Retention of Community College Students in Online Courses

Sarah Krajewski
Western Michigan University, sarahkrajewski@yahoo.com

Follow this and additional works at: https://scholarworks.wmich.edu/dissertations
Part of the Community College Leadership Commons, Higher Education Commons, and the Online and Distance Education Commons

Recommended Citation
Krajewski, Sarah, "Retention of Community College Students in Online Courses" (2015). Dissertations. 1180.
https://scholarworks.wmich.edu/dissertations/1180

This Dissertation-Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Dissertations by an authorized administrator of ScholarWorks at WMU. For more information, please contact maira.bundza@wmich.edu.
RETENTION OF COMMUNITY COLLEGE STUDENTS IN ONLINE COURSES

by

Sarah Krajewski

A dissertation submitted to the Graduate College in partial fulfillment of the requirements for the degree of Doctor of Philosophy Educational Leadership, Research and Technology Western Michigan University December 2015

Doctoral Committee:

Brandy Skjold, Ph.D., Chair
William Cobern, Ph.D.
John Geiser, Ph.D.
The issue of attrition in online courses at higher learning institutions remains a high priority in the United States. A recent rapid growth of online courses at community colleges has been instigated by student demand, as they meet the time constraints many nontraditional community college students have as a result of the need to work and care for dependents. Failure in an online course can cause students to become frustrated with the college experience, financially burdened, or to even give up and leave college. Attrition could be avoided by proper guidance of who is best suited for online courses. This study examined factors related to retention (i.e., course completion) and success (i.e., receiving a C or better) in an online biology course at a community college in the Midwest by operationalizing student characteristics (age, race, gender), student skills (whether or not the student met the criteria to be placed in an AFP course), and external factors (Pell recipient, full/part time status, first term) from the persistence model developed by Rovai. Internal factors from this model were not included in this study. Both univariate analyses and multivariate logistic regression were used to analyze the variables.

Results suggest that race and Pell recipient were both predictive of course completion on univariate analyses. However, multivariate analyses showed that age, race, academic load and first term were predictive of completion and Pell recipient was no longer predictive. The
univariate results for the C or better showed that age, race, Pell recipient, academic load, and meeting AFP criteria were predictive of success. Multivariate analyses showed that only age, race, and Pell recipient were significant predictors of success. Both regression models explained very little (<15%) of the variability within the outcome variables of retention and success. Therefore, although significant predictors were identified for course completion and retention, there are still many factors that remain unaccounted for in both regression models. Further research into the operationalization of Rovai’s model, including internal factors, to predict completion and success is necessary.
ACKNOWLEDGMENTS

I would like to thank my dissertation chair, Dr. Brady Skjold for mentoring and guiding me through the final stages of my degree. I appreciate her intellect, I am grateful for her quick feedback, and I value her friendship. I am also grateful to Dr. Cobern and Dr. Geiser for serving on my committee.

During this time in the doctoral program I gained wonderful lifelong friends. Leah and Tracy I am very grateful for your endless assistance and support! When I think of all the writing and thinking sessions we shared together I smile. I know that I could never have completed this degree without each of you! You were both key to my success and made the journey fun. Tracy, thank you! I could not have completed this dissertation without you. You are amazing! Leah, thank you! I never would have made it through the ERR requirement without our late Thursday night writing sessions and your willingness to guide me through the numerous procedures and processes that I missed by not being a full time student. Lauren, thank you! I appreciate all of your support and encouragement at each one of my presentations. I cannot express to you how much it meant for me to have you there. I would also like to acknowledge and thank Dr. Davis for his endless consultation and guidance. I greatly appreciate his assistance, and willingness to help me. He is a true intellect and professional.
Acknowledgements—Continued

Lastly, and most importantly, I would like to thank my loving family. John, you are an amazing husband and my best friend. Your undying support, encouragement, patience and love is why I was successful in this venture. I am so blessed to have you in my life. I want to thank my three wonderful children; Danny, Jack, and Katie. Each of you gave me the inspiration and confidence I needed to complete this degree. I’m so happy I have you! Time for a vacation!

Sarah Krajewski
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS........................................................................................................... ii

LIST OF TABLES......................................................................................................................... viii

LIST OF FIGURES....................................................................................................................... ix

CHAPTER

I. INTRODUCTION .................................................................................................................. 1
   Conceptual Framework of College Retention and Persistence................................. 3
   Development of Theoretical Models ..................................................................................... 3
   Bean and Metzner Conceptual Model of Nontraditional Student Retention........................... 7
   Rovai’s Composite Persistence Model ................................................................................. 12

II. LITERATURE REVIEW ....................................................................................................... 17
   Introduction .............................................................................................................................. 17
   The Intersection of Retention, Community Colleges, and Online Courses .......... 17
   General Overview of Community Colleges ............................................................................. 20
   Demographic Characteristics of Community College Students ......................... 21
   Investigating Online Learning Retention Studies ......................................................... 23
      Online Student Success and Gender ................................................................................. 24
      Online Student Success and Age ....................................................................................... 25
      Online Student Success and Race/Ethnicity ................................................................. 25
# Table of Contents—Continued

## CHAPTER

<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Online Student Success and Academic Readiness</td>
<td>26</td>
</tr>
<tr>
<td>Online Student Success and Enrollment Status</td>
<td>27</td>
</tr>
<tr>
<td>Academically Underprepared Students and Online Education</td>
<td>27</td>
</tr>
<tr>
<td>Liu and Liu</td>
<td>29</td>
</tr>
<tr>
<td>Stahl and Pavel</td>
<td>38</td>
</tr>
<tr>
<td>Fike and Fike</td>
<td>42</td>
</tr>
<tr>
<td>Aragon and Johnson</td>
<td>47</td>
</tr>
<tr>
<td>Wojciechowski and Palmer</td>
<td>51</td>
</tr>
<tr>
<td>Bantum</td>
<td>55</td>
</tr>
<tr>
<td>Jaggars and Xu</td>
<td>59</td>
</tr>
<tr>
<td>Community College Research Center</td>
<td>60</td>
</tr>
<tr>
<td>Summary</td>
<td>62</td>
</tr>
<tr>
<td>Research Questions</td>
<td>64</td>
</tr>
<tr>
<td>III. METHODOLOGY</td>
<td>66</td>
</tr>
<tr>
<td>Purpose</td>
<td>66</td>
</tr>
<tr>
<td>Study Procedures</td>
<td>66</td>
</tr>
<tr>
<td>Study Sample Description</td>
<td>66</td>
</tr>
<tr>
<td>Study Variables</td>
<td>67</td>
</tr>
<tr>
<td>Data Analysis</td>
<td>67</td>
</tr>
<tr>
<td>Logistic Regression Assumption</td>
<td>70</td>
</tr>
</tbody>
</table>
Table of Contents—Continued

CHAPTER

Common Assumptions for Both Regression Models ........................................ 70

Assumptions Model 1: Course Completion .................................................. 71

Assumptions Model 2: C or Better ......................................................... 72

Missing Data ............................................................................................ 72

IV. RESULTS ............................................................................................ 74

Research Question 1: Univariate Results ................................................. 74

Research Question 2: Univariate Results ................................................. 75

Research Question 2: Logistic Regression Analysis ................................. 77

Research Question 1: Logistic Regression Analysis .................................. 79

Summary .................................................................................................. 82

V. CONCLUSION OF RESULTS ................................................................ 83

Introduction ............................................................................................ 83

Findings ................................................................................................... 85

Student Characteristics ........................................................................... 86

Race ......................................................................................................... 87

Age .......................................................................................................... 87

Gender ..................................................................................................... 88

Student Skills .......................................................................................... 88

External Factors ....................................................................................... 89

Pell Grant ................................................................................................. 90
# Table of Contents—Continued

## CHAPTER

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enrollment Status</td>
<td>90</td>
</tr>
<tr>
<td>First Semester Students</td>
<td>91</td>
</tr>
<tr>
<td>Model Evaluation</td>
<td>91</td>
</tr>
<tr>
<td>Limitations</td>
<td>95</td>
</tr>
<tr>
<td>Future Studies</td>
<td>95</td>
</tr>
<tr>
<td>Conclusions</td>
<td>96</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>98</td>
</tr>
<tr>
<td>APPENDIX</td>
<td></td>
</tr>
<tr>
<td>A. HSIRB Approval Letter</td>
<td>107</td>
</tr>
</tbody>
</table>
LIST OF TABLES

1. Predictor Variables to be Tested ................................................................. 68

2. Frequency: Missing Data ............................................................................... 73

3. Univariate Summary of Individual Student Characteristic for
   Complete/Did Not Complete ............................................................................. 75

4. Results for Comparisons Between Those Who Earned a C or Better
   and Those that Completed the Course But Did Not Receive a C or Better ....... 76

5. Logistic Regression Results for the Completion Outcome ............................. 78

6. Five-fold Cross-validation RMES Comparison for Completion
   Regression Model ............................................................................................. 79

7. Logistic Regression Results for the C or Better Outcome ............................... 80

8. Five-fold Cross-validation RMES Comparison for C or Better
   Regression Model ............................................................................................. 82
LIST OF FIGURES

1. Rovai’s (2002) composite persistence model .................................................. 15
2. Model of the literature represented in this proposal ........................................ 19
3. ROC curve for completion regression model .................................................... 78
4. ROC curve for c or better regression model ..................................................... 81
5. Rovai’s model for student persistence ............................................................. 86
6. New predictor model for success and persistence in online courses ................. 93
CHAPTER I
INTRODUCTION

The Organization for Economic Cooperation and Development (OECD) assessment of international educational performance reported that the U.S. holds the 15th spot in the number of students entering and completing higher education. This is a stark contrast to the 2nd rank the U.S. once held (Baum & Payea, 2005). In 2010, President Obama pledged to take the U.S. to first place among countries with the most college graduates by 2020 (Johnson-Ahorlu, Alvarex, & Hurtado, 2011). Community colleges enroll approximately 40 percent of all undergraduates therefore improving student success at community colleges is imperative to achieve these goals (Staklis, 2010). Six prominent community college organizations signed a Call to Action, a pledge to increase by 50% the number of students with high quality degrees and certificates by 2020 (Mullin, 2010).

To meet Obama’s pledge of increasing the number of students who graduate in the U.S., many community colleges have increased the number of courses they offer online. Unfortunately, there is little research on the relationship between the unique characteristics of community college students and their ability to be successful in the online course delivery format. With increased demand for online learning, as well as more institutions of higher learning striving to provide diverse educational opportunities, online course delivery continues to grow as a viable means of providing increased access to a greater number of students. Very little research targeting the retention of community college students in online
courses has been reported. The question arises: Is the practice of increasing online course offerings to all community college students an appropriate utilization of economic resources, and effective in promoting students’ educational goals?

A key concern for those in higher education is that community college students are often unprepared academically for college level coursework. Suggestive evidence reports that less-prepared students may fare better in face-to-face than in online courses (Peterson & Bond, 2004; Standford-Bowers, 2008). According to the Achieve the Dream initiative database, 59% of community college students enrolled in at least one developmental course during the three years that students were tracked (Bailey, Jeong, & Cho, 2008). Community colleges report lower success and persistence rates among academically underprepared students in online courses (Jaggars & Xu, 2010). Community college students who chose to take online courses are less likely to complete and perform well in those courses. Additionally, students who took online classes early in their academic career were less likely to return to school in subsequent semesters than students who did not take online courses (Jaggars & Xu, 2010). The findings represent differences in online and face-to-face outcomes based on descriptive data (Xu & Jaggars, 2010). The results also suggest that online courses may exacerbate already persistent achievement gaps between student subgroups. For example, students with lower prior GPAs, male students, and black students showed difficulty persisting and being successful in online learning environments (Xu & Jaggars, 2013). Students who did complete online course sections were found to be 3% to 7% less likely to receive a C or better than students who took the face-to-face sections (Jaggars & Xu, 2010; Xu & Jaggars, 2011).
As community colleges work towards achieving the goal President Obama set to increase student graduation 50% by 2020 (College Board, 2012) persistence studies continue to evolve and incorporate new variables into the research. Figure 2 illustrates a current gap in higher education literature on retention of online learners at the community college. Emerging research recognizes that gaps do exist in the literature, as is evidenced by the movement from the four year perspective to the two year college as well as from in seat to online course success and retention studies. Opportunities still exist for future explanation of the basic theoretical concepts of Tinto, Bean, Metzner, and Rovai.

Conceptual Framework of College Retention and Persistence

Student success in college is defined as both persistence and educational attainment, or achieving the desired degree or credential. Each of these perspectives is based on academic preparation and the quality of students’ experiences during college. This chapter will look at the theoretical frameworks, applied to college student retention and persistence in higher education, that have emerged over the past several decades. The development and design of both Tinto’s Model of Student Departure (1975, 1987, 1993) and Bean and Metzner’s Conceptual Model of Nontraditional Student Retention (1985) are paramount to understanding the model discussed in this study, Rovai’s Composite Persistence Model (2002).

Development of Theoretical Models

Early attempts to explain persistence in higher education were based on psychological models. These models theorized that a student’s decision to persist was largely based on
previous success, attitudes, and norms that drive behavior through the formation of intent to learn. More recent models are grounded in these psychological models but explain persistence and attrition through student-institution “fit” by looking at student, institutional, and environmental variables along with specific themes, such as social integration of students into campus activities. Multiple retention models are available (Bean & Metzner, 1985; Rovai, 2002), though most branch from Vincent Tinto’s theories (1975, 1987, 1993).

Vincent Tinto (1975, 1987, 1993) put forward possibly the most influential attempt to explain the process of persistence in higher education, as a function of student-institution “fit.” Tinto (1975) first designed a theoretical model to explain all of the aspects and processes that influenced an individual’s decision to leave higher education, and how these processes interact to produce attrition. He then modified his theoretical model twice (1987, 1993) based on additional findings.

Tinto was known for impacting data collection in higher education to distinguish different reasons for student departure. Tinto argued that the reasons students left school were very important to decipher because that information could be used to identify students at risk and provide services and assistance to help students successfully complete their courses, programs, and eventually earn their desired degrees or certificates. Today, colleges and universities classify and identify different types of leaving behavior based on Tinto’s work. These include academic failure, voluntary withdrawal, permanent dropout, temporary dropout, and transfer.

Tinto’s (1975) original model, the Model of Student Departure (MSD), was first presented in a literature review on higher education. The MSD was based on the first theoretical model of the dropout process proposed by Spady (1970). Both Spady and Tinto
were sociologists by training. Spady is known for his modification of Durkheim’s (1951) hypothesis of suicide, which he applied to student attrition. The basis of the model was the assertion that suicide is predicted by the individual’s level of integration into the fabric of society. Individuals commit suicide because they are insufficiently integrated into society. Conversely, if an individual has an adequate social support network and sufficient moral integration, suicide likelihood is reduced. Spady (1970) based his educational model of the college dropout process on identifying characteristics such as family background, economic status, ability, and academic performance as elements that effect dropout decisions. He introduced the concept that social integration leads to student satisfaction; which brings increased institutional commitment and decreased probability of student attrition. Spady further claimed student attrition occurs when the individual is insufficiently integrated into either academic or social aspects of college life. Vincent Tinto took Spady’s initial work and built upon it. Tinto formulated the Model of Student Departure (MSD), in which he stated it was important to leave room for the individual characteristics of a person that would make them more likely to drop out of higher education. Durkheim’s (1951), Van Gennep’s (1960) and Spady’s (1970) theories were the foundation of Tinto’s MSD.

Tinto (1993) identified different student groups by means of group-specific interventions and policies, such as minority students, students with disabilities, students from low-income families, adult students, and transfer students. Tinto’s model encompassed individuals’ characteristics as affecting students’ pre-enrollment commitment to both their goal (i.e. degree attainment) as well as their commitment to the higher education institution. Individual characteristics are highlighted in the model as being important in influencing the individual’s persistence decision (race, sex, and academic ability), precollege experiences
(grade point average, academic and social accomplishments), and family background (social status, value climates and expectation climates).

Tinto stressed that there is an influence of an individual’s educational expectations on their probability of attrition. This was defined as the length of time the student expected to attend the institution, along with the importance the student placed upon the specific institution in which they attended. A significant variance in how committed individual students are to their specific educational institution was also of importance. For example, a great connection to the college will yield increased chances to persist, despite academic or social problems.

The central concept to Tinto’s model is the level of student’s integration into the social and academic systems of the college. Integration is explained in the SI model by Tinto (1993) as the process in which determines persistence or dropout which a student establishes membership in the college community. The higher the degree of integration of the individual into the college system, the greater will be the commitment to the specific institution and to the goal of persistence. Social integration refers to the social ties that result from daily interactions with peers and instructors. Academic interactions result from the sharing of information, viewpoints, and values common to the other members of the community.

Tinto indicated that social integration in college is directly related to persistence. While lack of social integration would lead to attrition, it would be most likely voluntary attrition. This is a result of students not feeling that they ‘belong’ and are not motivated to continue their studies. Social integration is especially important in the first year. Students who are socially integrated are motivated to continue working academically to remain enrolled. If a student does not feel socially integrated, they are more likely to drop out.
However, if the student does not feel socially integrated toward the end of their educational journey then attrition is less likely. Tinto also states that high levels of social integration increase academic integration.

Tinto (1993) states that social and academic integration are in fact not separate and discrete, but are indivisible. They are two areas of importance at different points throughout a students’ academic development. Tinto’s model states that academic integration directly influences the student’s goal commitment, while social integration directly influences commitment to the specific institution.

Bean and Metzner Conceptual Model of Nontraditional Student Retention

Bean and Metzner (1985) conducted a literature review and presented their 1985 article as a conceptualization. They concluded by proposing their own model, “Conceptual Model of Nontraditional Student Retention”, which was essentially a revision of Tinto’s (1974, 1986, 1993) models to further explain attrition of nontraditional students. Their specific objective was understanding the implications of and factors affecting non-traditional student enrollment. Bean and Metzner felt as though Tinto’s model did not adequately fit the unique situations of nontraditional students. Bean and Metzner (1985) defined a nontraditional student as “older than 24, does not live in a campus residence, or is a part-time student, or some combination of these three factors; is not greatly influenced by the social environment of the institution, and is chiefly concerned with the institution’s academic offerings especially courses, certification and degrees” (p. 489). They criticized Tinto’s model as a theory that is only applicable to the homogeneity of a four year, on campus
college setting, with limited applicability to non-traditional students, commuter students, and community college students. They argued that although Tinto has the most cited and documented theory of attrition to date in higher education he had predominantly dealt with traditional students at four-year residential institutions and did not address the many issues that nontraditional students face that differ from traditional students.

Bean and Metzner (1985) specifically suggested Tinto’s model was not useful for studying the attrition of older students, distance education students, and part time students. These students were described as having limited interaction within the college community as a support structure and instead draw support from peers, friends, family, and employers that are outside the institution of higher learning. Academic and social integration within the institution may be less beneficial for these non-traditional students. Conversely, traditional students draw support from friends and faculty within the institution.

Age, enrollment status, and residence were used as defining variables in their model. Age was described as having an indirect effect on attrition because older students reported to have more hours of employment, family responsibilities, and increased absenteeism than younger students. Enrollment status (full-time or part-time students based on credit hours) was an important predictor of persistence. Part-time students were described as being much more likely to withdraw than full-time students (Argon & Johnson, 2008; Conklin, 1997; Grimes & Antworth, 1996; Rajasekhara & Hirsch, 2000). Distinctive from previous models in its treatment of environmental influences, the nontraditional undergraduate student attrition model suggests that issues such as family responsibility, hours employed, and ability to pay for college may have direct, negative influences on student persistence.
Residence was listed by Bean and Metzner (1985) as the most important distinction between traditional and nontraditional students. Commuter students spend less time on campus than traditional students and therefore typically have less contact with faculty as well as with other students, both academically and socially. As a whole, non-traditional students tend to not reside on college campus, have less contact with faculty members outside of class, fewer extracurricular activities, and fewer friends at college.

Educational goals, high school performance, ethnicity, and gender were the four sub-categories used to define background variables in Bean and Metzner’s (1985) Model of Student Attrition. Educational goals were defined as the highest level of college education the student hoped to attain, how the student considered attaining the college education and the level of commitment a student had to attaining that education at the institution they were attending. High school academic performance, grade point average (GPA), was listed as the strongest pre-enrollment predictor of college. Bean and Metzner suggested ethnicity of nontraditional students should have a strong negative influence on GPA, they based this on the belief that poorer education is provided for minority students at the secondary level (Bean, 1985). Last, they suggested that gender may have an indirect effect on attrition based on family responsibilities that are greater for women and also less opportunity to transfer (Prince & Jenkins, 2005).

Four factors are identified that affect persistence: (a) academic variables such as study habits, course availability, academic advising and attendance; (b) background and defining variables such as age, educational goals, ethnicity, gender, and prior GPA; (c) environmental variables such as financing, hours of employment, family responsibilities, outside encouragement and opportunity to transfer; and (d) academic outcome, GPA, and
psychological outcomes such as goal commitment, stress and satisfaction while at college. The other box titled ‘Intent to Leave’ is defined as the student’s level of commitment to continuing or dropping. Bean and Metzner propose exogenous and endogenous factors directly or indirectly influenced a student’s decision to remain in school or dropout.

Psychological outcomes (satisfaction, goal commitment, and stress) are directly influenced by environmental and academic variables. Satisfaction measures the degree of enjoyment resulting from the course. Stress is listed as a negative influence on persistence (Bean & Metzner, 1985; Pascarella and Terenzini, 1986). Intent to leave is marked as the strongest predictor of course persistence.

Social integration is present to a much smaller extent in the Bean and Metzner (1985) model than in Tinto’s (1986) model. It is included for individual situations, but is assigned little importance (in Figure 2 it is denoted by a small box). Bean and Metzner state that nontraditional students do not make persistence decisions based on social life.

Academic variables include study skills, absenteeism, course availability, and certainty of a major. In the Bean and Metzner model, academic variables (such as study hours) directly influence academic outcomes (GPA). Poor academic outcomes can lead to involuntary departure (dismissal). The two strongest predictors of persistence are GPA and intent to leave. Academic variables can also be factors in a voluntary departure decision. They may cause an indirect route by causing a negative psychological outcome variable to develop (stress, or weakened goal commitment). These negative psychological variables may then lead to the intent to leave, which is finalized in an actual decision to drop out. If students take the correct courses for their ability then their GPA will allow them to remain at the institution.
Environmental variables measure factors such as lack of finance, employment, lack of encouragement, and family responsibilities. The institution has little control over these. Bean and Metzner (1985) state that environmental factors were more important than the academic variables (which correspond to the academic institutional system in Tinto’s model) for non-traditional students. They proposed two compensatory effects: 1) environmental support compensate for weak academic support, but academic support will not compensate for weak environmental support; and 2) nonacademic factors compensate for low levels of academic success, while levels of academic achievement will only result in continued attendance when accompanied by positive psychological outcomes from school (p. 92). In other words, for non-traditional students what occurs in their lives off campus is more important that than what is happening for them on campus.

Bean and Metzner state in their conclusion that students’ self-reported financial difficulty, and concern about the ability to finance college were positively related to attrition from college, putting too much pressure on their time, resources, and sense of well-being. Bean and Metzner introduce background and defining variables early in the model. Important characteristics including age, enrollment status, and residence as well as educational goals, high school performance, ethnicity, and gender are all expected to affect the interaction between nontraditional students and the institution.

Bean and Metzner (1985) concluded that little research has been conducted as it pertains to student retention of non-traditional students. Future research about non-traditional students is needed and it should (1) include external environmental variables, (2) contain multivariate research models, and (3) not be based solely on social integration (Bean & Metzner, 1985).
In conclusion, Bean and Metzner adapted Tinto’s model to better fit the nontraditional student. The most significant difference between Bean and Metzner’s model and Tinto’s model of student attrition has to do with the role of the institutional environment as a socializing agent. While there is a large body of evidence to suggest social integration is appropriate for traditional college students, nontraditional students interact with the college environment both less regularly and less intensely. As a result, Bean and Metzner deemphasized the impact of socialization on student persistence decisions. They offer in its place an enhanced focus on environmental variables and psychological outcome variables. Despite all indications of Bean and Metzner’s model as a good fit for community college student success research, few scholars have tested its effectiveness.

Rovai’s Composite Persistence Model

Rovai’s (2002) Composite Persistence Model was designed to better explain retention of nontraditional and online learners. The Composite Persistence Model (Rovai, 2002) was based on a thorough review of Tinto’s Student Integration Model (1993) and Bean and Metzner’s Student Attrition Model (1985) which are considered the most comprehensive previous frameworks. Rovai’s Composite Model (2002) argues that students who take classes online have additional needs, and therefore do not fit into the other models. His model is adapted to the needs of the online learner, specifically on the factors affecting a student’s decision to drop out of online courses. He declared it to be a better predictor of the persistence of nontraditional, community college students enrolled in online coursework.
Rovai Composite Model (Figure 1) is organized into two major categories: Prior to Admission and After Admission. The combined contexts serve to depict the needs students have during their online learning experience to persist. Under Prior to Admission there are two categories: Student Characteristics and Student Skills. These categories were present in both Tinto’s (1987) and Bean and Metzner’s (1985) model. Characteristics of age, gender, ethnicity, intellectual development, and academic performance as well as skill preparation prior to college were included because they affect student persistence (Bean & Metzner, 1985). Previous academic performance and completion of online learning courses had significant relationship. First year students were less likely to be successful in online-based courses (Schlosser & Anderson, 1994).

After admission experiences contains two categories: external and internal factors. External factors refer to environmental variable from Bean and Metzner (1985), it includes finances, hours of employment, family responsibilities, and outside encouragement. For example, regardless of academic preparation, if students cannot pay for college, make adequate child care arrangements, or adjust their work schedules, they are unlikely to persist in school. Rovai incorporated the impact of external factors in shaping students’ perceptions, reactions, and commitments, this was lacking in Bean & Metzner (1985) and Tinto’s (1987) model. These types of factors are especially significant to the community college student and online student (Rovai, 2003). Online students are very likely to be nontraditional (Workman & Stenard, 1996). Nontraditional students are generally associated with living away from campus, belonging to social groups that are not associated with college, having dependents, not being involved in campus organizations, and attending college part-time. Nontraditional student manage their time among their classes, work, families, and roles in the
community. There is often less time for campus involvement outside the classroom (Graham & Gisi, 2000)

Internal factors in Rovai’s Composite Model were taken from Tinto (1987) and Bean and Metzner’s (1985) models. They were modified to fit nontraditional students and online learners who are taking courses online as opposed to the other models which both theorize from the perspective of college students who attended class on campus. Rovai proposed that the degree to which these factors are fulfilled by the online student will impact the persistence. The needs are; consistency and clarity of the online program, policies and procedures, self-esteem, identifying with the institutions and feeling a part of it, social integration with other students and faculty, and access to support services (Workman and Stenard, 1996). Figure 1 depicts these internal factors.

Rovai’s model has not been operationalized, each construct has not been defined in a way for other researchers to replicate the design. Rovai’s constructs are: (a) background characteristics (b) initial commitment (c) academic and social integration (d) subsequent goal and institutional commitments; and (e) withdrawal decisions. Although Rovai continues to refer to categorical constructs, such as academic integration and social integration, he does not operationalize what these mean and which demographics would be assigned to academic integration. While it may be useful for theorists to know that academic and social integration matter, that theoretical insight does not inform institutions and practitioners how to define and measure either academic or social integration in an uniform way which will allow them to compare their integration to other studies at similar schools. The retention literature that refers to Rovai’s model often operationalize the constructs very differently. What seems to be
similar studies are widely divergent. Furthermore, it does not inform practitioners how to increase retention.

Figure 1. Rovai’s (2002) composite persistence model.

The organization and content of this chapter touched on retention models as presented by Tinto (1975, 1987, 1993), Bean and Metzner (1985) and Rovai (2002). Rovai’s (2002) Composite Model consists of two well cited, but very limitedly tested models that focus traditional students at four year institutions. Rovai’s Composite Model may be useful in
understanding persistence for this unique demographic of non-traditional and online students, but it is in its infancy in terms of empirically tested studies.
CHAPTER II
LITERATURE REVIEW

Introduction

Higher education literature examines institutional level retention and proposes that academic and co-curricular activities bring about student engagement and retention overall. Far less is known about retention at the course level, especially in community college courses. Very little research targeting the retention of community college students has been reported. It is understood that the demographic characteristics and academic preparedness of community college students is significantly different from the four-year college students (Bailey, Jeong, & Cho, 2008; McIntosh & Rouse, 2009) therefore it is not sufficient to apply the results from retention of students at traditional four year schools to students at community colleges.

The Intersection of Retention, Community Colleges, and Online Courses

Figure 2 represents a model of the literature represented in this proposal. Circle one represents the theoretical frameworks of the proposal, Rovai’s (2002) Composite Persistence Model. Rovai’s (2002) Composite Persistence Model concentrates on the online learner and nontraditional learners. Few online retention studies have acknowledge or tested this model instead they continue to reference either Tinto’s (1975, 1985, 1993) Student Integration Models (Fike & Fike 2009; Liu & Liu, 2000; Pascarella and Terenzini 1983) which are based
on traditional learners at residential four year colleges or Bean and Metzner’s (1985) Conceptual Model of Nontraditional Student Retention (Aragon and Johnson, 2008; Stahl and Pavel, 1992) which does not include online learners. Rovai’s Composite Model encompasses both of these other models while modifying the model to fit nontraditional, online learners.

Circle two represents community college education literature, including the demographics of community college students and retention studies at community colleges. Community college students and online students possess unique characteristics and attrition risk factors compared to the traditional university students (CCSE, 2012; Mendez, & Malcom, 2009; Xu & Jaggars, 2010). These risk factors include child care, work schedules, financial stress, and less support from friends and family. Many of these risk factors have been found to impact student success and retention, but these are also many of the same factors that make online courses appealing to students (CCSE, 2012; Mendez, & Malcom, 2009; Xu & Jaggars, 2010).

Circle 3 represents online learning and student retention at the course level. This area of research is very new. It has been found that many students who take online courses have additional risk factors that are similar to the risk factors of students at community colleges; child care, work schedules, financial stress, and less support from friends and family (CCSE, 2012).

Circle 4 is the focus of this proposal; online learning and retention at community colleges. The overlap of circles in Figure 4 indicates a gap in the current literature. Very little research has been done to explain the retention of community college students in online courses. The few empirical studies that have compared online to in-seat retention of
community college students suggest that students are substantially less likely to persist or successfully complete an online course (Carpenter, Brown, & Hickman, 2004; Jaggars & Xu, 2010; Xu & Jaggars, 2010).

For the purposes of this literature review the following pages will explore community college education, demographics of community college students, and retention of students at community colleges. The chapter begins with a discussion of the origination of community colleges from themes in the literature, it will then present studies that have examined various student characteristics to determine factors that increase student retention, specifically at community colleges.
General Overview of Community Colleges

Community colleges have their foundation in the Morrill Acts of 1862 and 1890, which have provided greater public access to higher education. Federal lands were given to create public institutions of higher education that focused on agriculture and technical sciences. In 1901, America’s first public community college, Juliet Junior College, was founded. Juliet Junior College was designed to accommodate members of the community who wanted to pursue a college degree but desired to stay within their community. The curriculum paralleled the first two years of a four year college or university. Junior colleges became popular solutions to provide communities a way to secure their own higher education institution, enhance the prestige of the local community, and increase the opportunity for more people to attend college (Pederson, 2000). The term “community college” has been the standard terminology for public two year institutions for the last thirty years, but they have also been referred to as “junior colleges,” “city colleges,” and “people’s colleges” (Cohen & Brawer, 2003, p.4).

During the severe economic downturn of the late 2000s community college enrollment surged 17% (Mullin & Phillippe, 2009). This was a result of the large numbers of factory closures and layoffs, which caused displaced workers to go back to community colleges to pick up the skills needed to be reemployed. In addition, there was an increase in younger students who would have gone to a university in better economic times.

Community colleges encourage part-time attendance and have lower tuition than universities. Typically community college students possess different characteristics than the traditional university student. The ease of access, low tuition, and open-door policies
have contributed to community colleges having a higher percentage of non-traditional students than universities (Cohen & Brawer, 1996).

Demographic Characteristics of Community College Students

Community colleges’ admission policies invite a very diverse group of students with a diverse set of needs, places for second chances, or new opportunities. As an open door institution, community colleges welcome students with all types of educational backgrounds. This includes individuals holding GED credentials, recent graduates from high school, and students returning to school to earn a second degree. Students attend community colleges for workforce training, to earn professional certificates, for continuing education credits, or as stepping stone to earn a bachelor degree. Retention and graduation rates of community college students are associated with demographic variables. Gender, race, and socioeconomic status are known factors associated with college success and degree attainment (Aitken, 1982; Bailey & Morest, 2004; Crawford & Jervis, 2011).

Academic preparedness determines success as students enter college (Greene, Marti, & McClanney, 2008). Community colleges differ from four year schools because they have the increased burden of making sure that students are indeed college ready. If students are not prepared, community colleges provide training through developmental education programs. A study based on 83 community colleges surveyed by the Achieve the Dream project found that 59% of students enrolled in at least one developmental course over three years (Bailey, 2009). Half of all first time community college students are in need of developmental education in English, reading, or math. Approximately twice the number of community
college students enroll in developmental courses compared to students at four-year schools (Levin & Calagno, 2008).

Most community college students attend classes while also working, caring for dependents, and juggling personal, academic, and financial challenges (Goldrick-Rab, 2012). The 2011 Community College Institutional Survey (CCIS) reported that 67% of full time students and 78% of part-time students work at least part-time while taking classes, and 53% of full time students and 60% of part-time students also care for dependents (CCSE, 2012).

Due to the relatively low tuition, community colleges are seen as a pathway to postsecondary education for financially challenged and minority students (Mendoza, Mendez, & Malcom, 2009; Rajasekhara & Hirsch, 2000).

These unique characteristics of community college students can impact student success and retention. In fact, fewer than half of entering community college students attain their stated goal of earning a degree or certificate within six years (Bailey, 2009). A longitudinal study between 1983 and 2008 by ACT, Inc.’s Educational Research Division reports that student persistence between freshman and sophomore semesters at public community college ranged between 51%-54%. This is significantly less than the 68% persistence rate at four year public institutions. Students from ethnic minority backgrounds are more likely to enroll on a part-time basis and are more likely to drop out before earning their degree or certificate (Mendoza, Mendez, & Malcom, 2009).
Investigating Online Learning Retention Studies

The number of students enrolling in online courses continues to grow in the United States (Allen & Seaman, 2011). Nearly one third of the total students enrolled at degree granting postsecondary institutions took at least one online course in 2010. That is more than six million students. At the same time, the attrition rate for online courses is extremely high (Argon & Johnson, 2008; Grimes & Antworth, 1996). Many community college nontraditional students, enroll in online courses because of the flexibility it allows for students with dependents or those who are employed. Enrollment in distance education courses have increased at institutions of all types but particularly at two year public community colleges (Newell, 2007). Distance learning often requires independent learning skills, study discipline, time management, and a high degree of motivation. These characteristics are not plentiful among developmental students at community colleges (Boylan, 2002).

It is important for the designers and educators of online courses to understand the characteristics of distance learners. These demographics and factors impact student success. Educators need to be aware of student demographics in order to develop high quality programs (Yukselturk & Bulut, 2007). The issue of attrition in online courses is important and remains a high priority in the United States (Nash, 2010).

At higher learning institutions, student retention at the course level is important in assessing the effectiveness of the cost of online learning compared to traditional classroom based teaching. Studies have shown that students who take online courses early in their career may be less likely to retain in school (Jaggars & Xu, 2010). For example, Jaggars (2011) reported that students who took one or more online courses their first fall semester
were significantly less likely to return the spring semester. Jaggars’ (2011) study showed retention rates 5% lower than those of students who took a campus-based curriculum, a non-significant, but measurable difference (69% vs. 74%). This type of information is important to institutions in order to guide students to sign up for courses that will assist them in retaining in school and meeting their desired goals.

The existing literature on course completion rates among distance learners focuses on studies trying to identify possible factors or demographics that influence retention in online courses. These findings have been broken down into categories of common factors (gender, age, ethnicity, enrollment status, and tuition fee waiver) in short summaries that highlight the primary conclusions from these articles, and how these conclusions relate to the persistence model.

Online Student Success and Gender

In a study by Aragon and Johnson (2008), females demonstrated a positive correlation to persistence in contrast to the male students. Conversely, Yukselturk and Bulut (2007) and Wojciechoski and Palmer (2006) found gender as a variable was unrelated to learning outcomes in online courses. Other factors may impact course persistence and gender. For example, Conger and Long (2010) found that male students arrive at college with lower high school grades than female students, and suggest this may explain some of the disparity in regards to gender and persistence. Females are more likely to apply for, receive, and respond to tuition and support which lowers the cost of school and may increase their
persistence and probability of graduation (Angrist, Lang, and Oreopoulos, 2006; Dynarski, 2007).

Online Student Success and Age

Wojciechowski and Palmer (2005) investigated student characteristics related to academic success in an online business course. Selected demographic and student characteristics were examined. One hundred and seventy-nine students participated in the study, with an average age of twenty-five. Students were considered successful if they received a “C” or better. Successful students were older and had taken online courses previously (Wojcichowski & Palmer, 2005). Aargon and Johnson’s (2008) study evaluated demographic characteristics of students enrolled in online courses at a rural community college in the Midwestern United States. They found that age was not a significant indicator of positive student persistence.

Online Student Success and Race/Ethnicity

In addition to the demographic characteristics of age and gender, Bowen, Chingos, and McPherson (2009) argued the importance of including race and ethnicity as a variable in any meaningful analysis of student persistence or attainment in distance education. They articulated the necessity to reduce the gross disparities in graduation rates that exist today among groups classified by race and socioeconomic status. According to the report by the National Center for Education Statistics (2010), institutions of higher learning experienced an increase in enrollment among various ethnic groups from 16.1% in 1980 to 30.9 in 2007.
Much of this change was attributed to rising proportions of Asian and Hispanic students. The proportion of students who are black was 13.9% in 2008, only an increase of 4.7% from 1980. Hispanic student enrollment increased by 9.2% during the same time period (National Center for Education Statistics, p.13).

The increase in enrollment among minority groups makes ethnicity an important variable to consider when investigating academic performance in online education, yet few studies have been conducted which examine the relationship between ethnicity and academic performance. Yukselturk (2009) found that white students successfully completed online course at higher rates than black students. These finding are supported in literature that found minorities were less likely to persist in online courses or programs (Carter, 2006; Jaggars, 2011). Conversely, Aragon and Johnson (2008) reported ethnicity was not significant to dropout or persistence in the online classroom.

Online Student Success and Academic Readiness

Academic readiness, as measured by placement scores in developmental reading, writing, and mathematics, was not significantly associated with the completion of online courses (Aragon & Johnson, 2008; Rovai, 2002). Student GPA was found to have the highest correlation to the final grade received in an online class (Aragon & Johnson, 2008; Wojciechowski & Palmer, 2006). This is consistent with the research conducted by Tidewater Community College (2001) which reported that students with a GPA under 2.00 were less likely to complete their online courses.
Online Student Success and Enrollment Status

Research indicates a high correlation between full-time enrollment and students achieving their educational objectives (Aragon & Johnson, 2008; Fike & Fike, 2008). A report released by the National Center for Education Statistics states that “Students who attend college part time are at a disadvantage when compared to their fulltime peers” (Walsey, 2007, p. 25). Unfortunately, many community college students are not in a position to enroll full time, because of their need to care for dependents or because of full-time employment (California Community Colleges, 2012). Nakajima (2008) reported that within a sample of 427 community college students, the number of credit hours students were enrolled in was a strong predicting factor for students to persist in their education. Colorado and Eberle (2010) also reported that community college students enrolled full-time with at least one online courses performed slightly higher than students enrolled part-time, though this was not a significant difference. Wojciechowski and Palmer (2005) investigated the relationship between enrollment statuses of 179 undergraduate online students: 144 of the students were enrolled part time and 46 students were enrolled full time. The results indicated that there was not a statistically significant relationship between enrollment status and success.

Academically Underprepared Students and Online Education

Over 50% of the students enrolled at community colleges take one or more developmental education course (Bailey, 2009). Currently in many schools there are no guidelines regarding who can and cannot sign up for online classes. However, many
institutions harbor specific concerns about online course performance among underprepared or traditionally underserved populations, who are already at risk for course withdrawal and failure (Jaggars & Bailey, 2010). Studies have reported that online learning may undercut academic progression among low-income and academically underprepared students (Jaggars & Bailey, 2010). Online course persistence studies suggest that online courses are often associated with less desirable course outcomes for underprepared students (Carpenter, Brown, & Hickman, 2004; Fike & Fike). Given the rapid growth of online courses in community colleges, it is important to verify that these courses do no harm to students’ academic success.

The results are really inconsistent but general trends suggest academic preparedness, factors appear to contribute to both success and retention of students in online courses. Research indicates a high correlation between full-time enrollment and students. Inconsistency regarding retention versus success data in regards to academic readiness, as measured by placement scores in developmental reading, writing, and mathematic was also found. The placement tests appear to be able to correlate to student final grade but are not consistent in predicting retention. Student GPA was found to have the highest correlation to the final grade received in an online class (Aragon & Johnson, 2008; Wojciechowski & Palmer, 2006).

This section focuses on existing literature looking at course completion rates among online learners using different retention models; Tinto’s (1993) SI model, Bean and Metzner’s (1985) Conceptual Model of Nontraditional Student Retention or Rovai’s (2002) Composite Persistence Model. In particular, the aim of these studies was to identify possible factors influencing retention or success in online courses. After reviewing each study
separately, a summary will highlight the primary conclusions from these articles, how these conclusions relate to the framework of retention of online science learners at community colleges, and identify questions that emerge from this literature.

One simple formula to ensure student success and attrition has not been discovered. Adult persistence is a complicated problem in colleges; Non-traditional students, traditional students, and online learners are all unique. It is not credible to attribute student attrition to any single characteristic, demographic, or situation but the more informed colleges and students are, the better their chance of having the tools needed for them to be successful in the learning environment they self-select into.

Liu and Liu

Liu and Liu’s (2000) *Institutional Integration: An analysis of Tinto’s Theory* examined the impact of social and academic integration on college students’ satisfaction and retention. Freshman (n=378) from a comprehensive state university in the Midwest responded to a survey that included questions on the variables of persistence, satisfaction, academic integration, social integration, academic performance, and demographics. These variables align with Tinto’s model. The surveys were matched with students’ academic records and transcripts to determine academic status.

The survey was originally sent to 1606 students, with 378 responses received back and evaluated. The authors expound that the characteristics of the responding students were compared to non-responding students and that they were similar with regard to age, ethnicity, high school performance and ACT scores. Details regarding how this was done, or of the
actual results were not shared, but the authors believe their data set was an accurate representation of the overall freshman class.

Persistence in this study was defined as ‘continued enrollment in the university after the semester in which data was collected.’ Student satisfaction was defined as ‘an evaluation of overall experience at the institution that is measured on a Likert scale of a seven.’ No further information regarding what was considered satisfied or not satisfied was disclosed. The overall experience was not further broken down into categories to show specific areas that the student was satisfied with and areas they were not.

The survey questionnaire of academic integration was included. It consisted of fifteen questions and the student responses were measured on a seven point Likert scale. It was not indicated if this survey had been previously used, validated, or where it had originated. The survey focused on the relationship between the student and the advisor as well as the extent to which the student felt the advisor was successful in this role. The other two survey questions asked if the course was ‘fair and appropriate’ and the extent to which the course ‘prepared you for more advanced courses.’

Social integration was also assessed in this study using a seven point Likert Scale. Students were asked to rate eight questions about how well students believed they fit into the school environment. The authors stated that the score was assessed by summing up the respondents’ ratings. Again, there was no indication of whether the statements in the survey were validated or appropriate for testing social integration. Two of the questions focused on the university staff being helpful, and sensitive to individual needs. Five questions asked if athletics, extra-curricular events, or cultural events were a valuable part of their experience.
The last question focused on whether the residence hall was an adequate environment for studying, but nothing about social integration at the residence hall was addressed.

Academic performance was indicated by grade point average which was collected from the students’ transcripts. Details regarding what grades were considered successful was not shared. Demographic variables in this study included gender, ethnicity, and age. These were chosen based on the demographics Tinto used. The academic integration evaluation aligned with Tinto’s SI model for the reason that it included both academic performance and faculty/staff interactions. The authors used regression analysis to analyze the impact of gender, academic performance, academic integration, social integration, ethnicity, and age upon the dependent variable of satisfaction. Next, a regression equation was used to evaluate the impact of those variables on retention. The results indicated that gender, age, and ethnicity had no impact on student retention. Academic integration, academic performance and student satisfaction influenced student retention and supported Tinto’s theory.

One limitation the study had was that it did not discuss where the survey instrument was attained and if it had been previously used, published, or validated. Again, no limitations were discussed but the instrument used to assess the students’ social integration seems to be a possible limitation of this study. Liu and Liu did not model their questions of social integration after Pascarella and Terenzini’s (1980, 1983) research even though they cite them as researchers who have extensively studied how to measure social integration and academic integration including student peer interactions, faculty interaction with students, goal commitment, faculty concerns for teaching, and student intellectual development. It is not evident what Lui and Lui’s questions were modeled from. Within the paper Lui and Lui defined social integration as ‘the ease of meeting friends and developing personal
relationships’ although the statements within the Likert scale questionnaire did not directly assess personal relationships or friendships. For example, one question had to do with living in a residence hall. If the student did not live in a residence hall that would have impacted the response, without prompting additional questions about social integration where he/she did reside. Another question focused on whether the student had the opportunity to attend cultural events such as plays, concerts and museums. It is possible that the students successfully integrated socially, but not as a result of the listed cultural events. Lastly, many of the social integration questions focused on the advisor/student relationship. This was not a focal point of Tinto’s theory, it is not clear why it is integrated so heavily into this study’s social integration questionnaire.

The results of this study conclude that social integration has no influence on the outcome of withdrawal. This study showed moderate support of Tinto’s theoretical framework. It should be stated that the social integration questions did not focus on any integration in more depth than if events were available or the students’ perception of staff and programs at the university.

The authors briefly state in the discussion section that social integration has no influence upon the outcome of withdrawal and commented ‘especially in community college settings.’ This comment was out of place because up to this point they had made no reference to community college students and no reference to community college studies. This study was completed at a state university.

Lui and Lui caution that higher education needs to assume the responsibility of installing a meaningful retention program to prevent students from dropping out. They coin the term ‘theory of alienation’ and recommend more practical research to be completed
around the topic of why students voluntarily withdraw without seeking help because of shame or embarrassment of ignorance.

A methodological shortcoming in this study is that such a small amount of students agreed to participate. In a study about retention it is particularly important to receive feedback from the students who dropped out. If that is not present there is a risk of self-selection distorting the data; those who were embarrassed about dropping the class will not give input.

The instrument used in this study did not adequately match the research cited definition of social integration by Tinto. Tinto simply described social integration within an institution as absence of isolation (1987). Students in this study were asked to rate eight questions on a 7 point Likert scale. The questions were specific to individual situations. Two of the questions focused on the university staff being helpful, and sensitive to individual needs. It would be possible for a student to rate the staff as being unhelpful and not sensitive but still being very socially engaged in other areas of school. Five questions asked if athletics, extra-curricular events, or cultural events were valuable part of their experience. Note that the questions did not ask if they attended these events, but instead if they felt it was ‘valuable.’ The last question focused on whether the residence hall was an adequate environment for studying. First, all students may have a different idea of what is ‘adequate environment for studying.’ Second, the residence hall being an adequate area to study does not directly correlate to Tinto’s definition of social integration. No questions focused about meeting other students and engaging in activities, which is what Tinto had stated social integration was, lack of isolation.
The instrument used to assess academic integration consisted of fifteen questions and student responses that were measured on a Likert scale. There was no indication of where this survey had come from, if it had been previously used, or if it was validated. The questions did not seem to focus on all areas of academic integration that Tinto had discussed. For example, thirteen of the fifteen academic integration statements on the Likert scale, started with the words “My advisor.” The instrument is not a good representation of Tinto’s definition of academic integration. Tinto’s (1993) definition of academic integration included directly influencing the students’ goal commitment, academic performance or faculty interactions.

In conclusion, the study partially validated Tinto’s theory. Academic integration, social integration, and academic performance all had positive influences on student satisfaction. Yet, social integration was not significant in the students’ decision to retain. A problem with this study was that it ignored Tinto’s suggestion of potentially important compensatory interactions among constructs. Tinto, for example discussed that in terms of influence on persistence, a high level of academic integration would compensate for a low level of social integration. Lui and Lui failed to mention the importance of these interactions.

The variables of academic integration, academic performance, and satisfaction were verified in both the Lui and Lui (2000) study, and Pascarella and Terenzini’s (1983) study that was discussed earlier. The only substantial difference between the findings in these studies lies in the impact of social integrations upon retention. Lui and Lui (2000) reported that social integration had no effect upon academic performance and persistence. They reported the correlation between academic and social integration was almost nil (r=.059). Conversely, Pascarella and Terenzini (1983) reported social integration had important direct
effects on persistence as well as indirect effects through their influence on the construct of institutional and goal commitment.

Pascarella and Terenzini (1983) recommended that Tinto should not break integration down into social and academic factors. Instead, integration should be viewed as a whole. They describe lack of integration as alienation. Lui and Lui (2000) pronounce lack of either social or academic integration as a structural problem, not an individual problem. It is possible that the experimental design and instrument used to collect data on social integration impacted the results. It would be interesting to see the results if Lui and Lui (2000) incorporated the operationalized constructs of social integration and questions from Pascarella and Terenzini’s (1983) study into their study. There is a strong need for consistency in design, methods, and instruments in order to gain insight into retention in success in higher education.

While Tinto’s (1986) written theory recognizes and discusses the relationship between consistent components of social, academic integration (personal/normative integration), and dropout behavior that are unrelated to institutional and goal commitments, his diagram (Figure 1) does not illustrate a direct path linking academic integration or social integration to dropout behavior. The written theory, however, does refer to experiences in the formal and informal systems that may lead to voluntary withdrawal. These discrepancies between his written theory and his diagram cause confusion and allow for the model to be interpreted differently leading to a diversity of applications.

Within Tinto’s explanation of his theoretical model, he states that although the model is diagrammed in a way that suggests path analysis, the diagram is in fact not meant to be a path model. It is interesting that he does not imply that path analysis is an appropriate method
to study dropout yet Pascarella and Terenzini (1983) used path analysis to successfully identified 80% of the persisters.

Tinto’s theory has been widely used across all areas of higher education, even though it does not precisely fit, nor was meant to fit, all types of institutions. Tinto (1993) himself explained that his model should be evaluated differently in two year schools, stating that ‘experiences in the academic settings of classrooms and laboratories of a two-year campus are much more important to persistence than they are in residential settings.’ This is because student time on campus is much more limited to class time than it is for students in a residential setting. Yet, retention studies from four-year community colleges are published with little or no comment on the applicability to retention at a commuter, two-year school. Possibly the researchers who are using his theory need to be cognizant not to generalize Tinto’s theoretical model to all student types in all forms of education. The model was not meant to predict all types of explanatory terms. Tinto states it was designed to “highlight the clearest explanatory terms specific types of relationships between individuals and institutions that may account for sufficient types of drop out behaviors.”

It is interesting to note that Tinto, himself, questioned the possibility of reducing the national rate of student attrition. In 1987 he stated that “we should be much more conservative in our projections referring our ability to significantly reduce dropout in higher education at a national level.” He further describes instead the chance of reducing attrition rates in certain subgroups of the population. Tinto (1987) has a striking fact illustrated in a graph: that for the last 100 years the dropout rate for universities in the USA has been constant at 45%, despite big increases in the participation rate and amount of public funding. Dropout rate was here defined as the ratio of undergraduate degrees awarded to those student
with first-time enrollment four years before. The second world war causes the only big wobble in the flat graph, and yet averaged across 10 years even there the rate is near-constant (because positive and negative blips cancelled out). Therefore, if global retention rates remain constant, it must mean that a reduction of attrition in one subgroup leads to an increase of attrition in a different subgroup.

However, Tinto did not indicate how these levels interact or provide indicators that would increase student persistence. Tinto’s model is based on experiences at four year colleges where students reside. The question of how commuter college students, non-traditional students, and community college students fit into this model needs to be examined. Tinto’s retention model does not address if and how the level of social or academic integration may vary depending on the type of school they were attending.

Tinto’s model is an overall picture of theoretical retention. Researchers have used techniques such as path analysis, and structural equation model testing to try to address this to some extent, but possibly not sufficiently. Essentially the question is, does Tinto’s model have potential for improving practice? There are so many ways in which a student's academic or social integration might be low: students drop out for diverse reasons, and having a general description does not tell you how to do something effective for each student. However in principle, it seems we could develop a detailed investigative instrument e.g. using questions to determine what the particularly bad issues are in each situation (each department of each institution); and then select an intervention specific to that diagnosis. As of now the education community is a long way from demonstrating this, though.
Stahl and Pavel

Stahl and Pavel’s (1992) study was included in the literature review because it was designed with the purpose of determining how well Bean and Metzner’s Nontraditional Undergraduate Student Attrition Model (1985) fit the community college. The research question was: To what extent does the Bean and Metzner model, as operationalized in the present study, account for the attrition of community college students? If the model does not "fit" the data, how can the model be modified to better account for persistence? The authors conclude that there is a deficiency in the number of studies validating the appropriateness of the Bean and Metzner model.

Community college students in an urban, multi-college district which enrolls about 10,000 students composed the sample for this study. Students enrolled in developmental or beginning reading, English, or math courses were surveyed (N=1519). The survey was administered during the 14th week of the semester. Information on grade point averages and subsequent re-enrollment were obtained through the college student information system after the fall 1990 semester began. The survey was completed by 665 students.

The data was analyzed using Structural Equation Modeling, a quantitative technique that estimates the covariance among the observed variables. Covariance was used to assess the hypothesized relationships among the constructs proposed by a theoretical model. They set up the structural equation model to fit Bean and Metzner’s theoretical framework. The sample size was acceptable since it was larger than the number of structural coefficients to be estimated.
Data was analyzed using two formal mathematical models and computer programs; LISREL (Jöreskog & Stirbom, 1989) and LISCOMP (Muthén, 1988). Both were used in order to take advantage of the strengths and overcome the weaknesses of each. LISREL is more widely used than LISCOMP, but LISCOMP may be better suited for student persistence research when using categorical data (T. Koehler, personal communication, Oct. 2014). Exploratory factor analysis was employed to suggest revisions to the model to improve its "goodness-of-fit" for community college students. This design was well suited for student persistence research when using categorical data.

Bean and Metzner’s model was not accepted as explaining retention at the community college in this study. The results from the Goodness-of-Fit indicator and Adjusted Goodness-of-Fit Indicator did not fall within the acceptable fit ranges. The critical matrices were not ‘positive definite.’ High school performance was negatively associated with retention in this study. In previous studies high school performance had been proposed as one of the best predictors of college persistence for traditional students in residential colleges (Astin, 1975; Bean & Metzner, 1978; Tinto; 1986). This study also reported the relationship of academic achievement to retention was negative (-1.199). Again, this data is contrary to the results from similar studies (Astin, 1975; Bean & Metzner, 1978; Tinto; 1986). It is important to note that in this sample more than half of the students stated they were in the bottom half of their high school class. All the students who participated in this study were all enrolled in developmental courses at the community college. The study contained a large number of younger students who were unmarried without children. This seems to be an indicator that the sample may not be heterogeneous enough to be a true test of the Bean and Metzner model. Bean and Metzner’s model define non-traditional students as “older than 24, does not
live in a campus residence, or is a part-time student, or some combination of these three factors; is not greatly influenced by the social environment of the institution” (p. 489). The population in this study did not fit Bean and Metzner’s definition.

This study was plagued with limitations. First, the population for this study was limited, only one subgroup from one institution. Less than half of the students that received the survey completed it. All of the students who completed the survey were enrolled in developmental or beginning reading, English, or math courses. Second, the survey results used were only from white students because not enough minority students completed surveys. Thirdly, the survey was administered the 14th week of the semester. Many of the students who were at risk may not be present in class this late in the semester. The survey methodology resulted in significant truncation of variance within the sample, and therefore the findings are of minimal value in assessing Bean and Metzner’s model. Lastly, the authors did not explain how they determined the reliability and credibility of the survey. It is difficult to make general conclusions about the overall application of a model when the selected group of students are all in remedial courses, and have academic deficiencies.

Stahl and Pavel’s (1992) conclusion proposed a new model with suggested modifications that were developed through structural equation modeling and by LISEREL analysis. The new model, Community College Retention Model, constructs included: academic variables (study habits and absenteeism), environmental variables (finances, employment, and family responsibilities), grade point average, psychological outcomes (satisfaction and stress), intent to leave (plan to re-enroll and transfer difficulty), and dropout. This new, proposed, model omitted the background and defining variables of age, gender, and ethnicity found in the Bean and Metzner Model.
Stahl and Pavel concluded that in order to effectively use a model to study retention patterns for student subgroups, it is best not to co-mingle the effects of the demographic variables with the other variables in the model. Residence was omitted as it was not necessary because students at the community college are commuter students. The Bean and Metzner variable of GPA was expanded to Cumulative GPA and Semester GPA. In all they came up with 22 measurement variables, not all significant and not all indicated reasonable influence. The authors recommended more studies need to be done to select variables that account for most variance.

In conclusion, Stahl & Pavel reported that the Bean and Metzner model was “an extremely weak fitting model” (1992, p. 19) based upon their analysis of data using structural equation modeling. The authors continued, as did Bean and Metzner from Tinto’s model, to propose their own model called the Community College Retention Model. Interestingly enough, their newly proposed model retained all the original variables introduced by Bean and Metzner except gender, race/ethnicity, and age. Justification for not including gender, race/ethnicity, and age was based on the fact that these attributes cannot be transformed by interaction with the college environment. Bean and Metzner’s model was an adaption of Tinto’s theoretical model for use with nontraditional students. The age variable should not be omitted as it is important in order to define student population as ‘nontraditional.’

Stahl & Pavel’s (1992) factor analysis resulted in a few differences in the association of variables, but overall the two models are quite similar. The number of students completing their survey was very low (whites=597; minorities=68).

Their dismissal of Bean and Metzner’s theory followed by a proposal of a very similar, reorganized theory seems to be misguided. Their own model was not tested, only
proposed as a better fit. Certainly additional research on community college students, with more heterogeneous samples than this study, must be conducted before any assessment of the utility of Bean and Metzner’s model can be made. Again, it is a struggle to test these models in appropriate ways, too many undefined constructs are being used. Constructs such as study habits, family responsibilities, and psychological outcomes are not defined in a way that allows for uniformity. A model that has all defining variables such as age, gender, ethnicity, GPA, enrollment status would allow for a common baseline for all studies to be able to equate their findings to other studies around the country.

Fike and Fike

Fike and Fike (2009) completed a quantitative, retrospective study to assess the predictors of student retention for first time in college (FTIC) students who are attending a community college. The study comprised of four years of data from 9,200 FTIC students attending a Texas, public, urban community college with an academic student population of approximately 10,000 students.

Fike & Fike (2009) based their study on the theoretical principles of Tinto (1993) in an attempt to describe ways in which the students and institutions interact with one another. The choice to use Tinto’s model is not explained. Tinto’s work is best used to explain student retention research on traditional-age students within the residential settings of the universities. Conversely, this study focused specifically on the community college student, whom possess different characteristics than the traditional university students.
This study stated that it aimed to provide greater insight into variables that differentiate community college students from university students and the possible factors that influence student retention. This information could assist institutions in retention by providing interventions with students who bring particular characteristics to the campus. The sample comprised of four years of data: 2001, 2002, 2003, and 2004 fall semesters. Student retention was operationally defined as remaining enrolled at the institution from the first-year fall semester to second year fall semester.

Student data was analyzed using SPSS. Descriptive statistics including means, standard deviations, and percentages were developed for the sample. Chi-square analysis was used to assess the distribution of student retention rates by academic year. Bivariate correlation coefficients were calculated to determine the association of student retention with predictor variables. Multivariate logistic regression models were also prepared to predict the odds of student retention while controlling for relevant confounders such as student demographic variables (Gelman & Hill, 2007). For all analyses the level of significance was 0.05.

Age was a variable, because in general community college enrolls more adult and returning students than universities. Ethnicity was another since community colleges are a primary entry point for minorities. A high proportion of community college students are not college ready therefore the number of students enrolled in developmental education was incorporated as a variable. The number of credits students were enrolled in was included because nearly two thirds of community college students attend part time. Tinto’s model refers to socio-economic status for this reason financial aid was selected as a variable. First generation college students were also looked at because they are more concentrated in
community colleges. The number of students who enrolled in an online course was included as an additional factor. The growth in online courses has led to this variable as one that needs to be assessed.

Descriptive statistics were provided in the study. Students were predominantly female (56%) and White (66%). Median age was 19, which is atypical of many community college studies and may be an indicator of a unique community college population sample. The median hours of credits were 12 for the first fall semester. About 22% of the students were enrolled in developmental reading and 66% of the students were enrolled in developmental mathematics.

Interestingly, taking online courses was shown to be a predictor of student persistence in this study. The authors suggest it was the flexibility provided by online courses. These findings provide a basis for institutions to offer flexible schedules, including quality online courses to meet the needs of their students. Research to further explore the association of online courses with student retention is needed. This result is contrary to other studies which have reported that online courses had increased students’ attrition rates, especially when taken the first year of community college (Jaggars & Xu, 2010). This inconsistency with prior research may be a consequence of the type or number of online courses that were offered at this institution. Further research is warranted in the area of retention and persistence at community colleges in online courses.

The regression models of this study indicate that the strongest predictor for retention was passing a developmental reading course. Students who scored at college level in reading on a placement test such as Accuplacer already possess college level reading skills. Therefore collectively these findings indicate the significant impact of reading skills on student success
and retention which was consistent with prior research (Dixon, 1993; Fleishauer, 1996).

Student Support Services is a federally funded TRiO program that is charged with increasing the rate of retention for students from disadvantaged backgrounds, including students from first-generation and low income families and those with physical or learning disabilities. Within this study a small number of students, (397, 3.9% of the student population), participated in Student Support Services, but the findings indicate that TRiO encourages student retention. Involvement in TRiO required students to meet regularly with their advisers, complete mid-semester grade checks, and complete a long-term plan of study. These requirements align with Tinto’s external factors within the college. The authors recommended that further research is done to strengthen the external validity of this finding.

A quarter of the students reported that their parents had some college-level education. It is important to note that this was an area that was self-reported, not verified, and was missing a large percentage of the data. The results reported parental education as positively associated with student persistence. This is consistent with national representative samples of first generation postsecondary students who have lower persistence rates than their counterparts (Nunez & Cuccaro-Alamin). Prior findings have also concluded that parental education is positively associated with student persistence (Hoyt, 1999; Summers, 2003; Wild & Ebbers, 2002).

The biggest limitation of this study is that the study did not employ an experimental design; a causal relationship between predictor variables and student retention could not be defined. Another limitation is that Fike and Fike did not report the actual questionnaire items used to operationalize the theory. By never discussing the design of questionnaire items, major theoretical issues are ignored. For instance does "social integration" mean integration
within that institution, or generally? Probably Tinto meant the former. Yet a student with no friends anywhere, and a student with plenty of friends who however are not enrolled at the same college are likely to show different tendencies to dropout. Scrutiny and discussion of individual questionnaire items is a good way to identify theoretical issues, and conversely avoiding such discussion also makes it likely that no two studies are measuring the same thing, yet are unable to determine this. Replication of this study with similar findings for other populations would strengthen the external validly of this study. Student data can be used to develop an understandable and workable plan to guide efforts toward effective interventions and increase student persistence.

Although the authors assert that community college students are unique, and have a different set of needs than the traditional university students, their study is still based on Tinto’s SI model. The lack of rationale of why Fike and Fike did not choose Bean and Metzner’s (1995) model of student dropout which focusses on non-traditional community college students, was not discussed and is considered a limitation to the study.

Findings support that passing developmental reading courses, taking Internet courses, participating in the Student Support Serves program, receiving financial aid, the number of hours for which the students enroll in the first fall semester, and the number of hours dropped in the first fall semester served as predictors of student persistence.

We could ask, and perhaps even find the answer to the questions: Which factor most determines student's success? (E.g. demographic such as social class, learning styles, or gender). However one reason Tinto's approach may be better than other ways of talking about this area is that it does not align with the simplistic question of whether the student or the university should be "blamed.” The metaphor of integration is about fit; it is not about
adapting, but about whether the institution or course and student go together well. Even more than that, integration is clearly the current outcome of a relationship of successive contacts which progressively modify that relationship, hopefully for the better. As a student has more successful interactions with a tutor, for example, they are more likely to feel more integrated with positive experiences, which will in turn teach them to ask for help in future, and to ask for it in a way that gets results from that individual tutor.

It is interesting that Tinto’s model is referenced to explain retention in community colleges yet students at community colleges commute and do not necessarily experience separation as described by “leaving home.” The use of Tinto in a community college retention study should raise questions regarding whether or not this model accurately describes the unique situations of community college learners. Surprisingly, Tinto’s models are continually the top cited models in retention studies at community colleges across the country (Lui & Lui, 2000; Wild & Ebbers, 2003; Wright, 2013; Xu & Jaggars, 2012).

Aragon and Johnson

Aragon and Johnson’s (2008) study aimed to provide further information about adult learners in online education in comparison with adult learners in other types of distance education. Community colleges are turning increasingly to online learning yet their online dropout rates are reported as 20% higher than in face-to-face classrooms (Park & Choi, 2009). Aragon and Johnson aimed to provide further information about students who enroll in online courses. This type of information can be used to guide institutions in developing and maintaining appropriate and adequate student support services (Deupin-Byrant, 2004). Community colleges often lack knowledge of who is enrolling in online courses and of the
differences between demographics, hours enrolled, and academic and self-directed learning characteristics of students who complete online courses (completers) and those that do not complete (non-completers) online courses. Their study also obtained students’ self-reported reasons for not completing the course.

Participants in this Aragon and Johnson (2008) study consisted of 305 students from a rural community college located in the Midwestern United States. Of the 305 students who participated in the study, 216 were female (71%) and only 89 (29%) were males. Additionally, the group consisted of a majority of white students, 248 (81%), the remaining 57 (19%) students were categorized as non-white students. The non-white students were not broken down further into ethnicity. The mean hours of enrollment as calculated on the 10th day of the study was 9.34 (SD= 5.30). Student GPAs ranged from 0 to 4.00, with the mean GPA being 2.16 (SD= 1.58) on a 4.0 scale. The study did not disclose what type of online courses these students were enrolled in.

This study used a combination of comparative and survey designs. A comparative research design was used to address research question one: “Is there a significant difference in demographic characteristics, enrollment (hours enrolled) characteristics, academic readiness, and self-directed learning readiness between students who complete and do not complete online courses?”

Each student’s demographics, enrollment, academic readiness, and course completion variables were all retrieved from the college’s Management Information System. Variables were queried and downloaded into an Excel database that was then imported into the SPSS for analysis. Students all took a self-directed learning readiness test, BISL, as part of their
tutorial program prior to the start of their online course to document their perceived self-readiness.

Chi-square was used to address the research question. When a significant difference was found as a result of the chi-square and t-test then a correlation coefficient was used to determine the degree of covariation between student characteristics and completion vs. non-completion in their online course. The Pearson product-moment correlation coefficient provided a numerical summary of the strength and direction of the linear relationship of the selected student demographics, enrollment, academic, and self-directed learning independent variables.

Research question two focused on the self-reported reasons for non-completion of online courses and was answered by survey design. Self-reported reasons were given by the students via phone interviews as to why they chose not to complete their online course. Content analysis was conducted to analyze the data collected. It should be noted as a study limitation that the authors were only able to collect 65 responses from the 116 non-completers. It is especially important to receive feedback from the students who dropped out to make sure the study is not biased by self-selected samples. This type of approach tried to both operationalize and establish parts of the overall model, piece by piece. Responses included the following reasons for course attrition: personal reasons (34%), course design and communication (28%), technology issues (18%), institutional mistakes (11%), and didn’t fit their learning style (9%). Although interviews could be thought of as a more informative data source, it should be noted that student will often create a rationalization, particularly dropouts, about why they were not successful. Students will be more likely to provide causes such as external factors to explain their attrition. It might be beneficial to develop a more
objective assessment of coding or possibly a panel of people to assess reasons for dropout; instead of self-report, to try to control for bias in the dropout’s own opinion on why they dropped out.

To the extent that this research is exploratory, results of this study provide insight by reporting that there were significant differences in gender and academic readiness, and completers enrolled in more online course than non-completers. It is key to point out that no significant differences were found in the characteristics of age, gender, ethnicity, financial aid eligibility, grade point average, and total hours enrolled. There was a significant difference between the number of hours in which completers and non-completers enrolled. Completers enrolled in more online hours than non-completers. The mean online hours of completers was 4.32 and non-completers hours mean was 1.48. The students completing their online course had a mean GPA of 2.47 while students who did not complete the course had an average 1.66 GPA. This is consistent with research conducted at Tidewater Community College (2001) and also with the studies by Fike and Fike (1999) and Rovai (2002).

However, several limitations must be considered in interpreting the study findings including the small number (N= 305) of students who participated in the study, and the uneven distribution of sex (71% female, 29 % male). Additionally, the group consisted of a majority of white students, 248, (81%) and the remaining 57 (19%) students were categorized as non-white students. The non-white student were not broken down further into ethnicity. The authors defined completion in this study as students who received a grade of an A, B, C, or D. Course non-completion was defined by a grade of an F, or those who dropped out.
In summary, it is evident that the current study is still far from being conclusive. Despite some deficiencies to the extent of size and heterogeneous sample, this study has provided some insight into differences between completers and non-completers of online courses. This study reported that students with lower GPAs are at greater risk of dropping their online courses. No significant association was reported between student placement in developmental reading, writing, or math.

Wojciechowski and Palmer

Wojciechowski and Palmer (2005) have similar interests as Aragon and Johnson (2008) in understanding student characteristics that can be predictors of success in online classes. In this study, the participants were all the students (N=179) who had taken an online business course at a small, rural, community college, in western Michigan between the fall 2000 semester through summer semester 2003. The same instructor using the same textbook taught all the courses providing instructional consistency.

In this study, Wojciechowski and Palmer (2005) do not articulate specific research questions. Instead, they explain their research purpose in more general terms as providing further insight into characteristics of successful (earning a C or better in the course) online learners to determine if commonalities could be identified. Previous research studies have looked at characteristics of learners but have taken the enormous variability of the online learners’ characteristics and disguised it by gathering samples of students and merging them into averages, which produces an erroneous typical learner (Carnevale, 2002; Colorado& Eberle, 2010). It is well documented that students who take online courses may appear
typical but there is a great degree of diversity within the student population (Cheung & Kan, 2002). Research on online education has reported that there is limited information on what accounts for individual student differences when taking online courses.

Wojciechowski and Palmer (2005) examined this diversity by using Pearson product-moment correlations to look at various characteristics and determine if there was a correlation between these and their grades in the online course. Various student characteristics were examined to determine their relationship to student success, which in this study was based on overall course grade. The variables were selected based on previous research findings that reported a possible connection to student success in online courses and availability to the data. The variables that were analyzed include: gender, age, previous online course completion, ACT English Scores, ACT Reading Scores, ACT composite scores, ASSET Reading Scores, ASSET Writing Scores, grade point average, previous withdrawal from other courses, semester format (16-week or 8-week), student status (full-time vs. part-time), and attendance at orientation session.

Descriptive and inferential statistics were used to analyze the data for individuals within two groups: (a) all students, and (b) those students receiving a C or better in the class (i.e., defined as “successful” for purposes of this study). At the descriptive level, simple means and frequency distributions were employed. At the inferential level, Pearson product-moment correlation coefficients (Pearson $r$) were used to determine any statistically significant relationships ($p < .05$) between each selected independent variable and the grade obtained in the online course. A regression analysis was performed using the grade in the online course as the outcome variable and GPA, orientation attendance, previous withdrawals, ASSET reading, previous online courses, age, ACT English score, student
status, gender, ACT composite, ACT reading, semester format, and ASSET writing as the predictor variables.

Wojciechowski and Palmer (2005) reported about 72% (R² = 0.720) of the variability of course grades could be accounted for by taking all of the statistically significant variables, in order from most significant to least significant; GPA, orientation attendance, age, previous online courses, ASSET reading scores, and number of previous withdrawals. Then they calculated a second linear regression analysis using only the two most significant variables, which in this case were GPA and orientation attendance. This regression model suggests that the six variables were statistically significant predictors of course grade and explained 72% (R² = 0.720) of the variability within the outcome variable. The authors went on to do another regression analysis using the two most significant predictors in the model, GPA and attendance at orientation, and found that 69.2% (R² = 0.692) of the variability could be accounted for using these two variables alone.

The second regression model required only two variables, the other four variables only accounted for 3% of the variability of the course grade and were not added to be used in the final model. The semester format chosen (16 week fall and spring semester or 8 week accelerated summer semester) indicated that there was a relationship between the semester format selected for the class, and the grade received for those receiving a C or higher. The authors mentioned that this relationship may reflect that many “non-community college” students register for a shortened 8 week format in the summer while they are on break from their four year schools. If this is indeed the case then this could be a deficiency in the study. The data from the shortened summer sessions may be altering all the outcomes and misrepresenting the results of the true community college students within all areas of this
study. Though this is a significant omission, I do not think it negates the outcomes of the study. It may be valuable to remove the data from summer sessions and recalculate the results.

Wojciechowski and Palmer’s (2005) study gives a unique breakdown of the relationship between 13 different demographic or learner characteristics and provides strong correlations between them. It was not surprising that GPA was found to have the strongest correlation with success. Other studies have also indicated the higher the academic performance obtained, the better the student will do in online courses (Anderson & Benjamin, 1994; Argon & Johnson, 2004; Fike & Fike; 2009; Moore & Kearsley, 1996).

A second key finding from the Wojciechowski and Palmer (2005) study was that attending an orientation session for the class was the second greatest predictor to success. I have not found any other studies that examined this issue for online courses nor had the authors. This strong correlation between having attended an optional orientation session and success in the course was a surprising result that could have strong ramifications for online courses in general.

The third strongest correlation involves the number of previous withdrawals from other (online or traditional) classes. Wojciechowski and Palmer (2005) report not finding any other studies that examined connections between previous course withdrawals and future success in other courses. There are studies that compare withdrawals from online and on-site but not that looked at past withdrawals. This is very noteworthy.

Another interesting outcome from their work was between a student’s ASSET reading score, a test given by many community colleges for academic panning and advising,
and their final grade in an online course. This was statistically significant for the overall population although this variable no longer served to distinguish between those earning a C or better. This means that a student’s ASSET Reading score could be used to predict whether a student would earn at least a C or better in the online course, but not the specific grade above that C. This supports previous research by Phipps & Merisotis (1999) that indicted a correlation between literacy level of the student and success.

In summary, the goal of this study was to explore the relationship between 13 different demographic or learner characteristics. Several correlations were found. This study has provided some insight into the identification of characteristics associated with successful online students, and could provide the necessary information for teachers and admissions personnel to suggest or discourage a student from registering for an online course. A student mistakenly placed into a course may encounter more difficulties and have reduced chances for success compared to an appropriately place student. Further studies should be undertaken with larger samples using a longitudinal study looking at factors influencing persistence and retention of community college students in online learning.

Bantum

Bantum (2013) used Rovai’s (2002) Composite Persistence Model to explore the relationship between course delivery and its impact on student success at community colleges. A non-experimental causal-comparative design was utilized. Data was collected from 112 campuses of the California Community Colleges for the duration of three academic years. The use of existing data at multiple community colleges to explore the research
questions is unique, transcending much of the existing literature that reports on single institution case studies.

The three research questions they used were: Are there statistically significant differences in student characteristics and situational factors between first year community college students enrolled in online and campus-based developmental English courses? Does course delivery format influence course persistence in first year students enrolled in online and campus-based developmental English courses? Does course delivery format influence course success in first year students enrolled in online and campus-based developmental English courses?

Information was collected from 188,204 first year community college students who were enrolled in developmental English course sometime between the years 2008-2011. Students with any missing values, students under the age of 18, and students enrolled during the summer were all eliminated resulting in a final data set of 145,601 students. It is interesting to note that the sample was comprised of 99% of students who took a campus based course and 1% (n=1,395) had self-selected to take an online based course.

Descriptive statistics were reported for this study. A binomial logistic regression was also performed with course format as primary outcome variable. Course persistence is the primary outcome variable in this study. Student characteristics of gender, age, race, enrollment status, and eligibility for tuition waiver were used as the predictor variables. The instrument of research was the database in the MIS system at the Chancellor’s Office of the California Community College. The statewide database can store very large numbers efficiently which increased the generalizability of the study’s findings.
Statistically significant differences in specific student characteristics and success were found between students enrolled in online and campus-based developmental English courses. Frequency distributions and Chi-square tests indicated that there were statistically significant differences in each of the five student variables: age, gender, race/ethnicity, enrollment status, and eligibility for tuition waiver.

An exploratory analysis was conducted of the relationship between course delivery format and course persistence; does course delivery format influence course persistence rates? Course persistence was defined as the student completing the course. Logistic regression analysis was performed to test the null hypothesis that course delivery does not influence course persistence. There was a statistically significant relationship found between course persistence and course delivery format. A coefficient of -.187 indicates a lower probability of students in the online course persisting in the course until the end of the term as compared to students in campus-based course. The relationship is significant. Statistically controlling for the other independent variables (gender, race, age, enrollment status, and eligibility for tuition fee waiver), the logistic regression analysis determined that the odds of online students completing the course were .829 times lower than the students in the campus based course.

Next, course success and student variables were explored by analyzing the differences in the rates of course success among the students by each independent variable. Course success was defined in this study as the student earning a C or higher. Binary logistic regression analysis was conducted to test the null hypothesis that course delivery format does not influence course success. The analysis results indicated a statistically significant relationship between course success and course delivery format. Statistically controlling for
the other independent variables (age, sex, race, enrollment status, eligibility for tuition fee waiver), logistic regression analysis determined that students enrolled in the online course were significantly less likely to receive a final grade of a C or higher than students in the campus-based course.

Cross-tabulation of frequency data was examined. Pearson’s Chi-Square test was also performed to measure the likelihood that the association between independent variables and course success was caused by chance.

Limitations of this study include that it only looked at demographic data, it only included a limited number of demographic variables were tested for. Additional variables such as high school G.P.A could be important indicators. Student skills, such as computer literacy and time management, were not tested for, although, they have been documented as statistically significant for course persistence and success. The study included some but not all of the external factors Rovai addressed such as family responsibilities, outside encouragement, and hours of employment. Additionally, the fact that students self-selected into the online and campus-based sections may have impacted the results. The study failed to discuss if equal resources, such as tutoring, were available to the online students and the on-campus students. This study focused on the students in a developmental course. It is not clear if these results would also be applicable to students who were not in developmental courses.

Further studies are needed in order to expand the body of knowledge about online learning and indicate factors that impact student success in online courses, especially for the diverse group of students that attend the community college. The ability to identify student success factors enables counselors and faculty to better counsel students regarding their ability to successfully complete the course.
Jaggars and Xu

Jaggars and Xu (2010) completed a study on first year community college students who were enrolled in introductory college-level English course online (N=1,052) or in-seat (N=12,921). Demographic data of the students’ were taken from the community college database system. They concluded that demographic characteristics had a statistically significant impact on success and retention in online courses taken the first year. They examined students from twenty three community colleges.

The demographic characteristic results of typical online learning students indicated that online courses were significantly more popular among females, English fluent students, students who applied for financial aid, students who had enrolled in computer literacy or developmental courses, students who were between the ages of 25-50 years old, and students who had attempted online courses before. Black students and Hispanic students were significantly less likely to take an online course in the first semester than were white students. The students enrolled voluntarily in distance education courses.

The authors reported that students enrolled in online courses exhibited an 8% lower chance of success than their traditional in-seat college counterparts do. When they separated students enrolled in remedial classes, the results decrease to a 9% lower chance of success for those students. They also reported that students with more formal education experience were more likely to complete distance learning courses successfully. Most of these students take education seriously, are highly motivated, committed, task oriented, and want to use the material they have gained.
Jaggars and Xu (2011) recommend increased technological support for students and more extensive training in online-teaching methods for faculty. Their recommendation involved increased institutional support though they did not recommend specifically what that would entail. It would have been helpful if they would have included how they came to this conclusion. Were they suggesting technological support could include an easily accessible online help desk?

Community College Research Center

The Community College Research Center from Columbia University released a research overview that focused on online success and retention at community colleges over a five year time frame. This study was unique in that it also was considered the most comprehensive and largest research study found that specifically evaluated community college students’ success and retention in hybrid courses. It seems that very few studies focused specifically on hybrid courses. If they do discuss them they do not clearly delineate them as unique from online courses. Hybrid courses were defined by this study as having an online component of 51% or less online. Previously all the studies found had either looked at online or face-to-face courses. Researchers tracked 51,017 degree seeking students in hybrid courses from 2004-2009 across the 34 community or technical colleges in the Washington State System. The study focuses on three areas: (1) patterns of online and hybrid course enrollment (2) under-prepared and college ready student retention and performance in face-to-face, hybrid, and online courses (3) consequent outcome of both the college ready students and the under-prepared students who were in online, hybrid, and face-to-face courses.
The dataset containing information on all 51,017 students included: demographics, employment information, grades, transcript data such as each course taken including if it was online, hybrid, or face to face, and any developmental or college ready course taken. Students were defined as college ready or developmental based on if they had ever enrolled in a developmental course. Across the sample 40% of the students took a remedial course. This is typical to the nationwide data of students at community colleges receiving remedial courses.

Demographics in the study included gender, race, age, socioeconomic status, and employment hours for each quarter. The last two demographics are not always obtainable within studies, so this was unique and enabled the study to precisely control for student baseline characteristics allowing for more accurate results. This study tracked students over a five-year time frame which increased the data’s validity. It was also completed across multiple institutions which was able to capture students who transferred between institutions as well as students who dropped out and then returned to school. The number of online, hybrid, and face-to-face courses offered at each institution was not included. If some school offered more courses that may impact the findings. Type of analyses were not discussed in detail. When looking at each analysis it was conducted with the student as the unit of analysis one time and at other times the course as the unit of analysis. Although it was denoted which they chose in each situation the authors did not disclose why the analysis varied. The results of the study were innovative and unique regarding how student characteristics were collected from the school database, and multi-level regression techniques were used to control for student characteristics.

The Community College Research Center of Columbia University broke the study down and looked at student demographics. They reported that students who enrolled in
online classes were typically more academically prepared, from higher socioeconomic groups, and more fluent in English than the average community college student. When the study controlled for student economic and educational characteristics, the failure and withdrawal rates increased.

The study found that students enrolled in hybrid courses, who were reported as having the same baseline characteristics as the population as a whole, did equally as well in hybrid courses as they did in face-to-face courses. This suggests that hybrid courses may pose fewer challenge for students than online courses. These results are similar to results discussed earlier from the Virginia Community Colleges (Jaggars & Xu, 2011; Xu & Jaggars, 2011). This study did not find any consistent or significant differences between hybrid and face-to-face completion rates. These results suggest that hybrid courses may be a better fit for students and pose fewer problems than online course. It is recommended that much more research be done in the area of hybrid vs. face-to-face courses at the community college level to see if they do indeed seem to have a much higher retention and success rate than online vs. face-to-face classes as this study has concluded.

Summary

Many of the studies described above stated that they were using a retention model as a framework without regard to the larger picture. They were limited by scope and methodology. Many studies sought to explain only a particular demographic segment within an institution. Retention studies are important to help a particular institution examine its own persistence, yet in general they are lacking a guiding set of principles to ensure that a big
picture can be drawn from their analysis. Without guiding principles these studies cannot be tied together to make the larger picture. Each study used different sample sizes, pulled from different sub-populations within the institution, and different statistical methods were used. Path analysis was used as the basis for some of the studies described above even though Tinto, Bean and Metzner, and Rovai did not design their models as ones that can be validated using Path analysis. Conversely other studies attempted to define predictor variables on either persistence or attrition using Chi-square analysis to examine the relationship. This medley of studies further limits our understanding of the overall picture. The information each study provided regarding indicators of success are important pieces of the puzzle, but because of the fragmented methodology and non-operationalized constructs they do not fit together neatly.

Each experimental design or study with the goal of validating or using one of the listed models seemed to vary dramatically, causing it to be difficult to come to a consensus on what constructs are the best indicators of retention. For example, academic integration in retention studies has been measured by a variety of factors, including: grade performance, self-efficacy questionnaires about students’ enjoyment of the subject, interviews about study patterns, academic success, and identification with one's role as a student. Rovai’s model states the importance of academic integration but each study who measured academic integration seemed to define it and measure it very differently. Consistency needs to be developed if progress in understanding retention is to occur.

The increase in the number of students taking classes creates a need to provide student support services to these students. The high attrition rate in online classes has created a need for schools anticipate students who may need additional assistance or may not be
well-suited for online or hybrid courses. The goal of this study is to investigate if there could be a regression equation that uses readily assessable variables from the school database to identify students who are at risk for attrition or failure in online courses.

Rovai’s model suggests that it is possible to predict which students will persist in online courses based on a variety of external or internal factors. Many of the internal factors Rovai lists are either open ended or not clearly operationalized such as learning styles, study habits, or interpersonal relations. This would make it difficult for a student or advisor to accurately insert a variable to these open ended constructs. An examples of these internal factors in Rovai’s model are teaching style, program fit, or course availability.

We are most interested in evaluating if Rovai’s Composite Model of Student Persistence could accurately predict student persistence and success in an online course without the internal factors construct. The purpose of this study was to use logistic regression to answer the following research questions.

Research Questions

How well does modified Rovai’s Composite Persistence Model predict retention and success in on-line biology courses without internal factors? Specifically we ask:

1. Which characteristics are the best predictors of retention in an on-line biology course with respect to predictive accuracy of a student’s course completion using demographic characteristics, student skills, and external factors?
2. Which characteristics are the best predictors of success (obtaining a C or better) in an on-line biology course with respect to predictive accuracy of a student’s grade using demographic characteristics, student skills, and external factors?
CHAPTER III

METHODOLOGY

Purpose

The issue of attrition in online courses at higher learning institutions remains a high priority in the United States (Nash, 2010). Failure or attrition in an online course can cause students to become frustrated with the college experience, financially burdened, and some even give up and leave college (Jaggar & Xu, 2010). Specifically, student retention at the course level is key to assess the effectiveness of the cost of online learning compared to traditional classroom based teaching. The rapid growth of online courses at community colleges have been instigated by student demand. Online courses meet the time constraint demands many nontraditional community college students have as a result of the need to work and care for dependents. Studies have shown that students who took online courses early in their career were less likely to retain in school the following semester (Jaggars & Xu, 2010; Jaggars, 2011; Wright, 2013). Attrition or failure could be avoided by proper guidance of who is best suited for online courses.

Study Procedures

Study Sample Description

This is a retrospective review of six semesters of data from students who enrolled in an online biology course at a Midwestern community college. The data used for
the study were obtained from the official Grand Rapids Community College admission records and main campus database. This database houses data for each student from the time of application to the college and is updated on a regular basis (every semester). A query was made through the Institution Research and Planning Department at Grand Rapids Community College. The data were de-identified prior to receipt by the author.

Study Variables

Various student characteristics examined included Student Demographic characteristics (age, gender, race), Students Skills (criteria to be placed in an AFP course based on their Accuplacer scores), and External Factors (Pell grant recipient, academic load (full/part-time), first term). These variables were selected for this study based on the availability and accessibility of data in community college databases that relates to Rovai’s Composite Retention Model. Rovai’s Composite model will be summarized in the conclusion sections in relation to the findings from this study.

Data Analysis

Prior to analysis data the categorical variables were coded (Table 1). The dataset was then imported into the statistical software and analyzed using Stata/IC 13.0 for Mac (StataCorp, College Station, TX).
Table 1

Predictor Variables to be Tested

<table>
<thead>
<tr>
<th>Outcome Variables</th>
<th>Coding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research question 1: Course Completion (received a</td>
<td></td>
</tr>
<tr>
<td>grade D or better).</td>
<td>0=No 1=Yes</td>
</tr>
<tr>
<td>Research question 2: Course Success (C or better,</td>
<td></td>
</tr>
<tr>
<td>excluding withdraws, drops, incompletes)</td>
<td>0=No 1=Yes</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Predictor variables</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at the time of</td>
<td>Years</td>
</tr>
<tr>
<td>the course</td>
<td></td>
</tr>
<tr>
<td>Ethnicity</td>
<td>0=Nonwhite 1=White</td>
</tr>
<tr>
<td>Gender</td>
<td>0=Male 1=Female</td>
</tr>
<tr>
<td>Met placement</td>
<td>0=No 1=Yes</td>
</tr>
<tr>
<td>criteria for</td>
<td></td>
</tr>
<tr>
<td>Developmental</td>
<td></td>
</tr>
<tr>
<td>Writing, Reading</td>
<td></td>
</tr>
<tr>
<td>or Math</td>
<td></td>
</tr>
<tr>
<td>First semester</td>
<td>0=No 1=Yes</td>
</tr>
<tr>
<td>enrolled in college</td>
<td></td>
</tr>
<tr>
<td>Academic load</td>
<td>0=part-time 1=full time</td>
</tr>
<tr>
<td>Pell Grant Recipient</td>
<td>0=No 1=Yes</td>
</tr>
</tbody>
</table>

Univariate analyses were performed on the data. This included using the Chi-square test for nominal variables and the t-test for quantitative variables. The groups compared included those that completed the course vs. those that did not complete the course (corresponding to research question 1), as well as those that received C or better vs. those that did not receive a C or better (corresponding to research question 2).
Two logistic regression procedures were performed. The dependent variables were either course completion or a grade of C or better. The independent variables for both models were age, race, gender, first term, Pell grant, academic load and met criteria for a developmental course as the predictor variables.

Model performance for each logistic regression was also assessed using two methods. First, a 5-fold cross-validation procedure was used to check the data for over-fitting. This procedure estimates how accurately the predictive model would perform in practice. The strategy of cross validation is sample reuse. The dataset was randomly split into five groups that were similar in size. One group was used to test the model that was developed using the other four groups. This step was repeated until all five groups were used in model development and testing. For the 5-fold cross-validation analysis, the root mean square error (RMSE) was used to evaluate for over-fitting. If the RMSE from the logistic regression and the five RMSEs from the cross-validation procedure are more than 15% different from one another, over-fitting would be further investigated (Harrell, 2001).

Next, the area under the curve (AUC) was also obtained to evaluate model performance. The AUC, which is derived from the receiver operating characteristic curve (ROC), tells us how well the model was discriminating between those that complete the course and those that do not complete, as well as between those that received a C or better and those that did not. The x-axis on the ROC curve represents false positive results, while the y-axis represents true positive results. For example, if the study outcome was completion of the course, a false positive result would mean that the model had predicted completion, but in reality the student had not achieved course completion.
The criteria from Hosmer-Lemeshow were used to evaluate the AUC (Hosmer and Lemeshow (2000, p. 162). These criteria suggest the following general rule for interpreting the area under the ROC curve: AUC = 0.5 suggests no discrimination (i.e., no better than flipping a coin), 0.7 ≤ AUC < 0.8 is considered acceptable discrimination 0.8 ≤ AUC < 0.9 is considered excellent discrimination AUC ≥ 0.9 is considered outstanding discrimination (extremely unusual to observe this in practice).

Logistic Regression Assumption

Prior to running the analysis the author investigated the assumptions associated with logistic regression for each model. These assumptions include: independence, linear relationship between the continuous predictor variables and the logit transformation of the dependent variable, no multicollinearity among the predictor variables, no outliers or influential points and categories for each variable are mutually exclusive and exhaustive.

Common Assumptions for Both Regression Models

The assumptions for independence can be assumed for both regression models, as each observation in the dataset represents a unique individual. In addition, all categories for each variable were mutually exclusive and exhaustive. The predictor variables were assessed for multicollinearity, using the condition index (Belsley, 1991). If the condition index were >30, multicollinearity would have been further investigated. The condition index for the variables was 13.24. As the condition index was well below the criterion of 30,
Assumptions Model 1: Course Completion

Linearity was assessed for the regression model. It was determined that there was a linear relationship between the continuous predictor age and the logit transformation of the dependent variable (complete yes/no), using the Box-Tidwell procedure (Tabachnick & Fidell, 2013). This technique requires the creation of an interaction term, which for this dissertation was between the natural log of the continuous variable, in this case age, and the original age variable. This interaction term is then added to the model prior to running a logistic regression. If the interaction term was significant (p<0.05) in the model, linearity would have been an issue. In this case, the interaction term was not significant (p=0.984), therefore, the assumption of linearity was met.

The assumption of no significant outliers or influential points was assessed using the Cook’s D and the residuals. A Cook’s D >1 and or a residual >3 or <-3 were investigated further. There were no values >1 for Cook’s D. However, there was one case where the residual exceeded 3. This case was removed to see if it was influencing the model. The original $R^2$ and $R^2$ value without the suspect case were compared, as well as the classification tables for both models to assess for significant changes. Changes between the model for the $R^2$ (original: 0.083 vs. no outlier: 0.083) and classification tables (original: 92.2% vs. no outlier: 92.1%) were not deemed significant enough to remove the suspect case. Lastly, all categories for each variable were mutually exclusive and exhaustive.
Assumptions Model 2: C or Better

The assumption of linearity was also assessed for the C or Better regression. Based on the results of the Box-Tidwell procedure, there was a linear relationship between the continuous predictor age and the logit transformation of the dependent variable (C or better). The interaction term for the age variable that was added to the model to test for linearity was not significant (p=0.204).

Lastly, the assumption of no significant outliers or influential points was assessed using the Cook’s D and the residuals. A Cook’s D >1 and or a residual >3 or < -3 were investigated further. There were no values >1 for Cook’s D. However, there was one case where the residual exceeded 3. This case was removed to see if it was influencing the model. The original R² and R² value without the suspect case were compared, as well as the classification tables for both models to assess for significant changes. Changes between the model for the R² (original: 0.120 vs. no outlier: 0.126) and classification tables (original: 86.1% vs. no outlier: 86.8%) were not deemed significant enough to remove the suspect case. Lastly, all categories for each variable were mutually exclusive and exhaustive.

Missing Data

The dataset was also reviewed for missing data among the variables. Table 2 illustrates the percent of complete data for each variable.
### Table 2

Frequency: Missing Data

<table>
<thead>
<tr>
<th>Variable</th>
<th>% Complete</th>
</tr>
</thead>
<tbody>
<tr>
<td>Race</td>
<td>650/687 (94.6%)</td>
</tr>
<tr>
<td>Gender</td>
<td>687/687 (100%)</td>
</tr>
<tr>
<td>First Term</td>
<td>687/687 (100%)</td>
</tr>
<tr>
<td>Pell Grant</td>
<td>687/687 (100%)</td>
</tr>
<tr>
<td>Met criteria for developmental course</td>
<td>687/687 (100%)</td>
</tr>
<tr>
<td>Academic load</td>
<td>687/687 (100%)</td>
</tr>
<tr>
<td>Age</td>
<td>687/687 (100%)</td>
</tr>
<tr>
<td>Complete</td>
<td>687/687 (100%)</td>
</tr>
<tr>
<td>C or better</td>
<td>629/629 (100%)</td>
</tr>
</tbody>
</table>
CHAPTER IV
RESULTS

The dataset included a total of 687 students. There were 629 students who completed the course and 58 who did not. Of those 629 students, 546 students earned a grade of a C or better. Females made up 64% (n= 440) of the dataset while males represented 36% (n=247). Race was broken down into white and nonwhite categories with 82% of the dataset (n=514) reported as whites and 18% as nonwhite (n= 136). The age of the students enrolled in the course was 25.3 ±7.5 years (mean±SD; 15.3 minimum, 62.7 maximum).

Research Question 1: Univariate Results

Table 3 illustrates the univariate summaries of individual student characteristics with respect to predictive accuracy of a student’s course completion using demographic characteristics, student skills, and external factors as it relates to Research Question 1. There were statistically significant (p<0.05) differences between students who completed the course and those that did not with regard to the proportion of students who received Pell grants and white and non-white students in each group. A sub-analysis of white vs African American students only showed that there was a statistically difference between the two groups (p<0.05). A higher proportion of white students completed the course compared to African American students (data not shown). No statistically significant differences were found for gender, whether or not it was a students’ first term, academic load, age or whether or not a student met the criteria to be placed in an AFP course.
Table 3

Univariate Summary of Individual Student Characteristic for Complete/Did Not Complete

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Complete</th>
<th>Did Not Complete</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>482/599 (80.5%)</td>
<td>32/51 (62.7%)</td>
<td>0.003*</td>
</tr>
<tr>
<td>Non White</td>
<td>117/599 (19.5%)</td>
<td>19/51 (37.3%)</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>228/629 (36.2%)</td>
<td>19/58 (32.8%)</td>
<td>0.596</td>
</tr>
<tr>
<td>Female</td>
<td>401/629 (63.8%)</td>
<td>39/58 (67.2%)</td>
<td></td>
</tr>
<tr>
<td>First Term</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First Term</td>
<td>87/629 (13.8%)</td>
<td>11/58 (19%)</td>
<td>0.285</td>
</tr>
<tr>
<td>Not first term</td>
<td>542/629 (86%)</td>
<td>47/58 (81%)</td>
<td></td>
</tr>
<tr>
<td>Pell Grant</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Received grant</td>
<td>325/629 (51.7%)</td>
<td>38/58 (65.5%)</td>
<td>0.043*</td>
</tr>
<tr>
<td>Did not receive</td>
<td>304/629 (48.3%)</td>
<td>20/58 (34.5%)</td>
<td></td>
</tr>
<tr>
<td>AFP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Met criteria</td>
<td>425/629 (67.6%)</td>
<td>45/58 (77.6%)</td>
<td>0.116</td>
</tr>
<tr>
<td>Did not meet criteria</td>
<td>204/629 (32.4%)</td>
<td>13/58 (22.4%)</td>
<td></td>
</tr>
<tr>
<td>Academic Load</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full time</td>
<td>275/629 (43.7%)</td>
<td>18/58 (31.0%)</td>
<td>0.062</td>
</tr>
<tr>
<td>Part time</td>
<td>354/629 (56.3%)</td>
<td>40/58 (69.0%)</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>25.3 +/- 7.6 years</td>
<td>25.6 +/- 6.6 years</td>
<td>0.482</td>
</tr>
</tbody>
</table>

*significant, (p value is less than 0.05)

Research Question 2: Univariate Results

Table 4 illustrates the univariate summaries of individual student characteristics, internal student characteristics, and external characteristics with respect to predictive accuracy of a student’s course grades as it relates to Research Question 2. There were
statistically significant (p<0.05) differences between students who received a C or better and those that did not with regard to students who received Pell grants, academic load, met AFP course criteria, age and the proportion of white and non-white students in each group. A sub-analysis of white vs African American students only showed that there was a statistically difference between the two groups (p<0.05). A higher proportion of white students received a C or better compared to African American students (data not shown). No statistically significant differences were found for gender and whether or not it was a students' first term.

Table 4

Results for Comparisons Between Those Who Earned a C or Better and Those That Completed the Course But Did Not Receive a C or Better

<table>
<thead>
<tr>
<th></th>
<th>C or Better</th>
<th>Not C or Better</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>429/519 (82.7%)</td>
<td>53/80 (66.3%)</td>
<td>0.001*</td>
</tr>
<tr>
<td>Non White</td>
<td>90/519 (17.3%)</td>
<td>27/80 (33.8%)</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>201/546 (36.8%)</td>
<td>(27/83 32.5%)</td>
<td>0.449</td>
</tr>
<tr>
<td>Female</td>
<td>345/546 (63.2%)</td>
<td>56/83 (67.5%)</td>
<td></td>
</tr>
<tr>
<td>First Term</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First Term</td>
<td>74/546 (13.6%)</td>
<td>13/83 (15.7%)</td>
<td>0.604</td>
</tr>
<tr>
<td>Not First Term</td>
<td>472/546 (86.4%)</td>
<td>70/83 (84.3%)</td>
<td></td>
</tr>
<tr>
<td>Pell Grant</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Received Grant</td>
<td>264/546 (48.4%)</td>
<td>61/83 (73.5%)</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>Did not Receive</td>
<td>300/546 (54.9%)</td>
<td>22/83 (26.5%)</td>
<td></td>
</tr>
<tr>
<td>AFP Courses</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Met Criteria</td>
<td>361/546 (66.1%)</td>
<td>64/83 (77.1%)</td>
<td>0.046*</td>
</tr>
<tr>
<td>Did not meet</td>
<td>185/546 (33.8 %)</td>
<td>19/83 (22.9%)</td>
<td></td>
</tr>
<tr>
<td>Academic Load</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full Time</td>
<td>228/546 (41.8%)</td>
<td>47/83 (56.6%)</td>
<td>0.011*</td>
</tr>
<tr>
<td>Part Time</td>
<td>316/546 (58.2%)</td>
<td>36/83 (43.3%)</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>25.5 +/- 7.7 years</td>
<td>23.9 +/- 6.5 years</td>
<td>0.044*</td>
</tr>
</tbody>
</table>

*significant, (p value is less than 0.05)
Research Question 1: Logistic Regression Analysis

A logistic regression analysis was performed using C or better as the outcome variable and gender, age, race, academic load, Pell recipient, met criteria for academic foundation program (AFP) course enrollment and first semester as the independent variables. There were 650 students included in the analysis. The Hosmer and Lemeshow goodness of fit statistic was $\chi^2 (6) = 1.94$ ($p=0.379$). This statistic was not significant therefore it is a well-fitting model. The Nagelkerke pseudo $R^2 = 0.083$, meaning the predictor variables explain 8.3% of the variability within the outcome variable.

Table 5 shows the results of the multivariate logistic regression. Significant predictors in the model included age, race, academic load and first term. For every additional year of age students were 1.1 times more likely to complete the course. White students were 2.4 times more likely to complete the course than non-white students. Full-time students were 2.1 times more likely to complete the course. Students in their first term were 2.6 times less likely to complete the course.

The ROC for the model is shown in Figure 3. The AUC for the model was 0.691. This value fell marginally short of the minimum value for acceptable discrimination as described by Hosmer and Lemeshow (2000). A five-fold cross-validation was performed to assess the model for over-fitting. The RMSE was used to evaluate for over-fitting. The comparison of the RMSEs in Table 6 showed that they were no more than 15% different from one another (Harrell, 2001). Based on this criterion, over-fitting was not an issue.
Table 5

Logistic Regression Results for the Completion Outcome

<table>
<thead>
<tr>
<th>Variable</th>
<th>OR</th>
<th>SE</th>
<th>P value</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>0.851</td>
<td>0.280</td>
<td>0.624</td>
<td>0.447</td>
</tr>
<tr>
<td>Age</td>
<td>1.060</td>
<td>0.029</td>
<td>0.034</td>
<td>1.00</td>
</tr>
<tr>
<td>Race</td>
<td>2.418</td>
<td>0.788</td>
<td>0.007</td>
<td>1.276</td>
</tr>
<tr>
<td>Academic Load</td>
<td>2.149</td>
<td>0.715</td>
<td>0.022</td>
<td>1.119</td>
</tr>
<tr>
<td>Pell Recipient</td>
<td>0.623</td>
<td>0.208</td>
<td>0.156</td>
<td>0.324</td>
</tr>
<tr>
<td>Met Criteria for AFP course</td>
<td>0.470</td>
<td>0.204</td>
<td>0.082</td>
<td>0.200</td>
</tr>
<tr>
<td>First Term</td>
<td>0.382</td>
<td>0.174</td>
<td>0.034</td>
<td>0.157</td>
</tr>
</tbody>
</table>

Figure 3. ROC curve for completion regression model.
Research Question 2: Logistic Regression Analysis

A logistic regression analysis was performed using course completion as the outcome variable and gender, age, race, academic load, Pell recipient, met criteria for academic foundation program (AFP) course enrollment and first semester as the independent variables. There were 599 students included in the analysis. The Hosmer and Lemeshow goodness of fit statistic was $\chi^2(6) = 4.86$ ($p=0.088$). This statistic was not significant therefore it is a well-fitting model. The Nagelkerke psuedo $R^2=0.120$, meaning the predictor variables explain 12.0% of the variability within the outcome variable.

Table 7 shows the results of the logistic regression. Significant predictors in the model included age, race, and Pell recipient. For every additional year of age, students were 1.1 times more likely to achieve a C or better in BI101. White students were 2.5 times more
likely to achieve a C or better in BI101 than non-white students. Pell recipients were 2.7 times less likely to achieve a C or better in BI101.

*Table 7*

Logistic Regression Results for the C or Better Outcome

<table>
<thead>
<tr>
<th>Variable</th>
<th>OR</th>
<th>SE</th>
<th>P value</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>1.013</td>
<td>0.270</td>
<td>0.960</td>
<td>0.601</td>
</tr>
<tr>
<td>Age</td>
<td>1.053</td>
<td>0.024</td>
<td>0.025</td>
<td>1.006</td>
</tr>
<tr>
<td>Race</td>
<td>2.518</td>
<td>0.718</td>
<td>0.001</td>
<td>1.440</td>
</tr>
<tr>
<td>Academic Load</td>
<td>0.699</td>
<td>0.184</td>
<td>0.173</td>
<td>0.417</td>
</tr>
<tr>
<td>Pell Recipient</td>
<td>0.368</td>
<td>0.107</td>
<td>0.001</td>
<td>0.208</td>
</tr>
<tr>
<td>Met Criteria for AFP course</td>
<td>0.560</td>
<td>0.198</td>
<td>0.101</td>
<td>0.280</td>
</tr>
<tr>
<td>First Term</td>
<td>0.501</td>
<td>0.205</td>
<td>0.091</td>
<td>0.225</td>
</tr>
</tbody>
</table>

The ROC for the model is shown in Figure 4. The AUC for the model was 0.707. This met the criteria for acceptable discrimination as described by Hosmer and Lemeshow (2000). A five-fold cross-validation was performed to assess the model for over-fitting. The RMSE was used to evaluate for over-fitting. The comparison of the RMSEs in Table 8 showed that they were no more than 15% different from one another (Harrell, 2001). Based on this criterion, over-fitting was not an issue.
Figure 4. ROC curve for C or better regression model.
Table 8

Five-fold Cross-validation RMES Comparison for C or Better Regression Model

<table>
<thead>
<tr>
<th></th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original Model</td>
<td>0.330</td>
</tr>
<tr>
<td>1</td>
<td>0.366</td>
</tr>
<tr>
<td>2</td>
<td>0.339</td>
</tr>
<tr>
<td>3</td>
<td>0.304</td>
</tr>
<tr>
<td>4</td>
<td>0.331</td>
</tr>
<tr>
<td>5</td>
<td>0.320</td>
</tr>
</tbody>
</table>

Summary

For the logistic regression analyses, these results indicated that there are predictive characteristics, for both course completion and earning a grade of C or better in an online course, that are available in higher learning databases. Predictive characteristics of students who completed the online course were age, ethnicity, academic load, and first term. Predictive characteristics for students who received a C or better in the course included race, age, and Pell recipient. These results are based on data present in higher institutional databases and do not include the information Rovai referred to as internal factors.
CHAPTER V
CONCLUSION OF RESULTS

Introduction

The organization and content of this paper touched on retention models as presented by Tinto (1975, 1987, 1993), Bean and Metzner (1985) and Rovai (2002). Rovai’s (2002) Composite Model. Vincent Tinto's Retention Model is the most cited and documented theory of attrition to date in higher education. Tinto’s model has been criticized (Bean & Metzner, 1985; Rovai, 2001) as a theory that is only applicable to the homogeneity of a four year, on campus college setting. His research predominantly dealt with traditional students at four-year residential institutions and did not address the many issues of nontraditional students that differ from traditional students. The model has limited applicability as it does not address the retention of non-traditional students, commuter students, and community college students.

Bean & Metzner's (1985) model proposed their own model, based on Tinto's work. They altered Tinto's Model to fit the unique characteristics of nontraditional students. Social integration and academic integration were replaced with environmental issues of nontraditional students face, such as family responsibility, hours employed, and ability to pay for college. While Bean and Metzner's model was meant to explain retention of nontraditional students, many of these characteristics also fit those of online learners. Online learners, though, have even less academic and social integration within the institution (Rovai,
Their limited interaction within the college community causes them to draw support from peers, friends, family, and employers outside of college.

Rovai’s (2002) Composite Persistence Model developed a model that focuses specifically on the retention of online learners. Based on a thorough review of Tinto’s Student Integration Model (1993) and Bean and Metzner’s Student Attrition Model (1985) Rovai addressed the idea that although online learners are often nontraditional they are unique. Online learners do not fit into the other models. Rovai’s addresses this and his model is adapted to the needs of the online learner. Specifically Rovai’s model focuses on the factors affecting a student’s decision to drop out of online courses as well as the variables that increase success in these courses. He declared his model to be a better predictor of the persistence of nontraditional, community college students enrolled in online coursework.

Rovai’s Composite Model was applicable to this study because it relates to the unique demographic of non-traditional and online student. Tinto, Bean and Metzner, and Rovai all developed models all requiring a large percent of self-report information that is not readily available in a dataset. Variables such as ‘Intent to leave’ and psychological outcomes (satisfaction, goal commitment, and stress) are difficult and time consuming for institutions to put it into practice. They also contain much bias, variability, and require the study to be retroactive instead of predictive.

As described throughout this paper, Rovai Composite Model (Figure 1) contains variables based on both Tinto’s (1987) and Bean and Metzner’s (1985) models. Examples include, but are not limited to, age, gender, race, intellectual development, academic performance and finances. Rovai incorporated the impact of external factors in shaping students’ perceptions, reactions, and commitments. These types of factors were lacking in
Bean & Metzner (1985) and Tinto’s (1987) model and are especially significant to the community college student and online student (Rovai, 2003).

Findings

This study was based on Rovai’s Model of Persistence (Figure 5). This study operationalized three of the four construct in Rovai’s Model of Student Persistence: Student characteristics, student skills, and external factors. The variables listed under each construct were chosen based on data that is easily available and assessable in school databases. The fourth box, internal factors, was eliminated. The purpose of this was to evaluate the predictive accuracy of a student’s course completion and course success (C or better) using demographic characteristics, student skills, and external factors. While it is agreed that the internal factors construct contain many important predicative variables, these variables all require students to self-report and are not easily accessible.

The rational of how each construct was operationalized, the results from this study, and their relationship to past research is explained individually below.
Student Characteristics

Rovai did not operationalize the construct of student characteristic box, yet he included examples of age, race, gender, intellectual development, and academic preparation. In this study, the construct of student characteristics was operationalized with age, race, and gender. These characteristics were chosen based on data that was readily available in the institutional database. High school GPA was not included in this study as a result of missing data. This predictor could be added in future studies. Intellectual development was mentioned by Rovai, but not explained further, and was not used in this study.

Figure 5. Rovai’s model of student persistence.
Race

The results supported the literature regarding ethnicity and retention and success in online courses. Race was broken down into White and Non-white categories. Non-white students, though under-represented in this study, were more 2.4 times less likely to complete the course than the White majority. This supports the retention literature that report minorities were less likely to persist in online courses or programs (Carter, 2006; Jaggars, 2011). Conversely, Aragon and Johnson (2008) reported race was not significant to dropout or persistence in the online classroom.

Research question 2 focuses on the outcome variable of students earning a C or better. Few studies have been conducted which examine the relationship between race and academic performance in online courses. This study reported that race was a significant predictor. White students were 2.5 times more likely to achieve a C or better in BI101 than Non-white students. These results supported research reported by Bean and Metzner (1985), Tinto (1993), and Rovai (2002) reporting that race had strong negative influence on course success. They based those reports on the belief that poorer education is provided for minority students at the secondary level.

Age

For every year of age, students were 1.1 times more likely to complete the online course. These results support Wojcicowski and Palmer’s (2005) study that reported successful students were older (Wojcichowski & Palmer, 2005). Conversely, age was reported by Aargon and Johnson’s (2008) as having an indirect effect on attrition because
older students reported to have more hours of employment, family responsibilities, and increased absenteeism than younger students.

Research Question 2, focused on the outcome variable of earning a C or better. The results indicated that age was a significant predictor in the model included age, race. For every year of age, students were 1.1 times more likely to achieve a C or better in BI101.

Gender

Gender was not a significant in the univariate or multivariate regression results for either of the research questions in this study. This is contradictory to a study by Aragon and Johnson (2008) that reported a positive correlation to persistence in contrast to the male students. It supports studies by Yukselturk and Bulut (2007) and Wojciechoski and Palmer (2006) who found gender as a variable was unrelated to learning outcomes in online courses. Other factors may impact course persistence and gender. This study did not support the studies that suggested that gender may have an indirect effect on attrition based on family responsibilities that are greater for women and also less opportunity to transfer (Tinto; Bean & Metzner, Prince & Jenkins, 2005; Rovai, 2001).

Student Skills

The student skills construct in Rovai’s model listed computer skills, time management, information literacy, reading, and writing skills. In regard to data that is available in institutional databases, the placement of student into academic foundation programs (AFP) was used for the student skill construct. In this dataset, students who score
below and 18 on the ACT test are required to take an Accuplacer test. The results from that test indicate whether or not the student is placed into AFP course(s).

Online course persistence studies suggest that online courses are often associated with less desirable course outcomes for underprepared students (Carpenter, Brown, & Hickman, 2004; Fike & Fike). According to the univariate regression results in this study, students’ ability to pass the academic foundation courses (AFP) was significant to their ability to earn a C or better. Although when controlling for other variables in the model, this variable was no longer a significant predictor of receiving a C or better. This variable did not make it into the model for either completion or C or better. Information was not available regarding when students were enrolled in the AFP courses, it is possible at this intuition to be enrolled in an AFP course while taking the online course. This may be one of the reasons for the varied results.

Bean and Metzner indicated that high school academic performance, grade point average (GPA), was listed as the strongest pre-enrollment predictor of college. This data was not readily available in this study’s dataset, it would be recommended to include this data in future studies.

External Factors

Rovai’s external factors construct includes a variety of examples including finances, hours of employment, enrollment status, family responsibilities, outside support, and life crisis. The specific database used in this study was not complete. Additional factors such as number of dependents, hours of employment, study habits, learning style, and self-esteem were not included. This information could be used in future studies if institutions had a more
complete database. Life crisis was not used because the dataset was not complete nor was the
definition of what constitutes a life crisis. A survey could collect this type of information
although this study’s goal was to use information that was already collected and easily
accessible to institutions for ease of use.

Pell Grant

The receipt of a Pell grant was used as a way to operationalize this construct. This is
information that was assessable in the database. The results indicated that receiving financial
aid is a not a predictor of student completion. Although, the results indicated that receiving
financial aid is a negative predictor of student success. Pell recipients were 2.7 times less
likely to achieve a C or better in BI101. This is consistent with Jaggars and Xu (2010)
findings that financial aid is associate with persistence and success and that students. The
greater the financial need the lower students’ completion rates are. (Lichtenstin, 2002).

Enrollment Status

Enrollment status was also used as a way to operationalize external factors in this
study. Enrollment status was defined as full time students (enrolled in 12 credit or more) and
part time students (enrolled in less than 12 credits). Academic load was a positive predictor
for completion in the online course, full-time students were 2.1 times more likely to complete
the course than part-time students. These results support research that enrollment status was
an important predictor of persistence. Part-time students were described as being much more
likely to withdraw than full-time students (Argon & Johnson, 2008; Conklin, 1997; Grimes
& Antworth, 1996; Rajasekhra & Hirsch, 2000).
In regards to Research question 2, academic load was not significant for earning a C or better in the Course. Enrollment status was not used as predictor variable in the model for C or better.

First Semester Students

Enrollment in the online course during their first semester of college was also used to operationalize the external construct box. Results indicated that students in their first term were 2.6 times less likely to complete the course than students who had taken more than one semester. This was not a positive predictive characteristic for univariant, but when multivariant regression was run it became a predictor. This study supports findings from Schlosser & Anderson (1994) that first year students were less likely to be successful in online-based courses. Conversely, first semester enrollment was not an indicator of earning a C or better. This study did not support the literature of first term students are less successful in online courses (Jaggars, 2011; Wright, 2013).

Rovai (2002) and Bean and Metzner (1987) indicated that high school academic performance, grade point average (GPA), was listed as the strongest pre-enrollment predictor of college. This data was not readily available in this study's dataset, it would be recommended to include this data in future studies.

Model Evaluation

The findings from the Logistic Regression Analysis Results listed age, race, academic load, and first term as the independent variables and completion as the outcome variable. This model performed fairly well, it is important to note the amount of variability that this
model explains is quite low. This implies that there are many other factors related to completion. This is not to be overlooked. Although, this study was designed to test the predictive ability of data that is easily attainable in institutional databases. It purposefully did not include the many other factors that also play a role in student persistence and success such as personal issues, family dependents, learning styles, and computer skills. The results are based on data present in higher institutional databases and do not include the information Rovai referred to as internal factors.

The logistic regression analysis results for research question 2 with the outcome of a C or better reported that significant predictors in the model included age, race, and Pell recipient. This model was a well performing model. These results are insightful as they are based on information found in institutional database and do not include the information Rovai referred to as internal factors. The internal factors box within Rovai’s model contains many internal factors that also influence persistence and success although they require qualitative collection techniques and/or additional information not available in the institutional database. These results are informative. They illustrate that the newly operationalized model, with the removal of the internal factors construct, may provide opportunities for institutions to use data as predictive tool. This could allow them to monitor students in need of assistance regarding persistence and success before they are statistics of attrition and failure.

Although significant predictors were seen in both regression models, the variables included in each model explained very little (<15%) of the variability within the outcome variables of retention and success. Therefore, it can be concluded that there are still many factors that remain unaccounted for that explain the outcomes of success and persistence.
Rovai’s internal factors construct includes a vast amount of information that cannot be easily operationalized or quickly collected. This study reports that the other three constructs (student characteristics, student skills, and external factors) may be able to act as indicators of potential attrition or failure.

Figure 6, The Predictor Model for Success and Persistence in Online Courses, is a revised version of Rovai’s Composite Persistence Model. Rovai’s categories of student characteristics, student skills, external factors and internal factors constructs. They are illustrated as boxes that directly relate to persistence. Rovai’s model is designed in a way that three constructs, student skills, student characteristics and external factors feed into an internal factors box. The findings of this study recommend an alteration to the model. Each of the four constructs directly contribute to persistence rather than indirectly contributing to internal factors and then persistence. In the new model (Figure 6) the four constructs are all presented as equal, contributing factors of persistence.

Figure 6. New predictor model for success and persistence in online courses.
With increased demand for online learning, as well as more institutions of higher learning striving to provide diverse educational opportunities, online course delivery continues to grow as a viable means of providing increased access to a greater number of students. Research targeting the retention of community college students in online courses needs to be addressed. The practice of increasing online course offerings to all community college students may be an appropriate utilization of economic resources for some students, but not for others. It is important to evaluate the factors that impact student retention in order to provide student services and guidance to students regarding their potential to be successful in an online course based on their characteristics, skills, and attributes. This inaugural research indicates that student characteristics, student skills, and external factors may be tools that could be used to assist students. By evaluating these constructs, institutions could provide guidance to students regarding their calculated potential to be successful in an online course.

This study evaluated the ability to predict student retention and success based on readily available data of student characteristics, student skills, and external factors. This information could be used to guide institutional decisions regarding online course development, enrollment, and support. The results of this retention study are note-worthy. The movement of the internal factors box, along with the operationalization of the constructs, provides increased opportunities for institutions to engage in retention studies as a preventative method.

This study is paramount in advancing retention research from theoretical, retroactive, studies to predicting student attrition and providing assistance to those who are at risk. Currently, retention theory has been presented in the literature as isolated, univariate,
retroactive studies that are being used to explain what has occurred. It is imperative for higher learning institutions to move beyond these retroactive retention research and practice.

Limitations

The study was limited to online students from a specific community college in the Midwest. The sample was drawn solely from students who were enrolled in one Biology course, therefore the sample was not necessarily representative of the overall population of online learners. Although the homogenous group was a weakness, it was also a strength of the study as variability was limited. The exact results from this study (predictor characteristics) may not be applicable to another class, because the variables may differ. The results do, however, show that it is possible to operationalize the constructs based on assessable data. It is possible that in another course this may not be replicable, based on data readily available.

Future Studies

Further research into the operationalization of Rovai’s model, including internal factors and other means of defining student characteristics, skills, and external factors, is recommended. This information may be beneficial to designate resources needed and assist students in making course modality choices. Administrators and counselors could use this data to identify students who are best suited for online courses in order to avoid attrition or failure. It is recommended that additional studies should be completed in additional courses, with different disciplines, and also with courses that have less uniformity of curriculum. The
Predictor Model for Success and Persistence in Online Courses could be incorporated into institutional databases as a predictive tool to flag students who are at risk of attrition or failure. This model illustrates a progressive, predictive, model based on information already present in each community college database. The themes revealed in this study provide a basis for developing an at-risk score to assist institutions in flagging students who may struggle in online courses. Despite suggestive themes from theory and research, the specific characteristics of students who are successful in online courses remain elusive.

Students continue to sign up for online courses because the online learning environment offers flexibility and convenience that is typically not available through traditional higher learning education. Attrition and failure in online courses is higher than in-seat courses. The research literature lacks a consistent set of constructs for predicting student success and completion. Opportunities still exist for further exploration of operationalization of retention theories in the literature to bridge the gap between theory and practice. Retention studies continue to evolve and incorporate new variables into the research, thus reflecting the changing dynamics in higher education. Additional studies are recommended to evaluate if valuable information could be obtained by determining if the demographic, student enrollment, or external characteristics could predict student retention and completion in online courses.

Conclusions

This study was unique in that it attempted to put theory into practice by operationalizing three of the constructs of Rovai’s retention theory with data that was attainable and present in the databases intuitions of higher learning. The use of the
operationalized constructs (student skills, student characteristics and external factors) resulted in a model with positive predictive ability for success.

This study is a step in the direction of developing a predictive model to guide current students toward choosing effective course work that increase student persistence and success. Not all students are able to be successful in the online course environment. The use of a predictive model may allow institutions to use readily available data to aid in the pursuit of increased retention and success. Attrition or failure could be avoided by proper guidance of who is best suited for online courses.
REFERENCES


StataCorp, College Station, TX.

University of Chicago Press, 5801 S. Ellis Avenue, Chicago, IL 60637.


Appendix A

HSIRB Approval Letter
Date: May 5, 2015

To: Brandy Skjold, Principal Investigator
    Sarah Krajewski, Student Investigator for dissertation

From: Amy Naugle, Ph.D., Chair

Re: HSIRB Project Number 15-05-03

This letter will serve as confirmation that your research project titled “Investigation of the Factors which Contribute to the Success and Retention of Students Enrolled in Hybrid Science Courses at a Community College” has been approved under the exempt category of review by the Human Subjects Institutional Review Board. The conditions and duration of this approval are specified in the Policies of Western Michigan University. You may now begin to implement the research as described in the application.

Please note: This research may only be conducted exactly in the form it was approved. You must seek specific board approval for any changes in this project (e.g., you must request a post approval change to enroll subjects beyond the number stated in your application under “Number of subjects you want to complete the study”). Failure to obtain approval for changes will result in a protocol deviation. In addition, if there are any unanticipated adverse reactions or unanticipated events associated with the conduct of this research, you should immediately suspend the project and contact the Chair of the HSIRB for consultation.

Reapproval of the project is required if it extends beyond the termination date stated below.

The Board wishes you success in the pursuit of your research goals.

Approval Termination: May 4, 2016