achievement and attitudes in a large enrolment tertiary English course (Unpublished doctoral dissertation). University of Southern Queensland, Queensland, Australia.
- Christensen, C. M., Horn, M. B., Caldera, L., & Soares, L. (2011). *Disrupting college: How disruptive innovation can deliver quality and affordability to postsecondary education*. Washington, DC: Center for American Progress / Innosight Institute.
- Christensen, C. M., Horn, M. B., & Johnson, C. W. (2011). *Disrupting class: How disruption will change the way the world learns*. New York, NY: McGraw Hill.
- Creswell, J. W. (2008). *Educational research: Planning conducting, and evaluating quantitative and qualitative research*. Upper Saddle River, NJ: Pearson/Merrill/Prentice Hall.
- Crisp, G. (2007). *E-assessment handbook*. New York, NY: Continuum.
- Davies, T. G., & Quick, D. (2001). Reducing distance through distance learning: The Community College Leadership doctoral program at Colorado State University. *Journal of Research and Practice*, 25, 607-620.
- Delaney, J., Johnson, A., Johnson, T., & Treslan, D. (2010). *Students' perceptions of effective teaching in higher education*. Board of Regents of the University of Wisconsin System. Retrieved from <http://www.uwex.edu/disted/conference>
- Division of Statistics and Scientific Computation, College of Natural Sciences. (n.d.)

- HLM. The University of Texas at Austin. Retrieved from http://ssc.utexas.edu/software/faqs/hlm#HLM_4
- Donavant, B. (2009, May). The new, modern practice of adult education: Online instruction in a continuing professional education setting. *Adult Education Quarterly*, 59(3), 227-245.
- Donohue, C. (2007, September). *Effective practice + instructional design = effective & engaging elearning*. Paper presented at the North Carolina Early Childhood/Education Curriculum Improvement Project, Raleigh, NC.
- Dougiamas, M. (2010, November). Keynote. Presented at the annual meeting of the North Carolina Community College Distance Learning Alliance Annual Conference, Raleigh, NC.
- Du, Y. (2004, January 1). Exploring the difference between concrete and abstract: Learning styles in LIS distance education. *Journal of Library & Information Services in Distance Learning*, 1(3), 51-64. (ERIC Document Reproduction Service No. EJ841669)
- Falvo, D. A. (2004, September/October). Constructing community in a graduate course about teaching with technology. *TechTrends: Linking Research & Practice to Improve Learning*, 48(5), 56-64.
- Farnsworth, K., & Bevis, T. B. (2006). *A fieldbook for community college online instructors*. Washington, DC: American Association of Community Colleges.
- Fox, C. (2010, June/July). Alabama gets on board. *T.H.E. Journal*. Retrieved from <http://thejournal.com/articles/2010/06/01/alabama-gets-on-board.aspx>
- Fox, C., & Donohue, C. (2006, November). Trends and promising practices in early

- childhood teacher education online: The view from New Zealand. *He Kupu: The Word*, 1(1), 29-34. Retrieved from <http://www.hekupu.ac.nz/index.php?type=journal&issue=3&journal=184>
- Gainor, S., Goins, R., & Miller, L. (2004, June). Using online modules is a multi-modality teaching system: A high-touch high-tech approach to geriatric education. *Gerontology & Geriatrics Education*, 24(4), 45-59. doi:10.1300/J021v24n04_04
- Garrison, D. R., Anderson, T., & Archer, W. (2000). Critical inquiry in a text-based environment: Computer conferencing in higher education. *The Internet and Higher Education*, 2(2-3), 87-105.
- Gillani, B. B. (2003). *Learning theories and the design of e-learning environments*. Lanham, MD: University Press of America.
- Githens, R. P., Crawford, F. L., & Sauer, T. M. (2010). *Online occupational education in community colleges: Prevalence and contextual factors*. Retrieved from National Research Center for Career and Technical Education website: http://136.165.122.102/mambo/component?option=com_remository/func,fileinfo/id,217
- Gokhale, A. A. (1995). Collaborative learning enhances critical thinking. *Journal of Technology Education*, 7(1). Retrieved from <http://scholar.lib.vt.edu/ejournals/JTE/v7n1/gokhale.jte-v7n1.html>
- Gordon, D. (2011, March 7). Return to sender. *T.H.E. Journal*. Retrieved from <http://thejournal.com/articles/2011/03/07/return-to-sender.aspx>
- Gozza-Cohen, M., & May, D. (2011, April). Traditional, online, and blended instruction: An investigation of outcomes in a graduate special education course taught in three different learning environments. Paper presented at the annual meeting of

American Educational Research Association, New Orleans, LA.

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

Haas, S., & Senjo, S. (2004). Perceptions of effectiveness and the actual use of technology- based methods of instruction: A study of California criminal justice and crime-related faculty. *Journal of Criminal Justice Education, 15*(2), 263-285. Retrieved from Academic Search Premier database.

Hatzipanagos, S., & Warburton, S. (2009, March). Feedback as dialogue: Exploring the links between formative assessment and social software in distance learning. *Learning, Media, & Technology, 34*(1), 45-59. doi:10.1080/17439880902759919

Haywood Community College. (2011a). College Catalog/Student Handbook. Clyde, NC: Author. Retrieved from <http://www.haywood.edu/pdf/Catalog%20files%2011/By%20Sections/WCatalog11-12.pdf>

Haywood Community College. (2011b). College facts. Clyde, NC: Author. Retrieved from http://www.haywood.edu/rie/college_facts

Haywood Community College. (2011c). End of course evaluation - 2011-12. Clyde, NC: Author.

Haywood Community College. (2011d). Who are our students? Clyde, NC: Author. Retrieved from http://www.haywood.edu/rie/who_are_our_students

Heiman, T. (2008). The effects of e-mail messages in a distance learning university on perceived academic and social support, academic satisfaction, and coping. *Quarterly Review of Distance Education, 9*(3), 237-248. Retrieved from Academic Search Premier database.

- Hill, J. R., Domizi, D. P., & Collier, N. D. (2011). Design and teaching: A bibliographic essay. As published in *Distances et Savoirs* (D&S-7/2009. A l croisee des recherches, pages 601 to 613). *American Journal of Distance Education*, 25(2), 91-102.
- Hiltz, S. R., & Turoff, M. (2005). Education goes digital: The evolution of online learning and the revolution in higher education. *Communications of the ACM*, 48(10), 59-64.
- Holder, B. (2007, October). An investigation of hope, academics, environment, and motivation as predictors of persistence in higher education online programs. *Internet & Higher Education*, 10(4), 245-260. doi:10.1016/j.iheduc.2007.08.002
- Houghton, J., & Sheehan, P. (2000). *A primer on the knowledge economy*. Melbourne City MC, Victoria, Australia: Centre for Strategic Economic Studies, Victoria University.
- HP/Intel. (n.d.). *Creating effective digital learning environments for higher education*. Retrieved from <http://guide2digitallearning.com/>
- Hrastinski, S. (2006, April). Introducing an informal synchronous medium in a distance learning course: How is participation affected? *Internet & Higher Education*, 9(2), 117-131. doi:10.1016/j.iheduc.2006.03.006
- Hsieh, P.-H., & Dwyer, F. (2009). The instructional effect of online reading strategies and learning styles on student academic achievement. *Educational Technology & Society*, 12(2), 36-50.
- Huett, J., Moller, L., Foshay, W., & Coleman, C. (2008, September). The evolution of distance education: Implications for instructional design on the potential of the

Web. *TechTrends: Linking Research & Practice to Improve Learning*, 52(5), 63-67. doi:10.1007/s11528-008-0199-9

Huhtanen, M., Ylä-Mella, J., Jeronen, E., & Keiski, R. (2008, December). A novel approach for undergraduate researcher education in engineering studies. *Journal of Baltic Science Education*, 7(3), 185-192. Retrieved from Academic Search Premier database.

Ice, P., Curtis, R., Phillips, P. & Wells, J. (2007). Using asynchronous audio feedback to enhance teaching presence and students' sense of community. *Journal of Asynchronous Learning Networks*, 11(2), 3-25.

Institut für Experimentelle Psychologie. (n.d.). G*Power (3) [Computer program]. Universität Düsseldorf: Author. Downloaded from <http://www.psych.uni-duesseldorf.de/abteilungen/aap/gpower3/>

Jackson, M., & Helms, M. (2008, September). Student Perceptions of Hybrid Courses: Measuring and Interpreting Quality. *Journal of Education for Business*, 84(1), 7-12. Retrieved from Academic Search Premier database.

Johnson, S. D., Aragon, S. R. Shaik, N., & Palma-Rivas, N. (2000). Comparative analysis of learner satisfaction and learning outcomes in online and face-to-face learning environments. *Journal of Interactive Learning Research*, 11(1), 29-49.

Jong, K. D., Moerbeek, M., & Van Der Leeden, R. (2010, May). A priori power analysis in longitudinal three-level multilevel models: An example with therapist effects. *Psychotherapy Research*, 20(3), 273-284.

Kelly, R. (Ed.) (2010). *Faculty Focus special report: Online course quality assurance: Using evaluations and surveys to improve online teaching and learning*. Madison,

WI: Faculty Focus/Magna Publications, Inc.

- Kelly, R. (Ed.) (n.d.). *Faculty Focus special report: Put to the test: Making sense of educational assessment*. Madison, WI: Faculty Focus/Magna Publications, Inc.
- Kerawalla, L., Minocha, S., Kirkup, G., & Conole, G. (2008, March). Characterising the different blogging behaviours of students on an online distance learning course. *Learning, Media, & Technology*, 33(1), 21-33. doi:10.1080/17439880701868838
- Kitsantas, A., & Chow, A. (2007, April). College students' perceived threat and preference for seeking help in traditional, distributed, and distance learning environments. *Computers & Education*, 48(3), 383-395. doi:10.1016/j.compedu.2005.01.008
- Kreft, I., & De Leeuw, J. (1998). *Introducing multilevel modeling*. London: Sage Publications.
- Lawrence, J., & Singhania, R. (2004, July). A study of teaching and testing strategies for a required statistics course for undergraduate business students. *Journal of Education for Business*, 79(6), 333-338. Retrieved from Academic Search Premier database.
- Levin, D., & Arafah, S. (2002). *The digital disconnect: The widening gap between internet savvy students and their schools*. Washington, DC: Pew Institute.
- Lim, J. H., Dannels, S. A., & Watkins, R. (2008). Qualitative investigations of doctoral students' learning experiences in online research methods courses. *Quarterly Review of Distance Education*, 9(3), 223-236.
- Lotkowski, V. A., Robbins, S. B., & Noeth, R. J. (2004). The role of academic and non-academic factors in improving college retention. *ACT Policy Report*. Retrieved

from http://www.act.org/research/policymakers/pdf/college_retention.pdf

Luke, D. A. (2004). *Multilevel modeling*. Thousand Oaks, CA: Sage Publications.

Lumina Foundation for Education, Inc. (2011). *The degree qualifications profile (beta)*.

Indianapolis: Lumina Foundation. Retrieved from the Lumina Foundation website

http://www.luminafoundation.org/publications/The_Degree_Qualifications_Profile.pdf

Lyons, J. (2004, August). Teaching U.S. history online: Problems and prospects. *History Teacher*, 37(4), 447-456. Retrieved from Academic Search Premier database.

Maas, C. J. M., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. (3), 86-92. doi: 10.1027/1614-1881.1.3.86

Macdonald, J. (2004). Developing competent e-learners: The role of assessment.

Assessment & Evaluation in Higher Education, 29(2), 215-226. *Methodology*, 1(3), 86-92. DOI

Mager, M., Heulett, S., & Karvonen, M. (2011, April). *Online discussion boards and student achievement: The impact of course design, instructor, and student variables*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Mager, M., Tignor, M., & Hipps, K. (2008). *Adapting hands-on courses for online delivery: A competency based approach*. Paper presented at Beyond the Horizon: North Carolina Community College System Conference 2008, Raleigh, NC.

Maki, R. H., Maki, W. S., Patterson, M., & Whittaker, P. D. (2000). Evaluation of a web-based introductory psychology course. *Behavior Research Methods, Instruments, & Computers*, 32(2), 230-239.

- Manochehri, N., & Young, J. I. (2006). The impact of student learning styles with web-based learning or instructor-based learning on student knowledge and satisfaction. *The Quarterly Review of Distance Education*, 7(3), 313-317.
- Mars, M. M., & Ginter, M. B. (2007). Connecting organizational environments with the instructional technology practices of community college faculty. *Community College Review*, 34(4), 324-343.
- Martinez, R., Liu, S., Watson, W., & Bichelmeyer, B. (2006). Evaluation of a web-based master's degree program: Lessons learned from an online instructional design and technology program. *The Quarterly Review of Distance Education*, 7(3), 267-283.
- McCoach, D. B. (2010). Hierarchical linear modeling. In G. R. Hancock & R. O. Mueller (Eds.), *The reviewer's guide to quantitative methods in the social sciences* (pp. 123-140). New York, NY: Routledge.
- Menchaca, M. P., & Bekele, T. A. (2008). Learner and instructor identified success factors in distance education. *Distance Education*, 29(3), 231-252.
- Moore, M. G. (1989). Three types of interaction. *American Journal of Distance Education*, 3(2), 1-6. Retrieved from http://www.ajde.com/Contents/vol3_2.htm
- Moore, M. G. (2002). Editorial. *American Journal of Distance Education*, 16(2), 61-64.
- Moore, M. G. (2007). *Handbook of distance education* (2nd ed.). Mahwah, NJ: Erlbaum.
- Nagel, D. (2010, March 5). National ed tech plan advocates radical reforms in schools. *T.H.E. Journal*. Retrieved from <http://thejournal.com/articles/2010/03/05/national-ed-tech-plan-advocates-radical-reforms-in-schools.aspx>
- Nagel, D. (2011, July 6). The disruption of blended learning. *T.H.E. Journal*. Retrieved from <http://thejournal.com/articles/2011/07/06/the-disruption-of-blended->

learning.aspx

National Center for Education Statistics. (2008). Distance education at degree-granting postsecondary institutions: 2006-2007. Washington, DC: U.S. Department of Education.

Neighmond, P. (2011). Think you're an auditory or visual learner? Scientists say it's unlikely. [Streaming media]. National Public Radio. Streamed from <http://www.npr.org/blogs/health/2011/08/29/139973743/think-youre-an-auditory-or-visual-learner-scientists-say-its-unlikely>

North Carolina Community College System. (2010, August 3). Distance education trends and emerging issues. Research brief presented at the Commission on Learning Technologies meeting, Raleigh, NC.

North Carolina Community College System. (2011). *Education Catalog*. NCCCS. Retrieved from http://www.nccommunitycolleges.edu/Programs/docs/Education-Catalog/Education_Catalog_19Sep2011.pdf

Offir, B., Lev, Y., & Bezalel, R. (2008, November). Surface and deep learning processes in distance education: Synchronous versus asynchronous systems. *Computers & Education, 51*(3), 1172-1183. doi:10.1016/j.compedu.2007.10.009

Oosterhof, A., Conrad, R-M., & Ely, D. P. (2008). *Assessing learners online*. Upper Saddle River, NJ: Pearson / Merrill / Prentice-Hall.

Ozkan, H. (2010). Cooperative learning technique through internet based education: A model proposal. *Education, 130*(3), 499-508. Retrieved from EBSCOhost.

Palloff, R. M., & Pratt. K. (2002). Beyond the looking glass: What faculty and students need to be successful online. In K. E. Rudestam & J. Schoenholtz-Read (Eds.),

Handbook of online learning: Innovations in higher education and corporate training (pp. 171-184). Thousand Oaks, CA: Sage.

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.

Pasher, H., McDaniel, M., Roherer, D., & Bjork, R. (2008, December). Learning styles: Concepts and evidence. *Psychological Science in the Public Interest, 9*(3), 105-119. doi: 10.1111/j.1539-6053.2009.01038.x

Picciano, A. G. (2002, July). Beyond student perceptions: Issues of interaction, presence, and performance in an online course. *Journal of Asynchronous Learning, 6*(1), 21-40.

Picciano, A. G., & Seaman, J. (Eds.) (2007). *K-12 online learning: A survey of U.S. school district administrators*. Babson Survey Research Group/Hunter College - CUNY/Sloan Consortium.

Picciano, A. G., & Seaman, J. (Eds.) (2008). *K-12 online learning: A 2008 follow-up of the survey of U.S. school district administrators*. Babson Survey Research Group/Hunter College - CUNY/Sloan Consortium.

Picciano, A. G., & Seaman, J. (Eds.) (2010). *Class connections: High school reform and the role of online learning*. Hunter College - CUNY/ Babson Survey Research Group.

Posner, M. A. (2011). The impact of a proficiency-based assessment and reassessment of learning outcomes system on student achievement and attitudes. *Statistics Education Research Journal, 10*(1), 3-14.

Powell, W. W., & Snellman, K. (2004). The knowledge economy. *Annual Review of*

Sociology, 30, 199-220.

- Precel, K., Eshet-Alkalai, Y., & Alberton, Y. (2009, April 1). Pedagogical and design aspects of a blended learning course. *International Review of Research in Open and Distance Learning*, 10(2). (ERIC Document Reproduction Service No. EJ844031)
- Ramaswami, R. (2009). Even! But no longer odd. *T.H.E. Journal*, 36(5), 39-44.
- Randall, B., Sweetin, J., & Steinbeiser, D. (2009). *Open source collaborative: Moodle assessment report*. Raleigh, NC: NCCCS. Retrieved from http://oscmoodlereport.files.wordpress.com/2009/08/osc_full_report.pdf
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: Sage Publications.
- Reuter, R. (2009). Online versus in the classroom: Student success in a hands-on lab class. *American Journal of Distance Education*, 23(3), 151-162.
- Rhode, J. F. (2009). Interaction equivalency in self-paced online learning environments: An exploration of learner preferences. *International Review of Research in Open and Distance Learning*, 10(1), 1-23.
- Richardson, J. C., & Swan, K. (2003). Examining social presence in online courses in relation to students' perceived learning and satisfaction. *Journal of Asynchronous Learning Networks*, 7(1), 68-88.
- Roberts, J. K. (2007, April). Group dependency in the presence of small intraclass correlation coefficients: An argument in favor of not interpreting the ICC. Paper presented at the annual meeting of American Educational Research Association. Retrieved from <http://www.hlm-online.com/papers/depend.pdf>

- Roschelle, J., Shechtman, N., Tatar, D., Hegedus, S., Hopkins, B., Empson, S., Knudsen, J., & Gallagher, L. P. (2010). Integration of technology, curriculum, and professional development for advancing middle school mathematics: Three large-scale studies. *American Educational Research Journal*, 47(4), 833-878.
- Rourke, L., Anderson, T., Garrison, D. R., & Archer, W. (2001). Assessing social presence in asynchronous text-based computer conferencing. *Journal of Distance Education*, 14 (2).
- Rudestam, K. E., & Schoenholtz-Read, J. (Eds.) (2002). *Handbook of online learning: Innovations in higher education and corporate training*. Thousand Oaks, CA: Sage.
- Ruhe, V. & Zumbo, B. D. (2009). *Evaluation in distance education and e-learning: The unfolding model*. New York, NY: Guilford Press.
- Schulte, B. (2011, March/April). Hybrid schools for the igeneration. *Harvard Education Letter*, 27(2), 1-4.
- Sher, A. (2004). *Assessing the relationships of student-instructor and student-student interaction with student learning and satisfaction in Web-based distance learning programs* (Unpublished doctoral dissertation). George Washington University, Washington, DC.
- Shu, P-C. (2005, Jan.). Identification of learning styles online by observing learners' browsing behaviour through a neural network. *British Journal of Educational Technology*, 36(1), 43-56.
- Sims, R., Dobbs, G., & Hand, T. (2002). Enhancing quality in online learning: Scaffolding planning and design through proactive evaluation. *Distance Education*, 23(2),

135-148.

Singer, J. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Education and Behavioral Statistics*, 24 (4), 323-355.

Smith, J. (2009). Benefits abound with online learning. *EdTech: Focus on Higher Education*, 5(2), 7.

Smith, P. L., & Ragan, T. J. (2005). *Instructional design* (3rd ed.). Hoboken, NJ: John Wiley & Sons.

SSI Inc. (n.d.a). HLM (6) [Computer program]. Author.

SSI Inc. (n.d.b). Use of deviance statistics for comparing models [Online guide for HLM software]. Author. Retrieved from http://www.ssicentral.com/hlm/help6/faq/Use_of_deviance_statistics_for_comparing_models.pdf

Stacey, E., & Rice, M. (2002). Evaluating an online learning environment. *Australian Journal of Educational Technology*, 18(3), 323-340. Retrieved from <http://www.ascilite.org.au/ajet/ajet18/stacey.html>

Staker, H. (2011). *The rise of K-12 blended learning*. Washington, DC: Charter School Growth Fund / Innosight Institute.

Stevens-Long, J., & Crowell, C. (2002). The design and delivery of interactive online graduate education. In K. E. Rudestam & J. Schoenholtz-Read (Eds.), *Handbook of online learning: Innovations in higher education and corporate training* (pp. 151-170). Thousand Oaks, CA: Sage.

Swan, K. (2003). Learning effectiveness: What the research tells us. In J. Bourne & J. C. Moore (Eds) *Elements of quality online education, practice and direction*.

- Needham, MA: Sloan Center for Online Education, 13-45.
- Swan, K. (2004). *Relationships between interactions and learning in online environments*. Newburyport, MA: The Sloan Consortium.
- Swan, K. P., Richardson, J. C., Ice, P., Garrison, D. R., Cleveland-Innes, M., & Arbaugh, J. B. (2008). Validating a measurement tool of presence in online communities of inquiry. *e-Mentor*, 2(24), 1-12.
- Symonds, W., Schwartz, R. & Ferguson, R. (2011). *Pathways to prosperity: Meeting the challenge of preparing young Americans for the 21st Century*. Cambridge, MA: Harvard Graduate School of Education. Retrieved from Harvard Graduate School of Education website. http://www.gse.harvard.edu/news_events/features/2011/Pathways_to_Prosperty_Feb2011.pdf
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). New York, NY: Pearson.
- Tulinko, J. (Senior Producer) & Glasser, C., Heus, J., Isaacs, S., & Wald, D. (Producers). (2005). *Declining by degrees: Higher education at risk* [Streaming media]. New York, NY: PBS. Streamed from NCLive Media Collection. http://0-media.nclive.org.wncln.wncln.org/content/begin_stream.php (WCU URL)
- Ulrich, J. W. (2009). *Test of a structural model to investigate the impact of instructor knowledge, attitudes and contextual constraints on intent to use Web 2.0 in online courses* (Unpublished doctoral dissertation). Western Carolina University, Cullowhee, NC.
- U.S. Department of Education, Office of Planning, Evaluation, and Policy Development. (2009). *Evaluation of evidence-based practices in online learning: A meta-*

analysis and review of online learning studies. Washington, DC: U.S. Department of Education. Retrieved October 1, 2009, from

<http://www.ed.gov/rschstat/eval/tech/evidence-based-practices/finalreport.pdf>

Virkus, S. (2008, July). Use of Web 2.0 technologies in LIS education: Experiences at Tallin University, Estonia. *Program: Electronic Library & Information Systems*, 42(3), 262-274. doi:10.1108/00330330810892677

Xu, D., & Jagers, S. S. (2011a). Online and hybrid course enrollment and performance in Washington State community and technical colleges (CCRC Working Paper No. 31). Retrieved from Community College Research Center website: <http://ccrc.tc.columbia.edu/Publication.asp?UID=872>

Xu, D., & Jagers, S. S. (2011b). The effectiveness of distance education across Virginia's Community colleges: Evidence from introductory college-level math and English courses. *Educational Evaluation and Policy Analysis*, 33(3), 360-377.

APPENDICES

Appendix A: Original Operational Definitions for Use in Pilot Phase 2

Appendix B: Instructions Given to Reviewer 1 for Use in Pilot Phase 3

Appendix C: Sample of Data Produced by Neutral Reviewer in Pilot Phase 3

Appendix D: Instructions Given to Reviewer 2 for Use in Pilot Phase 5

Appendix E: Data from Pilot Phase 6

Appendix A:

Original Operational Definitions for Use in Pilot Phase 2

Constructivist Teaching Methods

Identify activities in the course that allow or require students, either individually or in groups, to actively acquire learning materials from non-course locations and resources of their choice and to modify their roles as learners in order to determine how best to use those resources.

Student-Student Interaction (s-s)

Identify incidents of student-to-student communication within Moodle (e.g., within a discussion board).

Student-Teacher Interaction (s-t)

Identify incidents of student-to-teacher or teacher-to-*a particular student or subset of students* within Moodle. Note that communication from teacher to the entire class does not qualify.

Student-Content Interaction (s-c)

Identify activities in the course that include a triggering event (some story, question, or other stimulus that produces a sense of puzzlement), an opportunity for exploration of the trigger (students acquire or exchange information), an opportunity for integration (students connect the ideas and information acquired in phase two), and an opportunity for resolution (students apply the phase three integration to some new situation or stimulus).

Building Student Capacity (capacity)

Identify information or activities provided by the teacher with the expressed intention of preparing students for online learning or of increasing students' ability to accomplish increasingly difficult cognitive tasks.

Formative Assessment (formative)

Identify activities in which students receive evaluative feedback but which either do not count for a grade, count only for a complete/incomplete grade, or include opportunities for re-submission for an improved grade.

Measurable Learning Objectives (objectives)

Identify assignments that include a clear description of the skills or knowledge that students will acquire through a given exercise, a clear description of how they will demonstrate that acquisition for the purpose of a grade, and a clear description of the conditions under which that demonstration will occur.

Varied Teaching Activities (activities)

Identify the *kinds of activity required of students by teachers*. For instance, required textbook reading will count as one activity, required online reading will count as a second, required research will count as a third, required participation in a discussion forum as a fourth, required watching of a vodcast as a fifth, etc.

Varied sensory modalities (modalities)

Identify any activities in which two or more means of participation (visual, auditory, tactile, smell, taste) occur within a single course activity. Do not count reading, however.

Preprogrammed instructor communication (frequent)

Identify communication by the instructor to the entire class based on some pre-designated pattern that would be followed regardless of individual student behavior or activity in the course (e.g., preprogrammed announcements).

Synchronous Instruction (synchronous)

Identify incidents of student's use of real time communication with instructor or students as part of a given course activity.

Immediacy (immediacy)

Identify incidents of instructor's use of praise, solicitation of viewpoints, humor, or self-disclosure when that behavior is directed at individual students or groups in response to those students' behaviors.

Appendix B:

Instructions Given to Reviewer 1 for Use in Pilot Phase 3

Please evaluate only the first four content areas in 2011FA-PSY-281-IN1. The goal is to identify each incident of the 12 kinds of activities listed below and record them in the attached Excel spreadsheet.

Use your best judgment, but please make note of any questions that you have.

For all items, simply type the descriptor from Moodle (e.g., the Resource name or Quiz name) in the appropriate column. The goal is for a person to quickly be able to find that activity later, so you may want to identify the week or topic if the same descriptor occurs repeatedly (for instance, Topic 1 Study Guide, Topic 2 Study Guide). Descriptors may be used in more than one column.

Constructivist Teaching Methods (construct)

Identify activities in the course that allow or require students, either individually or in groups, to actively acquire learning materials from non-course locations and resources of their choice and to choose how best to integrate those resources to accomplish some task which does not have a clearly defined objective. This distinguishes between a traditional research paper, which is largely non-constructivist (usually limiting acceptable sources and requiring that the "objective" meet very specific and narrow criteria), and a project in which the student(s) define the objective and may approach that objective in a multitude of ways.

Student-Student Interaction (s-s)

Identify incidents of student-to-student communication within Moodle (e.g., within a discussion board).

Student-Teacher Interaction (s-t)

Identify incidents of student-to-teacher or teacher-to-*a particular student or subset of students* within Moodle. Note that communication from teacher to the entire class does not qualify.

Student-Content Interaction (s-c)

Identify activities in the course that include a triggering event (some story, question, or other stimulus that produces a sense of puzzlement), an opportunity for exploration of the trigger (students acquire or exchange information), an opportunity for integration (students connect the ideas and information acquired in phase two), and an opportunity for resolution (students apply the phase three integration to some new situation or stimulus).

Building Student Capacity (capacity)

Identify information or activities provided by the teacher with the expressed intention of preparing students for online learning. Two examples are technology skill building exercises and tutorials about how to navigate the course.

Formative Assessment (formative)

Identify activities in which students receive evaluative feedback but which either do not count for a grade, count only for a complete/incomplete grade, or include opportunities for re-submission for an improved grade.

Measurable Learning Objectives (objectives)

Identify assignments that include a clear description of the skills or knowledge that students will acquire through a given exercise and a clear description of how they will demonstrate that acquisition for the purpose of a grade.

Varied Teaching Activities (activities)

Identify the *kinds of activity required of students by teachers*. For instance, required textbook reading will count as one activity, required online reading will count as a second, required research will count as a third, required participation in a discussion forum as a fourth, required watching of a vodcast as a fifth, etc.

Varied sensory modalities (modalities)

Identify any activities in which two or more means of participation (visual, auditory, tactile, smell, taste) occur within a single course activity. Do not count reading, however.

Preprogrammed instructor communication (frequent)

Identify communication by the instructor to the entire class based on some pre-designated pattern that would be followed regardless of individual student behavior or activity in the course (e.g., preprogrammed announcements).

Synchronous Communication (synchronous)

Identify incidents of student's use of real time communication with instructor or students as part of a given course activity.

Immediacy (immediacy)

Identify incidents of instructor's use of praise, solicitation of viewpoints, humor, or self-disclosure when that behavior is directed at individual students or groups in response to those students' behaviors (not when directed at the entire class).

Appendix C:

Sample of Data Produced by Reviewer 1 in Pilot Phase 3

Phase 3 of the pilot of operational definitions resulted in the data below (only the first 15 lines of data are illustrated here). The shaded cells indicate the areas where Reviewer 1 and the author disagreed when using the second iteration of operational definitions (Appendix B). Those definitions were revised to produce the ones used for this study and validated in phase 5 (Appendix D).

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	construct	prpr	prpr	prpr	capacity	formative	objectives	approaches	modulatioz	request	synchrooz	immediacy	
2	Direct Project (Introduction to the course?) Forum	Introduction to the Course - Why are you taking this course? Forum	Introduction to the Course - Why are you taking this course? Forum	Module 1- as a whole, starting with introductory video through all the activities and quizzes	HOP Assignment	Module 1- M1 Quiz 1		Introduction to the Course - Syllabus Exam	Introduction to the Course - Video Introduction to the Course	Introduction to the Course - Course Syllabus		Introduction to the Course - Why are you taking this course? Forum	
3		Module 1- Question 1 Forum	Module 1- Question 1 Forum	Module 2- as a whole, starting with introductory video through all the activities and quizzes	Introduction to the Course - Messages Tool	Module 1- M1 Quiz 2		Introduction to the Course - News Report	Module 1- The Modules of the Course Resources	HOP Assignment		Module 1- Question 1 Forum	
4		Module 1- Question 2 Forum	Module 1- Question 2 Forum	Module 3- as a whole, starting with introductory video through all the activities and quizzes	Introduction to the Course - Video Introduction to the Course	Module 2- M2 Quiz 1			Module 1- About this Module	Introduction to the Course - Course Syllabus		Module 1- Question 2 Forum	
5		Module 1- Question 3 Forum	Module 1- Question 3 Forum	Module 4- as a whole, starting with introductory video through all the activities and quizzes	Introduction to the Course - Course Syllabus	Module 2- M2 Quiz 2			Module 1- Video: Early Treatment of Mental Disorders	Introduction to the Course - Project Choices		Module 1- Question 3 Forum	
6		Module 2, Question 1: Useful Models	Module 2, Question 1: Useful Models			Module 2- M2 Quiz 3			Module 2- About this Module	Introduction to the Course - Module Choices		Module 2, Question 1: Useful Models	
7		Module 2, Question 2: Useless Models	Module 2, Question 2: Useless Models			Module 2- M2 Quiz 4			Module 3- About this Module	Introduction to the Course - Research		Module 2, Question 2: Useless Models	
8		Module 3, Question 1: Reaction to Roy Tooms' Story	Module 3, Question 1: Reaction to Roy Tooms' Story			Module 2- M2 Quiz 5			Module 2- Ethical Concerns Raised by the Biological Model			Module 3, Question 1: Reaction to Roy Tooms' Story	
9		Module 3, Question 2: More About Assessment Techniques	Module 3, Question 2: More About Assessment Techniques			Module 2- M2 Quiz 6			Module 1- Introduction	Module 1- The Modules of the Course Resources		Module 3, Question 2: More About Assessment Techniques	
10		Module 3 - Module 3, Question 3: Choosing a Therapist	Module 3 - Module 3, Question 3: Choosing a Therapist			Module 2- M2 Quiz 7			Module 2- Video: Larry	Module 1- Introduction, Part 2		Module 3 - Physiological Events at the Neural Synapse	
11						Module 3- M3 Quiz 1			Module 2 - More about Social and Cultural Models	Module 1 - About this Module		Module 3 - Question 3: Choosing a Therapist	
12						Module 3 - M3 Quiz 2			Module 3 - Video: Mark	Module 1 - Modern Definitions of Abnormality			
13						Module 3 - M3 Quiz 3			Module 3 - Video: Marika	Module 1 - Historical Definitions of Abnormality			
14						Module 3 - M3 Quiz 4			Module 3 - Video: Julia and David	Module 1 - A Story About Abnormality... Josh Horton			
15						Module 3 - M3 Quiz 5			Module 3 - Video: Marital Therapy	Module 1 - The Treatment of "Abnormality"			
16						Module 3 - M3 Quiz 6				Module 1 - The Treatment of "Abnormality" Part 2			
										Module 1 - The Field of Abnormal Psychology			

Appendix D:

Instructions Given to Reviewer 2 for Use in Pilot Phase 5

Please evaluate only the first four content areas in 2011FA-PSY-281-IN1. The goal is to identify each incident of the 12 kinds of activities listed below and record them in the attached Excel spreadsheet.

Use your best judgment, but please make note of any questions that you have.

For all items, simply type the descriptor from Moodle (e.g., the Resource name or Quiz name) in the appropriate column. The goal is for a person to quickly be able to find that activity later, so you may want to identify the week or topic if the same descriptor occurs repeatedly (for instance, Topic 1 Study Guide, Topic 2 Study Guide). Descriptors may be used in more than one column.

Constructivist Teaching Methods (construct)

Identify activities in the course that allow or require students, either individually or in groups, to actively acquire learning materials from non-course locations and resources of their choice and to choose how best to integrate those resources to accomplish some task which does not have a clearly defined objective. This distinguishes between a traditional research paper, which is largely non-constructivist (usually limiting acceptable sources and requiring that the "objective" meet very specific and narrow criteria), and a project in which the student(s) define the objective and may approach that objective in a multitude of ways.

Student-Student Interaction (s-s)

Identify incidents of student-to-student communication within Moodle (e.g., within a discussion board).

Student-Teacher Interaction (s-t)

Identify incidents of student-to-teacher or teacher-to-*a particular student or subset of students* within Moodle. Note that communication from teacher to the entire class does not qualify.

Student-Content Interaction (s-c)

Identify activities in the course that include a triggering event (some story, question, or other stimulus that produces a sense of puzzlement), an opportunity for exploration of the trigger (students acquire or exchange information), an opportunity for integration (students connect the ideas and information acquired in phase two), and an opportunity for resolution (students apply the phase three integration to some new situation or stimulus).

Building Student Capacity (capacity)

Identify information or activities provided by the teacher with the expressed intention of preparing students for online learning. Two examples are technology skill building exercises and tutorials about how to navigate the course.

Formative Assessment (formative)

Identify activities in which students receive evaluative feedback but which either do not count for a grade, count only for a complete/incomplete grade, or include opportunities for re-submission for an improved grade.

Measurable Learning Objectives (objectives)

Identify assignments that include a clear description of the skills or knowledge that students will acquire through a given exercise and a clear description of how they will demonstrate that acquisition for the purpose of a grade.

Varied Teaching Activities (activities)

Identify the *kinds of activity assigned to students by teachers*. For instance, assigned textbook reading will count as one activity, assigned online reading will count as a second, assigned research will count as a third, assigned participation in a discussion forum as a fourth, assigned watching of a vodcast as a fifth, etc.

Varied sensory modalities (modalities)

Identify any activities in which two or more means of participation (visual, auditory, tactile, smell, taste) occur within a single course activity. Do not count reading, however.

Preprogrammed instructor communication (frequent)

Identify communication by the instructor to the entire class based on some pre-designated pattern that would be followed regardless of individual student behavior or activity in the course (e.g., preprogrammed announcements). Such communication should not convey course content, but should instead be facilitative or supportive. For instance, web pages that become active at a given date but that are focused on material that students are expected to learn for a quiz would not count, but an email to the entire class reminding them of a due date or offering a suggestion about how to find appropriate resources for a research paper *would* count. Most or all of this communication will be in the form of messages, emails, or announcements.

Synchronous Communication (synchronous)

Identify incidents of student's use of real time communication with instructor or students as part of a given course activity.

Immediacy (immediacy)

Identify incidents of instructor's use of praise, solicitation of viewpoints, humor, or self-disclosure when that behavior is directed at individual students or groups in response to those students' behaviors (not when directed at the entire class).

Appendix E: Data from Pilot Phase 6

Phase 6 of the pilot of operational definitions resulted in the data below. The first three courses were assessed by Reviewer 1 and then discussed with the author. Reviewer 1 then assessed the latter two courses. Chapter Four discusses this process in detail.

(Where the name of an assignment would have revealed the course identity, XXX has been used to mask that name.)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
	Course	Content area	Item	student directed	s-s	s-t	s-o	capacity	formative	objectives	approaches	modalities	programmed	synchronous	immediacy			
1																		
2	Course 1	Content area 3	XXX Report due				1				1	1						
3			XXX Report Guidelines				1				1	1						
4			XXX Report turn in								1							
5	Course 2	Content area 5	Week 5 text introduction material				1						1					
6			Unit 2A - voice and slides				1					1						
7			Unit 2A - slides only				1											
8			Unit 2B - voice and slides				1					1						
9			Unit 2b - slides only				1											
10			Week 05 Review Questions				1	1	1		1							
11			XXX2				1				1					1		
12			Discussion 4		1						1							
13	Course 3	Week 9	Week Nine Assignment	1			2	1			1							
14			Week Nine Web Review					1			1							
15			Week Nine Discussion		1						1							
16			Chapter Nine Quiz								1							

These three courses were assessed by Reviewer 1 first and then discussed with author. Green cells were agreed upon prior to discussion and both agreed that they met the operational definitions.

Yellow cells were initially disagreed upon, but upon discussion Reviewer 1 agreed that they did not meet operational definitions. Note that in the latter two courses, assessed after the earlier discussion, there was no disagreement on these items.

Red cells indicate disagreement on interpretation of the operational definitions.

17	Course 4	Content area 2	Introductory text at the top					1					1						
18			Assign 1				1				1								
19			Assign 2				1				1								
20			Assign 3	1			1				1								
21			Paper Assign	Same content as Assign 1, linked from 2 locations. Reviewer 1 and author agreed on Assignment 1.															
22			General Discussion Forum		1						1								
23			Defining XXX - Gp1		1						1								
24			Defining XXX - Gp2		1						1								
25			Defining XXX - Gp3		1						1								
26			Defining XXX - Gp4		1						1								
27	Course 5	Content area 5	Introductory text at the top										1						
28			M3 Reading Assign								1								
29			M3 Forum		1						1								
30			M3 Developing Understanding								1								
31			M3 Writing Assignment		1						1								
32			M3 Research								1								
33			M3 Quiz								1								

These two courses were assessed by Reviewer 1 after discussion concerning the first three had occurred. Agreement was much more consistent.

Green cells indicate agreement about the presence of a predictive variable and red cells indicate disagreement.