April 2018

Predictive Validity of Florida's Postsecondary Education Readiness Test

Alisa Murphy Žujović
University of South Florida, alisazujovic@hotmail.com

Predictive Validity of Florida's Postsecondary Education Readiness Test

by

Alisa Murphy Žujović

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Curriculum and Instruction
with a concentration in Educational Measurement and Evaluation
Department of Educational and Psychological Studies
College of Education
University of South Florida

Major Professor: Robert Dedrick, Ph.D.
John Ferron, Ph.D.
Yi-Hsin Chen, Ph.D.
W. Robert Sullins, Ed.D.

Date of Approval:
March 30, 2018

Keywords: developmental education, placement testing, community college

DEDICATION

To my loving husband and best friend, Ivan: this Ph.D. journey has been a long and arduous road…certainly not without its extreme challenges….but through it all, you've been my backbone and my strength. Thank you for loving me unconditionally and supporting my every endeavor. To my children, Dameion, CJ, Ivana, Luka and Mila, use this as proof that you're never too old, it's never too late, and there's never too much going on to achieve something you want in life.

To my parents and family, Sharlott & Julius Murphy, Nette Levingston, and Eference Murphy, you guys are my lifeline and I love you all dearly. At countless times, the memory of my loving and cherished sister, Kim V. Murphy encouraged me to keep moving forward. To my extended family, I love and thank you all.

Natalija (Kevo) i Marko (Ćale): Nemam reci da kažem koliko ja vas volim, i koliko sam zahvalna vama. Stavi ovo na terasu da svi vide! Smok! Za moju najbolju zaovu, na svetu (Ljiljana), volim te od srca.

To all of you, I dedicate this dissertation.

ACKNOWLEDGMENTS

"Appreciation" is not an adequate enough term to describe how thankful I am for the committee that guided me through this process.

To my major professor, Dr. Robert Dedrick, or 'Dr. D.', I don't know how to thank you for all you've done. Your constant advisement at any and every time has given me what I've needed to push through. To my committee (those former and present), Dr. John Ferron, Dr. Yi-Hsin Chen, Dr. W. Robert Sullins, Dr. Bill Young, Dr. Jeff Kromrey, and Dr. Don Dellow, none of this would have been possible without your guidance and assistance at every turn.

To the Florida Education Fund, thank you for accepting me as a McKnight fellow, and for supporting my educational dream.

To the many members of my Hillsborough Community College family, your motivation and constant encouragement have helped me more than you know. Dr. Nicole Jagusztyn, thank you for being another set of eyes on my programming. Dr. Ken Ray, thank you for your ear, and the reference book for people going through the dissertation struggle. To Dr. Sylvia Marion Carley and the late Dr. Gwendolyn W. Stephenson, thank you for being strong professional women that I continue to look up to. Elizabeth Key Raimer and Grace Glenn, thank you for showing up and supporting me at my final defense. Your presence was much appreciated! Finally, Dr. Paul Nagy—my mentor, my brother, forever my professional inspiration— thank you so much for the relentless urging to complete my doctorate and for holding me accountable. Let's toast to that coveted spot on your bookshelf!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The role of the community college is constantly evolving. At its inception in the early 1900's, the community college's broad focus was to provide quality, affordable education to the members of the community the college serves. Today, that focus remains the same, but has also morphed into one that meets the specific needs of its students. One of these needs that is a critical issue for community colleges relates to developmental education.

The assessment of developmental education has been a contentious subject among higher education institutions. Defining college readiness, methods describing how to measure it, and instruments with which to measure it, have all been issues that higher education researchers have debated. Using multilevel modeling, this study evaluated a customized developmental education assessment measure in a single community college in Florida, and its ability to correctly place students in appropriate courses.

The Postsecondary Education Readiness Test (PERT) was implemented in Florida in 2010 as the primary gauge of student readiness based on competencies identified by Florida's high school, college and university faculty. PERT assesses these competencies in the areas of mathematics, reading and writing. The courses of interest in this study were four math courses offered in community colleges across Florida: Developmental Math I (MAT 0018), Developmental Math II (MAT 0028), Intermediate Algebra (MAT 1033), and College Algebra (MAC 1105).

The sample for Developmental Math I consisted of 727 students in 64 sections; for Developmental Math II, 900 students in 197 sections; for Intermediate Algebra, 713 students in 328 sections; and for College Algebra, 270 students in 204 sections. Five models were formulated to investigate the predictive validity of the PERT with final grades in the

aforementioned math courses. These models also analyzed the relationships with student and course level predictors. Student level predictors included whether student had a first time in college status, student race/ethnicity, gender, student enrollment status (part-time or full-time), age, PERT score, and final grade in the math course. Course level variables consisted of employment status of instructor (part-time or full-time), the number of years the instructor had been employed, time of day of the course (day or evening), and the course delivery method (on campus or online).

Results of this study indicated that the PERT score was a significant predictor for Developmental Math I, Developmental Math II, and College Algebra showing a positive relationship with final grade in each of these courses. Four of the research questions inquired as to whether interaction effects with the PERT score and race, and PERT score and gender existed. No interaction were significant, which indicated that no differential predictive validity was evident. The remaining two research questions examined the level of variance associated with the student and course level variables. For Developmental Math I, Black students had lower final grades than White students, and older students performed better than younger students. In Developmental Math II, female students had higher final grades than males, and older students had higher grades. For the credit-level courses, in Intermediate Algebra, full-time students had higher final grades than part-time students, and once again, older students exhibited higher grades. In College Algebra, for the final model, only the PERT score was significant. No other student nor course level variables was found to be significant predictors of final grade.

These results are only a preliminary view of how PERT test scores relate to final math grades in only one institution in Florida. Statewide standard setting procedures are necessary in order to properly assess whether cut score for the PERT are appropriate, and to determine if this test is properly measuring the construct it intends in order to verify the reliability of the test items, and the validity of the test itself.

CHAPTER ONE

INTRODUCTION

There has always been, at the core of the community college mission, a need to provide quality education to individuals who come from various walks of life. At the community college, there are recent high school graduates, married students with families, older students looking to start a new career, and professionals looking for continuing education credits. The community college accepts all of these students, regardless of the knowledge, skills, and abilities the student possesses. Statistics compiled by the American Association of Community Colleges (2016) indicate that there are a total of 1,132 community colleges in the United States. Demographic data show that the average age for a community college student is 28. Fifty-seven percent are women and the racial/ethnic background is varied. Fifty-two percent of students reported that they were White, 18% Hispanic, 15% Black, 6% Asian/Pacific Islander, 1% Native American, and 9% reported as ‘Other’ or were of unknown ethnicity.

The community college at its inception with Joliet Junior College in 1901 was designed to assist in teaching students the reading, writing and math skills they were lacking upon entering postsecondary education. Cohen and Brawer (2003) note in *The American Community College* that “…the rise in remedial course enrollment occurred because student ability had sunk so low that college staff members, legislators, and the staff of the universities to which the students transfer had had enough” (p. 262). The role of the community college in present day is to educate and prepare its students for the workforce, and for further formal education. Ensuring that students are successful in college-level courses increases the likelihood that these students will become successful in their respective goals.

In view of the diversity in the student population and the goals of these students, developmental education in the community college has come to the forefront of these institutions. The National Association for Developmental Education (NADE) defines developmental education as

> …a field of practice and research within higher education with theoretical foundation in developmental psychology and learning theory. It promotes the cognitive and affective growth of all postsecondary learners, at all levels of the learning continuum. Developmental education is sensitive and responsive to individual differences and special needs among learners. Developmental education programs and services commonly address academic preparedness, diagnostic assessment and placement, development of general and discipline-specific learning strategies, and affective barriers to learning (http://www.nade.net/aboutdeved.html).

The K-12 sector has the responsibility of ensuring that graduating high school students have obtained all the necessary skills, knowledge and abilities set forth by state and national standards. Unfortunately, this is often not the case. Because community colleges have an open access policy, many make their way into postsecondary education lacking in vital areas, such as reading and math. The community college must then step in and assume the responsibility of helping students become 'college ready' through various academic services and programs, one of the most popular being developmental education courses.

College readiness, particularly in Mathematics, is one of the greatest obstacles underprepared college students have to overcome. The 2015 Texas Public Higher Education Almanac reports that 76% of students were enrolled in developmental math, compared to 65% in developmental reading and 60% in developmental writing. North Carolina developmental course statistics indicate that 41% enroll in developmental math and 36% enroll in developmental reading and English courses (Carolina Journal, 2014). The Florida College System reported that in 2012-13, the percentage of students enrolled in *at least* one developmental course was 18.3%. Of these, math comprised 63.1% of all developmental course enrollments, writing accounted for 18.7% of all developmental course enrollments, and reading accounted for 18.1% (Florida College System, 2014). Out of a need to increase the number of students successfully completing college courses, the idea of developmental education became an initiative worth addressing in the postsecondary education sector.

There are various levels of developmental math offered by community colleges in Florida. These include the lower level (MAT 0018, Pre-Algebra or Developmental Math I) and the upper level (MAT 0028, Beginning Algebra or Developmental Math II). These are both non-credit earning courses that, once passed, lead directly into the first college-ready math course, MAT 1033, Intermediate Algebra. Once a student has passed Intermediate Algebra he or she would be allowed to enroll in a number of upper level math courses, most commonly, MAC 1105, College Algebra. MAC 1105 has significant importance because of its reputation as a 'gatekeeper' course. A 'gatekeeper' course has been defined in Florida as a course that if passed, is positively correlated with a successful academic outcome. A successful academic outcome is defined as the student receiving an A, B or C in a course.

In order to assess college readiness and the need for developmental education, students take either a readiness or a placement test. Common tests used include the American College Testing (ACT) test, Scholastic Aptitude Test (SAT), and the ACCUPLACER (CPT) tests. Florida Statute 1008.30 states that there must be a common placement test in place for public postsecondary education. In Florida, the ACCUPLACER test was the test of choice for postsecondary institutions in assessing college readiness. The ACCUPLACER test suite are tests that are "…designed to assist institutions in placing student in to appropriate courses" (2014).The use of the ACCUPLACER stemmed from Florida Statute 1008.30(1), which specifically states in part:

> The State Board of Education, in conjunction with the Board of Governors, shall develop and implement a common placement test for the purpose of assessing the basic computation and communication skills of students who intend to enter a degree program at any public postsecondary educational institution. Alternative assessments that may be accepted in lieu of the common placement test shall also be identified in rule.

Also of note, is Florida Rule 6A-10.0315, which establishes the required test scores for student taking placement tests, and Florida Rule 6A-10.0318, which lays out the curriculum and competencies required for students to be deemed college-ready. The link between developmental education and an appropriate placement test is a very important one, and will be explained in more detail in the next chapter.

Statement of the Problem

Nationally, states seek methods of gauging student learning and readiness. Primarily, this is done through standardized testing in order to ensure the student is learning basic concepts. In Texas, the Texas Assessment of Academic Skills (TAAS) test, and in Florida, the Florida Standards Assessments (FSA) test are both used to record benchmarks of learning for students in the K-12 curriculum. Oftentimes, there exists a rift between what states define as readiness competencies and what standardized tests measure as readiness competencies. As Hodara, Jaggars and Karp (2012) comment in a report on improving developmental education and placement, states will occasionally become dissatisfied with the existing tests and opt to find or create a test that will better fit the needs of the curriculum. Colleges in Oregon, Virginia and North Carolina have made moves to develop customized placement tests. Florida has also recognized the discrepancy between its postsecondary readiness competencies and the limitations of the information of what commonly used placement tests provide. Mattern and Packman (2009) express that “The main goal of placement testing is to enroll students in courses that are aptly challenging to their current knowledge level so as not to bore or frustrate, which can lower motivation to perform. For this process to work, placement testing policies need to be continuously reviewed and evaluated to ensure that students are in fact being placed into courses that will maximize the probability of their success” (p. 1). In 2010, the state of Florida implemented one of the nation’s first state-wide customized placement tests specifically geared toward measuring college readiness based on identified postsecondary readiness competencies (PRCs). The Postsecondary Education Readiness Test (PERT) replaced the commonly used ACCUPLACER, or CPT, exam when the contract for the ACCUPLACER ended and state administrators sought out other methods of addressing the readiness issue.

The development of the Postsecondary Education Readiness Test stemmed from the 2008 findings from a Go Higher Florida task force, which recommended, in part, that:

- The State Board of Education…should adopt a common definition of “college and career readiness” for Florida; and
- Develop/ adopt high school/ postsecondary assessment (s) which are clear in purpose and function.

Following these recommendations, the officials at the Florida Department of Education organized a faculty workshop with over 70 cross-sector English/language arts and math faculty, including high school teachers, Florida college and state university faculty (Florida College System, 2010). Officials then decided that the upcoming expiration of Florida’s contract with College Board for the ACCUPLACER was the perfect opportunity to begin undertaking the process of customizing its own placement test by assessing postsecondary readiness competencies (PRCs) and developing an assessment instrument. These PRCs were created by high school, college and university faculty, and were identified as skills that were critical to college readiness in entry-level reading, writing and math (Brown & Lancashire, 2013). English

1101 (Freshman Composition Skills I), and MAT 1033 (Intermediate Algebra), were the courses identified as those being the “first credit bearing course” in each subject, and thus the courses for which benchmarks or competencies, would be identified (Florida College System, 2010). In January 2010, McCann Associates were chosen as the vendor to support Florida in this endeavor.

In April 2010, the Florida Department of Education solicited the help of faculty from five Florida school districts, two private postsecondary institutions, nine universities and 24 colleges (Florida College System, 2010). This group of individuals was tasked with reviewing the state’s Postsecondary Readiness Competencies (PRCs) and drafting preliminary Common Core College and Career Readiness Standards. Once the PRCs were developed, they were used to create the pool of test items. In June 2010, Florida colleges administered the first round of approximately 10,000 PERT pilot exams, from which data were gathered to build the item bank. In August 2010, the same faculty group re-convened in order to review all the items for relevance with the PRCs and to make modifications to items where necessary. The test was formally launched in October 2010 and was in use by all Florida community colleges by the end of spring 2011.

The PERT is an untimed, competency based computer-based test, which assesses competencies in three areas: Math, Reading and Writing. PERT determines student readiness for entry-level college credit courses, MAT 1033 (Intermediate Algebra) and ENC 1101 (Freshman Composition I). The PERT presents 30 questions per section for a total of 90 questions, and a student’s answers determine the difficulty of the next question. For each section, student scores are based on their responses to 25 operational items, and each section also includes 5 field test items that allow the test developer to monitor the test and continuously enhance the test bank. Students do not know which items are operational, and which are the field test items (Appendix A represents the PERT subject areas). The PERT also has a diagnostic component, which was introduced in fall 2011, but is used only on a voluntary basis.

Its scaled test scores range from 50 to 150, with course placement score ranges as listed in Table 1.

Table 1

Course Placement Score Ranges for Math, Reading and Writing

	Mathematics	Reading	Writing
Lower Level Developmental Education	50-95	50-83	50-89
Higher Level Developmental Education	96-113	84-105	90-102
College Ready (MAT 1033/ ENC 1101)	**114*-122**	**106*-150**	**103*-150**
College Algebra (MAC 1105) or Higher	123-150	---	---

*College ready cut score

The interim cut scores developed for the PERT were reviewed by both the Florida Department of Education (FLDOE) and McCann Associates. These cut scores were identified for placement into both levels of developmental education, college credit classes, and College Algebra (MAC 1105). It was decided by both parties that in order to minimize any drastic changes to course enrollment, the cut scores should mimic the ones previously set forth by the ACCUPLACER. The FLDOE used data from 2006-07, 2007-08, and 2008-09 to define the percentage of students placed into various developmental courses based on their ACCUPLACER results. This process involved selecting all students in the Florida database who were first-time-in-college (FTIC) students (to ensure they were FTIC's, researchers removed any students with course records in the previous year and those enrolled in dual enrollment courses), selecting the ACCUPLACER score for math, matching the math score to the student, and only keeping those with an ACCUPLACER score. Students were deemed 'College Ready' and 'Not College Ready' based on the minimum ACCUPLACER scores previously established (http://www.fldoe.org/core/fileparse.php/7724/urlt/0072380-pert.pdf). Those distributions were then matched to the expected distributions of the PERT in order to allow the interim cut score to mirror the current placement rates (http://www.fldoe.org/core/fileparse.php/7724/urlt/0072380-pert.pdf). An informational newsletter by the Florida College System (2010) notes "…each college is expected to use the interim cut scores for placement purposes until the Department has enough PERT performance and course outcome data to set final cut scores in State Board rule" . It is also stated that this may take approximately a year to acquire said data, however, as of the date of this publication, no final cut scores had been identified and publicly distributed. The test was fully implemented in the fall of 2011 when all 28 colleges were mandated to use this test for placement.

Once a student has applied to a postsecondary institution in Florida, he or she must then apply for financial aid (if needed), submit transcripts, complete a student orientation, and then contact an advisor. The advisor reviews the student's information and determines whether the student will need to take a placement test based on the cutoff score of a previously taken placement tests (i.e. SAT, ACT), or (either) the lack of a placement test score. Students encouraged to take the PERT exam do so at one of the college's campus testing centers, and the final score is put into the student database once the test is complete. The student is then placed in courses based on his or her score and where it falls within the range of cut scores for that particular course. Students meeting or exceeding the cut score for the math section of the test are placed in either Intermediate Algebra (MAT 1033) or College Algebra (MAC1105). Students not meeting the cut score are placed in either the upper level developmental math course or the lower level developmental math course. The present study focuses on students who took the PERT test in fall of 2012 and subsequently enrolled in developmental courses (MAT 0018 or MAT 0028) as a result of their score. The fall 2012 term was used as a starting point in order to find enough students who had taken any initial developmental course and subsequently enrolled and completed either MAT 1033 (Intermediate Algebra) or MAC 1105 (College Algebra). Figure 1 represents the path of the math curriculum.

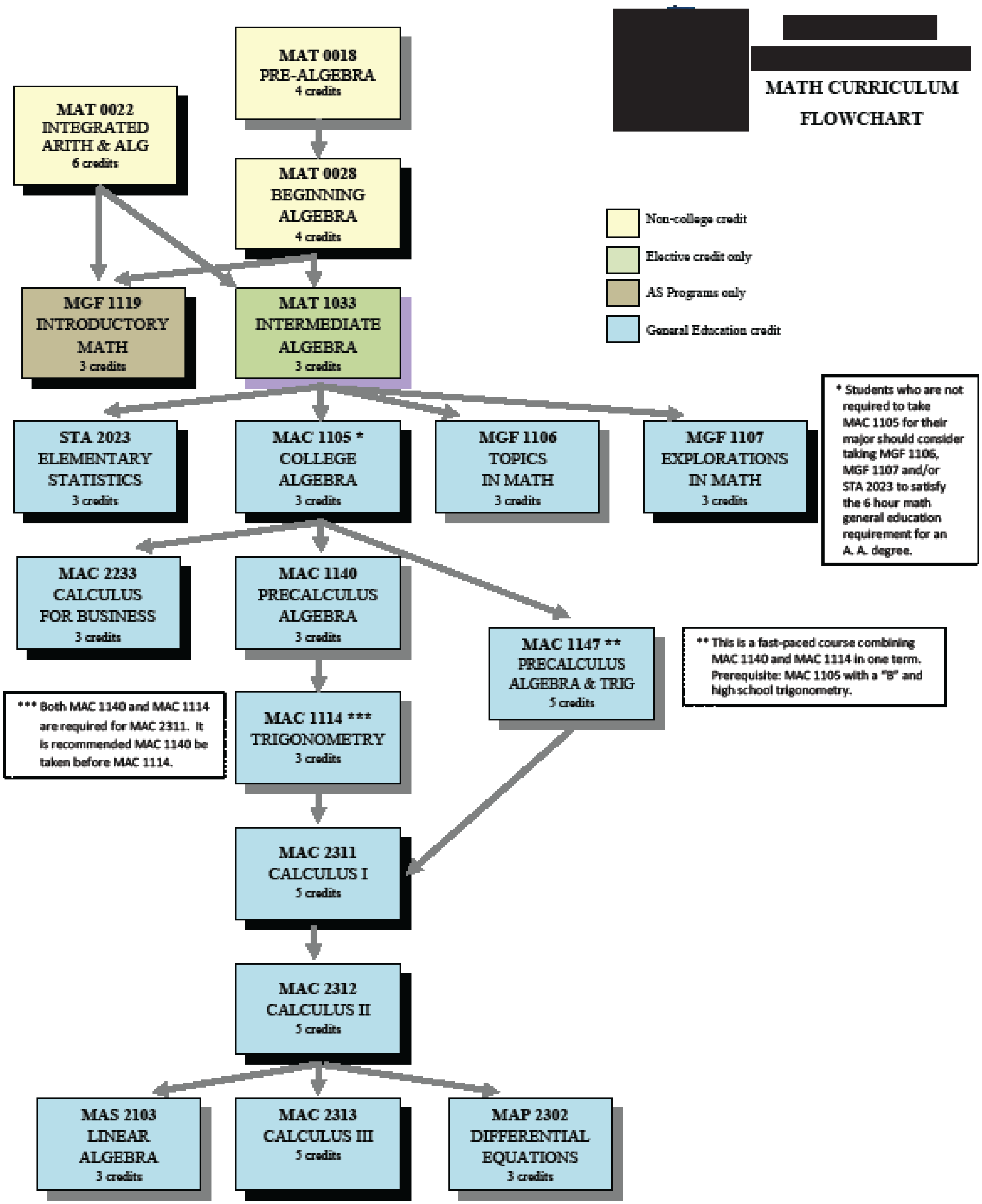


Figure 1. College Math Curriculum Flowchart

Currently, no published validity information on the PERT has been publicly distributed. According to the *Standards for Educational and Psychological Testing* (2014) when a new measure is developed, it is the responsibility of both the test developer and the test user to address issues of validation and to ensure that validity evidence is obtained (p.11). There are five sources of validity evidence:

- Evidence based on test content;
- Evidence based on response processes;
- Evidence based on internal structure;
- Evidence based on relations to other variables; and
- Evidence based on consequences of testing.

This research provides an opportunity to see how accurately the test places students in both developmental and credit courses, relations to other variables, and to review whether the student was successful in these courses. It is important to note that a number of factors could affect student success in courses. For instance, the test could be fine, but the cut scores could be inappropriate. Also, the test and the cut scores could be adequate, but variability could exist among the instructors. In addition to reviewing PERT scores, it is also pertinent to look at these relations across instructors and across demographic information of the students. The PERT is a fairly new test, and as of this publication, there has yet to be any validity information distributed. An analysis of the test provided an opportunity to see how the test is working and to see how well students are performing.

Purpose of the Study

The purpose of this study was to begin the collection of evidence that supports the validity of the PERT scores. This research seeks to investigate the predictive validity of the Postsecondary Education Readiness Test (PERT) in relation to student final course grade in the following courses: MAT 0018 (Developmental Math I), MAT 0028 (Developmental Math II), MAT 1033 (Intermediate Algebra) and MAC 1105 (College Algebra). The focus is on math due to state data showing that a higher percentage of students requires developmental math when compared to developmental reading and developmental writing. This study seeks to determine the significance of student and course section variables on student final course grade in Intermediate Algebra and College Algebra. This research also looks at differential predictive validity in terms of how the relationships between the PERT scores and course outcomes varied as a function of student gender and/or student race.

Several previous studies that have examined various placement tests for differential predictive validity (i.e., “when a test systematically over- or under-predicts the criterion variable by the subgroups”) have indeed shown that evidence of differential predictive exists. For example, Noble, Crouse et al. (1996) conducted a study of differential prediction in course placement for ethnic and gender groups based on ACT scores and subject area grade averages and found that both the ACT and SAT slightly over predicted course success for blacks and males when compared to whites and females. Analyzing differential predictive validity for the PERT is important in order to determine whether there is, in fact, evidence of predictive validity.

This study also looks at the variation in these predictive relationships across course sections at one community college using multilevel modeling techniques. The use of multilevel modeling was warranted given that students are nested within course sections. Analyses were based on archival data collected in the fall 2012. Data from this period were used due to the large number of students that are enrolled in each course to be analyzed. Because the PERT is a new assessment instrument, it is anticipated that results from this study would add to the body of knowledge pertaining to this test.

Research Questions

Research questions that were explored in this study include:

1. To what extent does the predictive relationship of the PERT with final grades in Developmental Math I (MAT 0018) and Developmental Math II (MAT 0028) vary based on student gender?
2. To what extent does the predictive relationship of the PERT with final grades in Developmental Math I (MAT 0018) and Developmental Math II (MAT 0028) vary based on student race/ethnicity?
3. To what extent does the predictive relationship of the PERT with final grades in Intermediate Algebra (MAT 1033) and College Algebra (MAC 1105) vary based on student gender?
4. To what extent does the predictive relationship of the PERT with final grades in Intermediate Algebra (MAT 1033) and College Algebra (MAC 1105) vary based on race/ethnicity?
5. What combined student and course section-level variables (i.e., first-time-in-college status, student race/ethnicity, student gender, student enrollment status, student age, instructor employment status, course time of day, course delivery method) are significant predictors of student final course grade in Developmental Math I (MAT 0018)? In Developmental Math II (MAT 0028)?
6. What combined student and course section-level variables (i.e., first-time-in-college status, student race/ethnicity, student gender, student enrollment status, student age, instructor employment status, course time of day, course delivery method) are significant predictors of student final course grade in Intermediate Algebra (MAT 1033)? In College Algebra (MAC 1105)?

Limitations of the Study

This study may be limited in that within this particular community college, there is no one particular methodology in the way the classes are delivered. Each instructor is given the academic freedom to teach the class as he or she wishes. They all, however, must ensure that the student learning outcomes for that particular class are met. Class delivery methods, tests and assignment of grades are factors that are determined solely by the instructor. Additional studies at other community colleges would assist in evaluating generalizability of the results of the present study.

Another limitation is that there are no item-level data available, therefore item-level analyses were not able to be performed. Finally, due to the fact that the existing cut scores for the PERT were 'borrowed' from the previous placement test, the accuracy of these cut scores could be a limiting factor. Uncertainty as to the accuracy of proper placement by this test could have an effect on the results of the predictive validity analyses.

Definition of Terms

College Readiness. Operationally defined as "level of preparation a student needs in order to enroll and succeed—without remediation—in a credit-bearing general education course at a postsecondary institution that offers a baccalaureate degree or transfer to a baccalaureate program (Conley, 2011).

Cut Score. Score that determines the minimum performance level at which a student passes an assessment test.

Developmental Education. Instruction through which a high school graduate who applies for any college credit program may attain the communication and computation skills necessary to successfully complete college credit instruction. (May also be referred to in this study as '*Remedial Education*'.)

First Time in College (FTIC). Indicates a student who has entered any college for the first time and does not have an existing college transcript.

Florida College System (FCS). The collection of the 28 public colleges in Florida which offer the Associate's degree.

Gatekeeper Course. A course for which successful completion positively correlates with academically successful outcomes.

Placement Test. Test given to students that determines the skill/ ability level for a specific subject (i.e., mathematics, English, or writing).

Postsecondary Education. Education beyond high school that happens at the college or university level.

Reliability. Achieved when a test-taker's scores on a test remain consistent throughout repeated administrations of the same test or an alternate test form (Crocker & Algina, 1986).

Standard Setting. Refers to the process used to establish the cut scores for an assessment test.

Validity. Determined when a test measures what it purports to measure. Validity gives support to the assumption that a test or measure really works. According to the *Standards for Educational and Psychological Testing*, validity "refers to the degree to which evidence and theory support the interpretations of the test" (American Educational Research Association, American Psychological Association, National Council on Measurement in Education, 2014, p.11)

Importance of the Study

Given recent legislative changes, assessing how students are faring in developmental and college ready courses will have important implications for the future of developmental education in Florida. In general, it is good practice to review placement scores as a matter of institutional policy. Standard 12.13 of the *Standards for Educational and Psychological Testing* (2014) states: "When test scores are intended to be used as part of the process for making decisions for educational placement, promotion, or implementation of prescribed educational plans, empirical evidence documenting the relationship among particular test scores, the instructional programs, and desired student outcome should be provided. When adequate empirical evidence is not available, users should be cautioned to weigh the test results accordingly in light of other relevant information about the student" (p. 199). In the case of determining placement into developmental courses, evidence that supports the degree to which student placement rates are appropriate is an extremely important component of the educational assessment process. An appropriate score means that students are successfully completing the course in which they are placed,

Organization of the Study

Chapter one introduces the study. It includes the statement of the problem, purpose of the study and the research questions, definitions of the terms used within the narrative, the importance of the study, and the organization of the study. Chapter two consists of a review of the literature to include a discussion of college readiness, developmental education policies in the United States as a whole, and in individual states; tests used to assess college readiness; and methods in establishing cut scores in educational tests.

Chapter three is the methods section, which describes the participants, the process of data collection, the statistical analyses performed, and a summary of the chapter. Chapter four presents the results of the research study, along with an explanation of the findings. Finally, chapter five is a discussion of the results, implications of the study, and recommendations for future research.

CHAPTER TWO

REVIEW OF LITERATURE

The sections in the literature review will include background information on the community college and on developmental education in the community college; the variety of different tests used to assess readiness in Florida; the development of the Postsecondary Education Readiness Test (PERT) as an assessment instrument; the importance of establishing both validity and standard setting policies in educational placement tests; and a discussion of multilevel modeling and how it relates to this study.

Background Information

According to Thomas Bailey (2008), "…a majority of community college students arrive unprepared to engage effectively in the core function of the college" (p. 1). This statement and its implications for the academically unprepared student can have a potentially devastating effect. Being college ready means that a student has the knowledge, skills and abilities to be successful in college. Conley (2011) defines success as "…completing entry-level courses at a level of understanding and proficiency that makes it possible for the student to consider taking the next course in the sequence or the next level of course in the subject area" (p. 1).

With such a diverse population of students entering postsecondary education, college counselors and student advisors must ensure that placement test scores used are reliable and valid. The ACT and the SAT tests are both standardized college entrance exams that give a general assessment of student college readiness. If the minimum scores for these tests are not met, a placement test may be given in order to ascertain the level of remedial education the student would need. Remedial education seeks to improve students' knowledge, skills and abilities so that they may be better prepared to enroll and succeed in higher level college credit courses.

Developmental education as a process has long been studied. From the inception of college preparatory programs in colleges in the late 1800's, educational professionals have debated what constitutes college readiness, what skill sets students need to possess, and valid ways with which to measure college readiness. Casazza (1999) notes that the idea of developmental education is a comprehensive process that requires looking at the student holistically. This view of developmental education as a process has allowed higher education professionals to hypothesize how to best address this idea. How to best define what developmental education is, how to measure it, and how to improve it are all matters that have arisen in the higher education sector. Through the process of reviewing postsecondary readiness competencies, university and college faculty and staff members in Florida engaged in a conscious effort to make a change in its K-12 and postsecondary curriculum, and in the way student readiness is measured. Developing the PERT involved a set of practices that required these individuals to assess the current state of the K-12 and postsecondary curriculum, evaluate the accuracy of the readiness competencies, and assist in the development of a measure with which to assess these readiness capabilities.

Developmental Education in the Community College

Bailey and Cho (2010, p. 1) noted that "…addressing the needs of developmental students is perhaps the most difficult and most important problem facing community colleges" (2010). Often, there is a negative connotation associated with taking developmental education courses; however these courses can be beneficial to a student who is not academically ready for the college curriculum. In the 1950s and 60s, Clark (1960) noted that a phenomenon known as 'cooling out' is used to describe students' lowering their academic expectations and accepting their limitations. This is especially evident when it comes to community colleges because of its 'open-door, access for all' policy. Years later, Deil-Amen and Rosenbaum (2002) recognized the stigma attached to this ‘cooling out’ occurrence, and studied methods on how to encourage students to maintain their goals and aspirations in spite of their lack of academic skills when entering community college. In their study, four key conditions- (1) high school students’ college aspirations, (2) college-for-all counseling, (3) the large number of students in remedial courses, and (4) the association between the number of remedial courses and college dropout - were identified as areas community colleges must address when managing students whose ambitions may conflict with their academic readiness. The first condition highlights the fact that some high school students don’t recognize the relationship between school achievement and educational aspirations. The article points out that some students may wait until college to put effort into learning, not realizing that high school grades can affect their college careers. The second condition points out the evolving role of guidance counselors in high school—from a view that would focus on “keeping it real” with students—meaning guiding them towards careers or jobs that they’re already good at-- to one that adopts a “college-for-all” approach. This “everyone can go to college” mentality can be the downfall to some students who are not ready for stringent college curriculums. The third condition is pretty self-explanatory, in that it recognizes that there are a growing number of community college students who are simply not ready for college-level coursework; and finally, the last condition, which explains that there is a correlation between the number of remedial courses taken and the risk of dropping out of college.

Some of the most significant initiatives addressing developmental education have been sponsored by the Lumina Foundation and Achieving the Dream (ATD). Achieving the Dream was developed in 2004 by the Lumina Foundation in order to improve student success in higher education, and its Developmental Education Initiative specifically, was aimed to assist 15 colleges in six states on how to make developmental education more streamlined (Quint, Jaggars et al., http://www.mdcinc.org/projects/developmental-education-initiative). To date, the

ATD network includes over 200 two- and four-year colleges from 36 states that provide data, which inform policy decision-making.

The Developmental Education Initiative had as its primary goal, finding methods to assist students in moving through developmental courses quickly or eliminating the need to take them at all (Quint, Jaggars et al., 2013). A number of experimental strategies were employed in order to add to the body of knowledge when considering the challenges of developmental education. These particular strategies involved implementing study skills courses, tutoring, advising, placement test preparation, and employing methods to make instruction more relevant and engaging (Quint, Jaggars et al., 2013).

A number of studies have explored developmental education, including its method of delivery, sequence of courses and even the establishment of student support programs that assist the student in successful completion of the developmental education curriculum (Bailey, Jeong, & Cho, 2009; Le, Rogers, & Santos, 2011; Perry, Bahr, Rosin & Woodward, 2010). These studies seemingly seek to ensure that measuring student learning outcomes is consistent among institutions, and even at the state level.

The assessment of developmental education has gone through several iterations in the state of Florida. Beginning in 1984 with the adoption of Florida Legislative Rule 6A-10.0315, Florida community colleges were required to address the issue of developmental testing, placement in developmental courses, and administration to all FTIC students using the Florida College Entry-Level Placement Test, also known as the CPT. In 1997, all Florida community college were mandated by law to administer the College Placement Test (CPT). Students who failed to meet the standardized cut-off were to be placed in developmental courses. The rule was modified in August 2012 to specifically state that "…first-time-in-college degree seeking students and students who have not met college level competency either through the completion of developmental education requirements in the Florida College System or have not been awarded credit for college-level coursework in the area of deficiency shall be tested for reading, writing, and mathematics proficiency prior to the completion of initial registration, using the Florida Postsecondary Education Readiness Test" (Florida Rule 6A-10.0315). Prior to implementing the Postsecondary Education Readiness Test, Florida had used College Board's ACCUPLACER exam since 1996 as the state mandated placement exam. The ACCUPLACER will be discussed in more detail in a later section.

As previously stated, in 2012-13, the percentage of students enrolling in any developmental course in Florida was 18.3%. Nationally, the percentage of students enrolling in developmental courses in community colleges is slightly higher than that of Florida. texasewell, Lavin et al. (2006) quote in their research of the National Education Longitudinal Study, also known as NELS:88, that 58% of the students participating in NELS:88 took remedial courses in community college. Additionally, a more recent 2011-12 National Postsecondary Student Aid Study (NPSAS:12) reported that the percentage of Associate's Degree students who reported ever having taken a remedial course while in college was 39.5%. Possibly in an effort to quell both the number of students enrolling in developmental education courses and those failing in those courses, in July 2013, the state of Florida approved Senate Bill 1720, which, among other things, modified the implementation of developmental education. It states that two groups of students are exempt from taking a common placement test or enrolling in developmental education. They include:

- Any student who entered 9th grade in a Florida public school in 2003-04 or later, and who earned a Florida high school diploma; or
- Any student who is an active duty member of the United States Armed Services.

What this bill theoretically assumes is that any student meeting these criteria is deemed college ready. This change was implemented in October 2013, so longitudinal data are not available to assess the accuracy of this assumption.

Of consequence are the numbers of students that are successfully completing developmental courses. Bailey, Jeong and Cho (2009) note that in a sample of Achieving the

Dream community college students, only 33% of students referred to any developmental math complete all of the developmental sequence. Out of a need to increase the number of students successfully completing college courses, the idea of a more streamlined developmental education curriculum became an initiative worth addressing in the postsecondary education sector.

Assessment of College Skills and College Readiness

David Conley (2011) operationally defined college readiness as "…the level of preparation a student needs in order to enroll and succeed—without remediation—in a credit-bearing general education course at a postsecondary institution that offers a baccalaureate degree or transfer to a baccalaureate program" (p. 1). Palomba and Banta (1999) offer the following as a definition of assessment: "…the systematic collection, review, and use of information about educational programs undertaken for the purpose of improving student learning and development" (p. 4). Both college readiness and the methods with which to assess it have been at the forefront of many higher education institutions, particularly in community colleges.

Conley (2011) goes on to further identify the four facets of college readiness, including (1) key cognitive strategies; (2) academic knowledge and skills; (3) academic behaviors, and (4) contextual skills and awareness. Key cognitive strategies describe the "…intelligent behaviors necessary for college readiness" (p. 9). The purpose of highlighting these strategies is to emphasize that these behaviors need to be automatic and need to happen without thinking. Academic knowledge and skills consist of a combination of cognitive strategies and content knowledge, which is mastered through applying a more extensive set of cognitive skills. Academic behaviors entail "…behaviors that reflect greater self-awareness, self-monitoring, and self-control on the part of students in relation to a series of processes and behaviors necessary for academic success" (Conley, 2011, p.12). The final facet, contextual skills and awareness, pertains to students' ability to use their own knowledge and awareness to cope and adapt to their surroundings, even if the surroundings are vastly different from the one they are accustomed to.

The recognition that educational skills needed to be assessed in higher education began in the early 1800’s when the issue was discussed in an article called the Yale Report. While this report focused on setting a specific curriculum for the students entering the university, its overarching theme was to provide the proper program of study for students. The lack of college readiness issue persisted in the late 1800’s when students entering Harvard University were not prepared to meet the criteria of college-level courses. This resulted in Harvard modifying its curriculum to better meet the needs of students showing deficits in reading, writing and math (Cassaza, 1999). The issue of the unprepared college student was a pervasive problem throughout the United States, as evidenced by multiple state legislatures addressing developmental or remedial education.

By its very existence, the community college is an institution that promotes access for all students, oftentimes, regardless of the academic preparedness of the student. Therefore, testing students in order to assess their knowledge of the skills and abilities, deemed necessary for success in the postsecondary setting, is a priority among many community colleges. Perin (2006) questioned whether the community college can adequately conjoin the concept of access for all students with the requirement of the community college to provide a quality education to all.

In a policy brief authored by Michael Lawrence Collins for Achieving the Dream, it was documented that in 2008, 27 states require community colleges to assess students for developmental placement. Of those, 21 required a specific assessment test or tests to be used and within that group, 19 states required the use of cut scores in determining placement into developmental courses (Collins, 2008). In 2012, a national survey administered by the National Assessment Governing Board and Westat reported that of the 3,650 institutions in the study, 71% used a placement test specifically for math. Among the 970 public 2-year institutions, 100% used a placement test for math.

In Florida, the concept of addressing readiness has been a task undertaken by a number of college administrators, faculty, and organizations alike. In 2004, the Florida community college system elected to participate in the Achieving the Dream initiative. This national initiative, sponsored by the Lumina Foundation, has at the crux of its mission the goal of increasing community college student completion rates in hopes of improving the individual's long-term economic situation. Achieving the Dream boasts a network of over 200 higher education institutions across 34 states, all with the ultimate goal of improving student readiness of at-risk students, particularly in community colleges.

Initially, the question of whether or not community college students were prepared to enter and succeed in college stemmed from data from a 2004 review of Florida's K-20 longitudinal data that presented statistics regarding the completion rates of Florida's community college students. At the time, it was revealed that 41% of full-time FTIC students were obtaining an Associate's degree or certificate within three years. An investigation of the Florida Comprehensive Assessment Test (FCAT), a high school standardized assessment test to assess college readiness, revealed that some students who were passing this test were not passing college placement tests (Burdman, 2011). This led to the question as to whether the FCAT was an accurate measure of readiness, whether the college placement tests were an accurate measure of readiness, whether neither test was appropriately measuring student readiness, or if a combination of placement test, student GPA and/or student grades was the better measure of student postsecondary readiness and success.

Using Grades as a Measure of Success

The use of grades to determine how well a student understands a particular subject has long been a disputable topic, however some educators believe that the use of grades is a valid measure of student achievement. The current study looked at grades as a criterion measure of final course grade in MAT 0018, Developmental Math I; MAT 0028, Developmental Math II; MAT 1033, Intermediate Algebra; and in MAC 1105, College Algebra. In the 1950's, Bendig (1953) acknowledged the importance of researching the reliability of course grades. He noted two reasons for this exploration being vital: 1) these academic ratings, which would become part of the student's permanent record, could directly affect future education and job opportunities, and 2) the usage of grades as predictor variables in research. A number of research studies have been performed in order to evaluate the degree to which grades relate to postsecondary success (Camara & Millsap 1998; Camara & Echternacht, 2000; Patterson & Mattern, 2013; Patterson, Mattern, & Kobrin, 2009). These studies use correlational statistics in order to more accurately define the relationship between grades and their relation to certain placement tests.

Allen (2005) studied the validity of grades when measuring academic achievement, and found that the primary focus should be in educating teachers in assessment and measurement principles. The validity of grades is important when attempting to gauge student learning, and therefore may have a significant impact on student readiness. Allen noted that there are three primary reasons that professors are uncomfortable using grades to determine student academic achievement: (1) opinions of what constitutes good academic achievement varies from teacher to teacher, and are therefore subjective; and (2) sometimes, a student's body of work in a class cannot be accurately summed up in a single academic mark (i.e., grade). The teacher must determine if the grade is an honest reflection of the student; and (3) due to a combination of the two previous reasons, teachers will often assign grades that are inconsistent. If grades are used accurately and are based on a reliable measurement approach, they can be a very powerful tool in identifying student success.

Tests to Assess College Readiness in Florida

In Florida, there are currently five tests that may be used to assess college readiness and for course placement in postsecondary education: The Florida Standards Assessment (FSA) (Formerly the Florida Comprehensive Assessment Test or FCAT), the

ACCUPLACER/CPT, the ACT (formerly America College Testing), the Scholastic Aptitude Test (SAT) and the current state-mandated test, the Postsecondary Education Readiness Test (PERT). The Postsecondary Education Readiness Test is a computer adaptive exam, which is comprised of both a placement component as well as a diagnostic component. While these are the most commonly used, there are other tests that are frequently utilized, sometimes in conjunction with the aforementioned. In addition, colleges may utilize a multiple measures approach in order to assess college readiness, or as in the case of Florida, a customized test may be developed. Table 2 shows the state-approved cut scores for four of these tests. What follows is a brief description of each test. Appendix B details the percentile ranks for these various tests.

Table 2

Florida College-Ready State-Approved Tests and Their Cut Scores

	Mathematics (MAT 1033)	Reading (ENC 1101)	Writing (ENC 1101)
Grade 10 FCAT 2.0	---	262	---
ACT	19	19	17
SAT-I, The College Board	440	440 (Verbal)	
ACCUPLACER, The College Board	72	83	83

Florida Comprehensive Assessment Test (FCAT)

The FCAT began in 1998 in order "…to increase student achievement by implementing higher standards" (FCAT Briefing Book 2007, p. 13, http://www.fldoe.org/accountability/assessments/k-12-student-assessment/history-of-fls-statewide-assessment/fcat/). It is an assessment test given to students in grades 3 through 11. It is intended to measure students' achievement of the Sunshine State Standards in math, reading, science, and writing. These standards were established in 1996 in order to bring attention to educational accountability with the K-12 sector.

Students taking the FCAT are rated by levels 1 through 5, with Level 1 being ‘low’ and Level 5 being ‘high’. The FCAT reports two types of scale scores: a grade-level scale score and developmental scale score. On the scale score per grade level, a student can earn between 100 and 500 points. On the developmental scale, which encompasses grades 3 through 11, students can earn from 0 to 3000 points (FCAT Briefing Book, 2007).

In March 2014, the Florida Department of Education commissioner announced that for the 2014-15 school year, the FCAT would be replaced with the Florida Standards Assessment (FSA) due to the “…more rigorous standards in place to help Florida students succeed” (FLDOE Press Office Memo, 2014). The FCAT test was considered to “…no longer serve the purpose of measuring student progress and achievement” (2014).

American College Testing (ACT) and the Scholastic Aptitude Test (SAT)

The American College Testing (ACT) test is a multiple-choice assessment test that is given to high school students in order to evaluate college readiness in the areas of English, Math, Reading, and Science. There is also an optional writing test. Three scores are reported in the English test: a total test score based on 75 items, a subscore in Usage/Mechanics, and a subscore in Rhetorical Skills. In Math, there are four scores reported: a total score based on 60 items, a subscore in Pre-Algebra/Elementary Algebra, a subscore for Intermediate Algebra/Coordinate Geometry, and a subscore in Plane Geometry/Trigonometry (ACT Technical Manual, 2007). The number of items and the time allotted varies for each subject matter. The composite score for the ACT ranges from 1 to 36. A number of studies have been performed for the ACT, including those that assess its validity for college readiness benchmarks. In 2007, 76,122 students from 92 colleges for English Composition, 33,803 students from 85 colleges for College Algebra, and 14,136 students from 31 colleges for Biology participated in a study geared towards establishing readiness benchmarks for common first-year college courses based on ACT scores. The study also included five different courses for the Social Science analyses: history, psychology, sociology, political science, and economics. The results for the social science courses were based on 53,705 students from 45 colleges (ACT Technical Manual, 2007). Results from this study indicate that students who score 18 or higher for English, 22 or higher for College Algebra, 21 or higher for Social Science, and 24 or higher for Biology give a college student a 50% chance of earning a B or better in those respective college benchmarks.

The Scholastic Aptitude Test (SAT) is a curriculum-based, high-stakes test developed by The College Board that measures academic achievement as students prepare for postsecondary and career opportunities (SAT Educator's Handbook, 2012-13). There are three sections of the test: Critical Reading, Math and Writing. The SAT includes nine subsections, including a 25-minute essay, with each subsection timed separately. Students receive a score on each of the three sections, with the scores reported on a 200 to 800 point scale in 10-point increments.

ACCUPLACER

ACCUPLACER was developed by The College Board, and is a test that primarily assesses whether a student is prepared for college by using a computer-adaptive approach to "tailor" the test for each student (College Board, 2014). In addition to the computer-adaptive component, The ACCUPLACER has a written essay exam. The test consists of topics that measure students' abilities in various areas by administering a set number of questions per area: arithmetic (17 items), elementary algebra (12 items), college-level math (20 items), reading comprehension (20 items), sentence skills (20 items), English as a second language, writing skills, and computer skills. The ACCUPLACER also has a diagnostic component for reading comprehension, sentence skills, arithmetic and elementary algebra. ACCUPLACER scores range from 20 to 120 for multiple choice tests. The version of ACCUPLACER that Florida used consisted of 52 items (20 items each for reading and writing, and 12 for math). A listing of

ACCUPLACER math content areas can be seen in Appendix C (College Board ACCUPLACER Program Manual, 2014). The ACCUPLACER also has a diagnostic component for reading comprehension, sentence skills, arithmetic and elementary algebra. ACCUPLACER scores range from 20 to 120 for multiple-choice tests.

The ACCUPLACER was implemented in Florida by the legislature on June 30, 1997. Common cut scores were established state-wide, which set forth the “college-ready” cut-scores for the state. Currently, the ACCUPLACER is still used as a method to determine college readiness in the postsecondary sector, however, it is no longer the mandated test for the state of Florida.

The College Board makes recommendations to schools on setting cut scores and follows guidelines set forth by the AERA Standards as well as best practices in the field. These recommendations are available in its ACCUPLACER program manual (2014). In a meta-analysis conducted by Mattern and Packman (2009), a collection of ACCUPLACER validity studies were evaluated in order to assess the appropriateness of the placement policies. The studies came from 17 institutions, including 14 community colleges. These studies reviewed the validity of ACCUPLACER test for math related tests (i.e., Arithmetic, Elementary Algebra, and College-Level Math) in courses from basic mathematics to precalculus. For verbal related tests (Reading Comprehension, Sentence Skills, and Writing), institutions primarily reviewed composition or reading courses. In each placement report, the following were provided:

- A correlation between the ACCUPLACER test score and course success;
- The percentage of students accurately placed in a given course; and
- The probability of success in a given course given different ACCUPLACER scores.

In this research study, course success was operationally defined in two ways; as (1) a grade of “B” or higher in the course and (2) a grade of “C” or higher in the course. The percentage of students placed accurately was determined by that student’s success in the course. The results of the test indicated that the ACCUPLACER did, in fact, "…support a moderate-to-strong relationship between test scores and subsequent course performance" (2009). The meta-analysis on the ACCUPLACER will be discussed in more detail in a later section.

While the ACCUPLACER is still a valid test for evaluating college readiness, the move within the state of Florida to develop a new set of postsecondary readiness competencies (PRCs) and aligning these competencies with the Florida curriculum prompted the opportunity for Florida to set new ground with a customized placement test aimed directly at assessing these PRCs. Therefore, the PERT was designed to meet this need.

Validity of Placement Tests

Hughes and Clayton (2011) posed an important question in a report discussing developmental assessment in the community college: *Do placement tests predict future performance?* Their research noted that in some previous studies of predictive validity in placement tests, the mere use of correlation coefficients may be flawed. This could be due, in part, to the issue of restriction of range, which may decrease the coefficient. Weber (2001) defines restriction of range as occurring when "…design or circumstances abbreviate the values of one of both variables being correlated" (p. 4). In the present study, range restriction could be an issue simply because the scores being reviewed fall within a certain range (i.e., students in lower level math have scores PERT scores between 50 and 95, upper level math have scores between 96 and 113, etc.). A study of this manner would effectively restrict the results because the researcher is only assessing within a limited range of scores. This phenomenon is known as explicit selection, and is defined by Crocker and Algina (1986) as occurring when "the test being validated is used for selection purposes before its validity has been established" (p. 227).

With the development of this new placement test (PERT) there is a need to evaluate the psychometric properties of this measure to insure that the scores are reliable and valid for the placement purposes for which the scores are being used. The *Standards for Educational and Psychological Testing* (2014) state that when a new measure is developed, it is the responsibility of both the test developer and the test user to address issues of validation and to ensure that validity evidence is obtained (p.12). There are five sources of validity evidence: (a) test content; (b) response processes; (c) internal structure; (d) relations to other variables; and (e) consequences of testing. More specifically, Standard 12.13 of the *Standards for Educational and Psychological Testing* (2014) states: “When test scores are intended to be used as part of the process for making decisions for educational placement, promotion, implementation of individualized educational programs…, then empirical evidence documenting the relationship among particular test scores, the instructional programs, and desired student outcome should be provided. When adequate empirical evidence is not available, users should be cautioned to weigh the test results accordingly in light of other relevant information about the student” (p. 199). In the case of determining placement into developmental courses, evidence is an extremely important component of the standard setting process. It is required in a study such as this, when empirical evidence is needed in order to ensure that students are being placed correctly in different math courses. The relationship between student scores on placement tests and outcomes in the courses in which they are placed must be monitored on a regular basis so that the validity of the placement test can be established and maintained.

Validating educational tests involves assessing one or more types of validity. These types, recognized in the field of measurement as criterion-related validity (encompassing concurrent and predictive validities), content validity, and construct validity, are sometimes merged together in various permutations. Messick (1989) defined validity as “…an overall evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of interpretations and actions on the basis of test scores or other modes of assessment” (p. 21).

Standard 1.18 of the *Standards for Educational and Psychological Testing* (2014) states: “When it is asserted that a certain level of test performance predicts adequate or inadequate criterion performance, information about the levels of criterion performance associated with given levels of test scores should be provided". The standard also states "regression coefficients are more useful than correlation coefficients…when describing patterns of association between tests and other variables" (p. 28). More specifically, Standard 1.19 states, "If test scores are used in conjunction with other quantifiable variables to predict some outcome or criterion, regression (or equivalent) analyses should include those additional relevant variables along with the test scores". In the current study, I am examining how various factors such as student gender, race/ethnicity, PERT score, and instructor variables may relate to students' final course grade in college-level courses (Intermediate Algebra and College Algebra).

Sawyer (2007) reviewed how the users of test scores could leverage those scores more effectively within the educational sector. As an example, he referenced a common initiative among colleges and universities: that of attempting to predict academic success based on specific student variables. In this particular study, he identified the selection variables as being the average of the college preparatory course grades a student took in high school (HSA) and the student's ACT composite score (ACT-C). The outcome variables of the study were academic success (a dichotomized variable) and GPA. This study defined academic success by a student completing the first year with a 2.0 (C) or higher, or 3.0 (B) or higher (2007). The type of methodology used to analyze this study was logistic regression, which provided evidence that the two selection variables may suggest "incremental predictive validity" (Sawyer, 2007).

There are a number of dissertations and research studies alike that have analyzed the predictive nature of specific placement tests. As previously mentioned, Mattern and Packman (2009) conducted a meta-analysis of the predictive validity of the ACCUPLACER placement scores. Their report reviewed 47 studies conducted at 17 institutions, 14 of which were community colleges between the years of 2001 and 2006. The researchers examined the validity of the ACCUPLACER for a number of courses, including math, reading and writing. ACCUPLACER used an overall cut value of 50 percent because "…this value maximizes the percentage of cases correctly classified…" (2009, p. 3). They used correlations between the ACCUPLACER test scores and course success, percentage of students correctly placed, and the probability of success in a given course given different ACCUPLACER scores in order to conduct the meta-analysis. Mattern and Packman (2009) used a computer program by Schmidt and Le that utilized meta-analytic procedures in order to correct for sampling error, range restriction, and measurement unreliability. Their findings revealed a moderate to strong relationship between test scores and subsequent course performance. Specifically, the meta-analysis found that when success was defined as a "B" or higher, students were correctly placed in the correct math course between 64.5% and 66.5% of the time. In the same study, when success was defined as "C" or higher, the percentage of students correctly placed in math courses ranged from 73% to 84%.

In 1997, Day conducted a study on the assessment of a Computerized Placement Test (CPT) developed by ETS, the COMPASS placement test developed by ACT, and a paper/pencil instrument, the Academic and Assessment Placement Program Pretest (AAPP) in Tennessee higher education institutions. The researcher conducted a series of chi-square analyses to determine the significance of the relationship between placement scores and final grades in math courses, followed by Pearson correlations to assess the predictive validity between the variables. The results indicated that, in this particular study, there were low correlations between the placement score and the final grade in the math course. The researcher states that the results may be due, in part, to a small sample size. In this study, every effort was made to ensure that an appropriate sample size was utilized for a more reliable result.

A placement validity report on the COMPASS reading and math tests by Mzumara and Shermis (2001) utilized logistic regression to determine the optimum cutoff scores for math and writing courses at Indiana University Purdue University Indianapolis (IUPUI). These cutoff scores were originally set by a previous validation study, and were established based on empirical evidence (Mzumara, Shermis, & Fogel, 1998). Studies by LaForte (2000) and Verbout

(2013) reviewed correlations to assess the relationships between COMPASS and first year students' GPA, and relationships between ethnicity, COMPASS score and course completion, respectively. The current study seeks to ascertain whether or not the cut-off scores for the PERT test are appropriately placing students in the correct course.

Reviewing placement tests at the postsecondary level has several implications for the student and for the institution alike. It is important to ensure that the student is not taking a class he/she does not need if he/she is college ready; and conversely, that the student is not being placed into a college-level course for which he/she is not academically prepared. Developmental education courses cost money, and do not count toward a student's college credits. Properly placing students in the correct class would alleviate the financial burden of students paying for courses that are not relevant for a degree. Proper placement would also decrease time to degree for community college students who may typically already have other life priorities.

Bias, Differential Validity and Differential Prediction

Hierarchical models, or multilevel models, are often used in educational research because of the need to compare outcomes within nested data structures. Ma, Ma, and Bradley (2008) note the importance of using multilevel modeling in assessing math achievement outcomes among students within schools and across different schools. The practice of assessing the relationships between placement scores and various other outcomes at the student level and the classroom level can provide a more comprehensive view of the student, and could possibly lead to practices that would improve student success by decreasing time spent taking developmental courses. Multilevel models are discussed in a later section.

Arguably, one of the most important issues with test creation is that of ensuring that the test is free of bias. Meade and Fetzer (2009) define bias as "…the extent to which a person's expected observed score on a given test doesn't match that person's true score" (pp. 739-740). They also note that "when bias is present, test scores within groups may have meaning, but comparisons across groups may not be appropriate" (2009, p. 738). In 1968, Cleary analyzed the differences in scores on the SAT among Black and White students. Her study set precedence in the field of test fairness and bias by recognizing bias as occurring when "…if in the prediction of a criterion for which the test was designed, consistent nonzero errors of prediction are made for members of the subgroup. In other words, the test is biased if the criterion score predicted from the common regression line is consistently too high or too low for members of the subgroup" (p. 115). Her results indicated that, at the time, no bias was detected for two of the sample colleges in the east, but there was bias identified in the sample college located in the southwest.

In 1996, Noble et al. conducted a study of differential prediction in course placement for ethnic and gender groups based on ACT scores and subject area grade averages (SGA). The findings suggested that both the ACT and SGA slightly over-predicted course success for blacks and males when compared to whites and females.

Camilli and Shepard offer their own perception of bias, noting that there are two main ways to investigate it: an internal method and an external method (1994). The goal of the internal method, they cite, is a comparison between "…true group differences and bias in the measurement" (p. 15), or basically reviewing item-level bias and differential item functioning (DIF). The more commonly used external method seeks to determine bias through approaches such as predictive validity modeling.

Linn (1982, 1984, 1990) was a pioneer in the field of selection bias and on the topic of differential prediction. He has written extensively on admissions testing and the most appropriate ways in which to interpret these tests. The overarching conclusion for much of his research is that when deciding the placement of students, test scores should be used very responsibly and if possible, in combination with other information. The PERT test, given that it is in its preliminary stage, needs to be assessed to reveal whether or not there are significant differences in scores based on the gender or race of the students taking the exam. The population of students from which the current study was drawn is racially and ethnically diverse and this information is important should statistically significant differential prediction be found.

Standard Setting in Educational Placement Tests

The importance of establishing standard-setting policies, of setting cut-off scores, and the complexities involved in both these processes are noted in a number of studies (Berk, 1986; Chinn, 2006; Cizek, 1996; Collins, 2008; Livingston & Zeiky, 1982; Prince, 2005). The *Standards for Educational and Psychological Testing* state that no one method for setting cut scores is sufficient (2014).

Berk discusses that there are 38 methods for standard setting and standard adjustment, as well as two types of criteria to consider when evaluating standard setting methods. Within the two types of criteria, technical and practical, there are 10 standards that are advised. The technical component of standard setting is defined by Berk (1986) as "…the extent to which a method satisfies certain psychometric and statistical standards that would render it defensible to experts on standard setting" (p. 140). These criteria were garnered from three sources: 1) expert opinion based on research and logical assessments; 2) standards set forth by the *Standards for Educational and Psychological Research* (AERA, APA, NCME Joint Committee), and 3) relevant legal decisions. These technical criteria as listed by Berk include: (1) the method should yield appropriate classification information. The PERT should make sure that it clearly distinguishes between students passing and failing the placement exam; (2) the method should be sensitive to examinee performance. Every effort should be made to ensure that the items are at an adequate level of difficultly in order to accurately assess the student's mastery; (3) the method should be sensitive to instruction or training. A student's level of achievement prior to taking any placement test should be taken into consideration; (4) the method should be statistically sound. Accurate psychometric data for the PERT should be provided so that scores can be correctly interpreted; (5) the method should identify the "true" standard (the "true" standard refers to the true score scale). The difference between the true score scale and the observed score scale should be recognized, and decisions on whether to use one or the other should be explained; and (6) the method should yield decision validity evidence.

The practical component of standard setting ensures that the methods can be realistically implemented. These include:

7. The method should be easy to implement;
8. The method should be easy to compute;
9. The method should be easy to interpret to laypeople; and
10. The method should be credible to laypeople.

Best practice in cut score decision-making calls for the cut scores to be monitored by the institution regularly, and to be reviewed whenever there are curriculum changes. Morgan and Michaelides note that cut scores be reviewed every five to seven years for relevancy (2005). The PERT was implemented in 2010 and cut scores were modeled after the previous placement exam, the ACCUPLACER. Cut score recommendations by The College Board encourages each institution to establish their own ACCUPLACER cut scores based on their own factors and data (http://media.collegeboard.com/digitalServices/pdf/accuplacer/accuplacer-method-for-setting-cut-scores.pdf). Understanding the proper procedure when developing cut scores is important for new measures like the PERT, and these procedures should be followed as a review of the PERT scores lend to more accurate standard setting. Information regarding the PERT indicated that the test developer of PERT would provide updated cut scores based on its scores. As of spring 2017, that information was not available.

Multilevel Modeling in Education

Data in many educational topics can sometimes require specialized statistical analyses. Often, the questions to which a researcher may want to find the answers involve additional variables. By utilizing multilevel techniques, researchers are able to analyze nested data and review the relationships that may occur at different levels, and take other variables into consideration. Raudenbush and Bryk (2002) note that this statistical method is known by a variety of names: multilevel-, mixed level-, mixed effects-, random effects-, random coefficient-, and covariance components-modeling. A very common example of this type of modeling is in a study such as this, where students are nested within course sections, with those course sections are nested within college campuses, and those college campuses are nested within a college.

Pedhazur (1997) specified “multilevel analysis uses information from all available levels (e.g., student, classrooms, schools), making it possible to learn how variables at one level affect relations among variables at another level. Moreover, multilevel analysis affords estimation of variance between groups as distinct from variance within groups” (p. 692). Pedhazur further stated that multilevel models “yield more realistic standard errors” than ordinary least squares (OLS) estimates (p. 692).

The use of multilevel models, or hierarchical linear models, in education has increased greatly in recent years in part due to the realization that the educational hierarchy naturally falls right in line with the structure of multilevel modeling. Prather (2007) utilized a two-level hierarchical linear model to describe teacher effectiveness by using a reading test as the criterion measure within a nested structure. Hudson (2015) also implemented a multilevel model approach in order to analyze the relationship between motivational factors and student academic achievement.

Lee (2000) identifies three main problems that occur when analyzing multilevel data with single-level methods: aggregation bias, misestimated standard errors, and heterogeneity of regression. Aggregation bias, she states, occurs when a variable takes on different meaning and has different effects at different levels of aggregation (2000). Misestimated standard errors can occur when the researcher mistakenly treats individual cases as independent when they are not. Finally, heterogeneity of regression occurs when, for example, the relationship between the characteristics of students and the outcome being measured vary across schools and may be functions of group-level variables (Lee, 2000). Multilevel models give the researcher an opportunity to analyze and compare many components, such as group level differences, cross level interactions and variance across levels (Raudenbush & Bryk, 2002). The process by which to perform multilevel modeling is discussed in the following chapter.

Chapter Summary

College readiness in the community college has long been an area of concern among educators. To assess readiness in Florida, law-makers instituted admissions and placement exams such as the FCAT (replaced in Spring 2015 with the Florida Standards Assessment), the ACT, the SAT and the ACCUPLACER. If a student is deemed not college-ready, a series of developmental education courses are suggested in order to assist the student in gaining the knowledge, skills and abilities needed in order to be successful in college-level courses. In 2008, driven by the desire to address newly established postsecondary readiness competencies, Florida developed a customized placement test known as the Postsecondary Education Readiness Test (PERT) to be used state-wide. The purpose of this research is to utilize multilevel model in which students (level 1) are nested within classes (level 2) in order to analyze the predictive abilities of the PERT for student course grade in order to assess its validity and how other factors, such as instructor employment status (part-time/ full-time), instructor years of employment, course section size, gender, and race may affect that validity.

CHAPTER THREE

METHOD

The purpose of this study is to investigate the predictive validity of the Postsecondary Education Readiness Test (PERT) and assess its predictive validity of student course grade in the following courses: MAT 0018 (Developmental Math I), MAT 0028 (Developmental Math II), MAT 1033 (Intermediate Algebra) and MAC 1105 (College Algebra). This study also analyzed the variation of the relationships by student-level variables and course section variables. The focus was on math due to state data showing that a higher percentage of students require developmental math when compared to developmental reading and developmental writing. Because the PERT is a new assessment instrument, a major benefit of this study is that it will add to the body of knowledge pertaining to this test. A benefit more specific to the college is an analysis of the scores used for placement in math courses, and whether or not that placement is accurate.

Participants

This study used student data from a large multi-campus community college (HCC) in the southeast, with an unduplicated headcount enrollment of over 46,000 students for the 2012-13 year. For this same year, the college reported a diverse student population consisting of 21% Black/ African American, 27% Hispanic, and 46% White (HCC Factbook 2013). The student body is also diversified by gender, with a reported 55% female students.

The population consisted of all students who took the PERT test in fall of 2012 and subsequently enrolled in a developmental math course, MAT 1033 or MAC 1105 as a result of their score. The fall 2012 term was used as a starting point in order to find enough students who had taken any initial developmental course and subsequently enrolled and completed either MAT 1033 (Intermediate Algebra) or MAC 1105 (College Algebra). With this approach, the time frame of the data reviewed was from August of 2012 through June of 2016. For the fall term of 2012, the college offered 64 sections of MAT 0018 (Developmental Math I) and 197 sections of MAT 0028 (Developmental Math II). This population did not include students who were enrolled in a dual enrollment course in the same year in which the PERT was taken. For the purposes of this research, any student with incomplete or missing course grades were removed from this data analysis. Confidentiality of the records was ensured by reporting, analyzing, and maintaining the data on a secure server on the college's network. The risk of a breach of confidentiality was minimal.

Research Questions

Research questions that are explored in this study include:

1. To what extent does the predictive relationship of the PERT with final grades in Developmental Math I (MAT 0018) and Developmental Math II (MAT 0028) vary based on student gender?
2. To what extent does the predictive relationship of the PERT with final grades in Developmental Math I (MAT 0018) and Developmental Math II (MAT 0028) vary based on student race/ethnicity?
3. To what extent does the predictive relationship of the PERT with final grades in Intermediate Algebra (MAT 1033) and College Algebra (MAC 1105) vary based on student gender?
4. To what extent does the predictive relationship of the PERT with final grades in Intermediate Algebra (MAT 1033) and College Algebra (MAC 1105) vary based on race/ethnicity?
5. What combined student and course section-level variables (i.e., first-time-in-college status, student race/ethnicity, student gender, student enrollment status, student

age, instructor employment status, course time of day, course delivery method) are significant predictors of student final course grade in Developmental Math I (MAT 0018)? In Developmental Math II (MAT 0028)?

6. What combined student and course section-level variables (i.e., first-time-in-college status, student race/ethnicity, student gender, student enrollment status, student age, instructor employment status, course time of day, course delivery method) are significant predictors of student final course grade in Intermediate Algebra (MAT 1033)? In College Algebra (MAC 1105)?

Data Collection

Prior to any data retrieval, the researcher sought approval from the large multi-campus community college's Institutional Review Board (IRB) and the University of South Florida's IRB in order to ensure that the study followed federal mandates protecting personal and identifiable data. Upon approval, the researcher, who is employed at the college, retrieved the data. The number of records to be reviewed was estimated to be 500,000; however after a careful review of the data, the final datasets included less than 5,000 students in each course. The information was gathered from archival files previously submitted to the State and were analyzed using SAS 9.3 software.

Variables

The student level variables included student race/ethnicity, gender, enrollment status (Part-time or Full-time), age, and whether the student was first-time-in-college (FTIC). The students' final grade in Developmental Math I (MAT 0018), Developmental Math II (MAT 0028), Intermediate Math (MAT 1033), College Algebra (MAC 1105), were included along with the student's PERT score. If a student took any course more than once, or took the PERT score more than once, that was not included as a variable in the current study, as only one grade and a single PERT score is included in a student's data file.

The course section variables included instructor employment status (Part-time or Full-time), number of years the instructor was employed at the institution, time of day of the course section (Day or Evening), and the course section delivery method (Campus or Online). Table 3 lists the variables and provides information about how the variables were coded.

Multilevel modeling was utilized in analyzing the relationships between the PERT and the final grades in either of MAT 0018 and MAT 0028 (developmental math courses), MAT 1033 (Intermediate Math) and MAC 1105 (College Algebra). The PERT score served as a predictor variable, while the criterion measure was the course grade in those respective courses. The main independent variable of the multilevel study was the PERT score. Student level variables were considered Level-1 variables while course section variables were categorized as Level-2 variables in the multilevel models.

Table 3

Variables and Description

Variable Description	Coding
Course Section-Level:	
Employment Status of Instructor	1= Full Time 0= Part time
Number of years the instructor has been employed by the institution	Continuous
Time of day of course	1= Day (Before 3pm) 0= Evening (After 3pm)
Course delivery method	1= Campus 0= Online
Student-Level:	
First Time in College indicator	1= Yes 0= No
Student race Black Hispanic Other White (used as reference group	 1=Black 0=Non-Black 1=Hisp. 0=Non-Hisp. 1=Other 0=Non-Other 1=White 0=Non-White
Student Gender	1= Female 0=Male
Enrollment Status Dichotomous variable for student status	1= Full Time 0= Part Time
Age of the student	Continuous
PERT scaled score	50 to 150
Grade in lower developmental math course (MAT 0018), if taken	4= A 3= B 2= C 1= D 0= F, FX
Grade in upper developmental math course (MAT 0028), if taken	4= A 3= B 2= C 1= D 0= F, FX
Grade in MAT 1033 (Intermediate Algebra), if taken	4= A 3= B 2= C 1= D 0= F, FX
Grade in MAC 1105 (College Algebra), if taken	4= A 3= B 2= C 1= D 0= F, FX

Preliminary Analysis

The data mining process required extraction from multiple data files. Data were selected and recoded based on the variables of interest within the study. Dummy coding of most of the variables was utilized to facilitate interpretation. The remaining variables were left as continuous and categorical. An examination of a number of factors preceded the formal statistical analysis. This preliminary analysis was done in order to screen for missing data across variables, examine the relationships missing data may have with other variables, and to determine whether there were outliers in the data. Missing data were addressed in the initial assembly of the dataset by using statistical methods to ensure that each subject had all required variables. Any cases with missing data were eliminated from the dataset/ Means, standard deviations, and other descriptive statistics were calculated for these data and are described in chapter four (see Tables 4 through 7).

Statistical Analysis

Analyses of these data were conducted using two-level multilevel models in order to examine the predictive relationship between final grade in each of the four math courses and PERT score, relationships among variables, and relationships within and between course sections. A multilevel model was utilized where students were the first level and the course section was the second level (i.e., students nested within course sections). All tests of significance were conducted at the $p< .01$ level.

The first step involved fitting an unconditional means model, which analyzes the amount of variance that is explained by a model with no predictor variables. The next steps involved entering student level variables, interaction effects, and course level variables at specific intervals. The first model (1a) included student level variables student race, student gender, student enrollment status, student age and first-time-in-college status. The decision was made a priori to allow the PERT score-course grade relationship to vary randomly across sections after assessing all models. Student age was group mean centered prior to inclusion in the models. This model is represented by the following equation:

Model 1a: Final Grade = $\beta_0 + \beta_1(Black) + \beta_2(Hispanic) + \beta_3(Other) + \beta_4(Gender) + \beta_5(Enrollment\ Status) + \beta_6(Age) + \beta_7(FTIC) + R_{ij}$

The second model, 1b, included all variables from the first model, plus the PERT:

Model 1b: Final Grade = $\beta_0 + \beta_1(Black) + \beta_2(Hispanic) + \beta_3(Other) + \beta_4(Gender) + \beta_5(Enrollment\ Status) + \beta_6(Age) + \beta_7(FTIC) + \beta_8(PERT) + R_{ij}$

The third model, 1c, included all Model 1b variables, plus an interaction of gender by PERT score:

Model 1c: Final Grade = $\beta_0 + \beta_1(Black) + \beta_2(Hispanic) + \beta_3(Other) + \beta_4(Gender) + \beta_5(Enrollment\ Status) + \beta_6(Age) + \beta_7(FTIC) + \beta_8(PERT) + \beta_9(Gender\ x\ PERT) + R_{ij}$

The fourth model, 1d, included all Model 1b variables, plus an interaction of race by PERT score:

Model 1d: Final Grade = $\beta_0 + \beta_1(Black) + \beta_2(Hispanic) + \beta_3(Other) + \beta_4(Gender) + \beta_5(Enrollment\ Status) + \beta_6(Age) + \beta_7(FTIC) + \beta_8(PERT) + \beta_{10}(Black\ x\ PERT) + \beta_{11}(Hispanic\ x\ PERT) + \beta_{12}(Other\ x\ PERT) + R_{ij}$

Model 2 included all student variables for each course, and the inclusion of the course section variables. Models were repeated for each of the courses being analyzed (MAT0018, MAT0028, MAT1033 and MAC1105). In all models, student age and PERT score were grand mean centered. Student variables and course section variables were included as fixed predictors in all models. In model 1b, PERT score was included as a random variable, along with the intercept.

Chapter Summary

The purpose of this study was to investigate the predictive validity of the Postsecondary Education Readiness Test (PERT) and assess its predictive validity of student course grade in the following courses: MAT 0018 (Developmental Math I), MAT 0028 (Developmental Math II), MAT 1033 (Intermediate Algebra) and MAC 1105 (College Algebra). A multilevel modeling approach was used due to the nested structure of the data and was used to analyze the relation between the students' PERT scores and final grades in the selected math courses across the multiple sections of the course.

CHAPTER FOUR

RESULTS

This research study examined the predictive relationships of Florida's Postsecondary Education Readiness Test (PERT) with students' final grades in selected math courses in the community college. This study also examined whether these predictive relationships varied by students' gender and race. This chapter provides descriptive information about the sample, as well as the findings related to the research questions. This chapter describes the student samples used in the research, the statistical measures used to analyze the variables, and finally the results of the analyses.

Description of the Samples

Descriptive statistics were used to describe the sample used in this study. The main sample consisted of students who took the PERT math test in fall 2012. Four sub-samples were established from that sample. They included students who 1) tested into the Developmental Math I course, MAT 0018, and subsequently enrolled in lower Developmental Math I (n=727); 2) tested into Developmental Math II course, MAT 0028, and subsequently enrolled in upper Developmental Math II (n=900); 3) tested into Intermediate Algebra course, MAT 1033 (n=713), and subsequently enrolled in Intermediate Algebra; and 4) tested into College Algebra course, MAC 1105 (n=271), and subsequently enrolled in College Algebra. Tables 4 through 7 present the demographic makeup of MAT 0018, MAT 0028, MAT 1033, and MAC 1105.

Table 4

Descriptive Statistics for Level 1 Variables: All Courses

	MAT 0018		MAT 0028		MAT 1033		MAC 1105	
Variable	***n***	**%**	***n***	**%**	***n***	**%**	***n***	**%**
Gender	**710**		**874**		**703**		**263**	
Male (0)	278	39.2%	371	42.4%	331	47.1%	160	60.8%
Female (1)	432	60.8%	503	57.6%	372	52.9%	103	39.2%
Race	**727**		**900**		**713**		**271**	
Black	210	28.9%	216	24.0%	124	17.4%	35	12.9%
Hispanic	213	29.3%	287	31.9%	233	32.7%	89	32.8%
White	274	37.7%	345	38.3%	311	43.6%	113	41.7%
Other	30	4.1%	52	5.8%	45	6.3%	34	12.6%
First Time in College	**723**		**899**		**710**		**269**	
No (0)	435	60.2%	443	49.3%	528	74.4%	213	79.2%
Yes (1)	288	39.8%	456	50.7%	182	25.6%	56	20.8%
Age	**727**		**900**		**713**		**271**	
18-28	514	70.1%	780	86.7%	629	88.2%	222	81.9%
29-38	118	15.9%	77	8.5%	61	8.6%	40	14.8%
39-48	70	9.1%	33	3.7%	17	2.4%	5	1.8%
49-68	25	4.9%	10	1.1%	6	0.8%	4	1.5%
Student Enrollment Status	**727**		**900**		**713**		**270**	
Part-time (0)	342	47.0%	389	43.2%	320	44.9%	122	45.2%
Full-time (1)	385	53.0%	511	56.8%	393	55.1%	148	54.8%

Note. MAT 0018=Lower Developmental Math I. MAT 0028=Upper Developmental Math II. MAT 1033=Intermediate Algebra. MAC 1105=College Algebra.

When looking at the breakdown of the student demographics across all courses, female students constitute more than half of the samples for all courses except in MAC 1105. In MAT 0018, MAT1033, and MAC 1105, more than half of the students were classified as first time in college. The largest percentages of students in the samples were between the ages of 18 and 28, and were full-time students. Finally, students classified as being of Hispanic ethnicity made up the second largest category after students classified as White.

Table 5

Descriptive Statistics for Level 2 Variables: All Courses

	MAT 0018		MAT 0028		MAT 1033		MAC 1105	
Variable	***n***	**%**	***n***	**%**	***n***	**%**	***n***	**%**
Instructors	47		73		84		70	
Employment Status								
Full Time (1)	14	30%	20	37.4%	36	43.9%	36	51.4%
Part Time (0)	33	70%	53	72.6%	46	56.1%	34	48.6%
missing						*2*		
Years Teaching								
1-5.99	23	49%	37	50.7%	37	45.1%	27	38.6%
6-15.99	13	28%	24	32.9%	37	45.1%	32	45.7%
16+	11	23%	12	16.4%	8	9.8%	11	15.7%
missing						*2*		
Number of Sections Taught								
1	31	66%	16	21.9%	21	25.0%	24	34.3%
2	15	32%	21	28.8%	15	17.9%	16	22.9%
3	1	2%	14	19.2%	7	8.3%	13	18.6%
4	0		16	21.9%	11	13.1%	5	7.1%
5	0		3	4.1%	10	11.9%	0	
6-10	0		3	4.1%	19	22.6%	11	15.7%
11-15	0				1	1.2%	1	1.4%
Time of Sections Taught	***n*=64**		***n*=197**		***n*=328**		***n*=204**	
Day (1)	45	70%	131	68.2%	208	70.3%	134	76.6%
Evening (0)	18	30%	57	29.7%	88	29.7%	41	23.4%
N/A			4	2.1%				
missing			*5*		*32*		*29*	
Course Delivery Method								
Campus (1)	64	100%	192	98.0%	281	87.3%	171	84.2%
Online (0)	0		4	2.0%	41	12.7%	32	15.8%
missing			*1*		*6*		*1*	

A review of the course section variables show that higher percentages of instructors teaching MAT 0018, MAT 0028 and MAT 1033 were part-time instructors. A majority of the instructors taught less than 16 years, and taught between one and five sections. Across all courses, the sections were primarily taught during the day, and were delivered in an on-campus setting.

Table 6

Final Grade by PERT Scores

PERT (MAC 1105)	*n*=271	Success (A, B, C)	Fail (D, F, FX)
143-150	14	100.0% (*n*=14)	0.0% (*n*=0)
133-142	50	88.0% (*n*=44)	12.0% (*n*=6)
123-132	207	76.8% (*n*=159)	23.2% (*n*=48)
PERT (MAT 1033)	***n*=713**	**Success (A, B, C)**	**Fail (D, F, FX)**
119-122	217	79.3% (*n*=172)	20.7% (*n*=45)
114-118	496	75.2% (*n*=373)	24.8% (n=123)
PERT (MAT 0028)	***n*=900**	**Success (A, B, C)**	**Fail (D, F, FX)**
108-113	195	85.1% (*n*=166)	14.9% (*n*=29)
102-107	358	80.2% (*n*=287)	19.8% (*n*=71)
96-101	347	76.1% (*n*=264)	23.9% (n=83)
PERT (MAT 0018)	***n*=727**	**Success (A, B, C)**	**Fail (D, F, FX)**
85-95	420	82.4% (*n*=346)	17.6% (*n*=74)
75-84	198	71.7% (*n*=142)	28.3% (*n*=56)
65-74	57	71.9% (*n*=41)	28.1% (*n*=16)
55-64	34	50.0% (*n*=17)	50.0% (*n*=17)
50-54	18	44.4% (*n*= 8)	55.6% (*n*=10)

Table 7

Summary of Course Information

Course	MAT 0018	MAT 0028	MAT 1033	MAC 1105
Instructors	47	73	84	70
Course Section	64	197	328	204
Students	727	900	713	271

Note. MAT 0018=Lower Developmental Math I. MAT 0028=Upper Developmental Math II. MAT 1033=Intermediate Algebra. MAC 1105=College Algebra.

Tests of normality were run for all models in all courses, and the residuals were approximately normally distributed with constant variance. The graphical output for each of these courses can be found in Appendix D.

Model Specification

The primary question in this research study is whether specific student and course-section variables can successfully predict final grade in the math course of interest and whether the PERT scores incrementally add to the explained variance in students' final grades.

The first model, 1a, was fit to include the student level variables race (Black, Hispanic, and Other), gender, enrollment status, age and first-time-in-college status as predictors. This model is written as

Final Grade = $\beta_0 + \beta_1(Black) + \beta_2(Hispanic) + \beta_3(Other) + \beta_4(Gender) + \beta_5(Enrollment\ Status) + \beta_6(Age) + \beta_7(FTIC) + R_{ij}$

The second model, 1b, included all variables from the first model, plus the PERT:

Model 1b: Final Grade = $\beta_0 + \beta_1(Black) + \beta_2(Hispanic) + \beta_3(Other) + \beta_4(Gender) + \beta_5(Enrollment\ Status) + \beta_6(Age) + \beta_7(FTIC) + \beta_8(PERT) + R_{ij}$

The third model, 1c, included all Model 1b variables, and an interaction of gender by PERT score:

Model 1c: Final Grade = $\beta_0 + \beta_1(Black) + \beta_2(Hispanic) + \beta_3(Other) + \beta_4(Gender) + \beta_5(Enrollment\ Status) + \beta_6(Age) + \beta_7(FTIC) + \beta_8(PERT) + \beta_9(Gender\ x\ PERT) + R_{ij}$

The fourth model, 1d, included all Model 1b variables, plus an interaction of race by PERT score:

Model 1d: Final Grade = $\beta_0 + \beta_1(Black) + \beta_2(Hispanic) + \beta_3(Other) + \beta_4(Gender) + \beta_5(Enrollment\ Status) + \beta_6(Age) + \beta_7(FTIC) + \beta_8(PERT) + \beta_{10}(Black\ x\ PERT) + \beta_{11}(Hispanic\ x\ PERT) + \beta_{12}(Other\ x\ PERT) + R_{ij}$

Model 2 included all student variables, and the inclusion of the course section variables. Models were repeated for each of the courses being analyzed (MAT0018, MAT0028, MAT1033, and MAC1105).

Data Analysis

The first step in the statistical modeling process was to formulate an unconditional model in order to calculate the Intraclass Correlation Coefficient (ICC). This ICC tells the researcher how much variance of the final grade in math occurs due to the differences among course sections. The Level 1 equation for the unconditional model is written as $Y_{ij} = \beta_{0j} + R_{ij}$, in which Y_{ij} is the final grade of student *i* in the *j*th section, β_{0j} is the average grade for section *j*, and *R* is the student level error term. The Level 2 equation is written as $\beta_{0j} = \gamma_{00} + u_{0j}$, in which β_{0j} is intercept for the *j*th section, and γ_{00} is the grand mean outcome of all students with u_{0j} being the random effect associated with the *j*th section. The unconditional model is expressed as a combination of the level 1 and level 2 regression equations:

Final Grade= $\gamma_{00} + u_{0j} + R_{ij}$

MAT 0018 Models

Table 20 displays the output from the unconditional model. In the unconditional model run for the MAT 0018 student sample consisting of 727 students in 64 course sections, one can conclude that the average final grade in MAT 0018 is 2.27, and that there is more variation within sections (1.90, *p*<.01) than among the different sections (0.12). Calculating the intraclass correlation coefficient (ICC), 0.12 / (.12 + 1.90) reveals that approximately 6% of the variance in final grade in MAT 0018 is due to differences between course sections.

Table 8

Estimates for MAT 0018 Unconditional Model

Fixed Effect	*Estimate*	*Standard Error*	*t Value*	*Pr > |t|*
Intercept	2.27	0.07	33.38	<.01
Variance Components	*Estimate*	*Standard Error*	*Z Value*	*Pr > Z*
Intercept	0.12	0.05	2.30	.01
Residual	1.90	0.10	18.21	<.01

Note. Intraclass correlation coefficient = .12/ (.12 + 1.90) = .059

MAT 0018 Model 1a: Entering Student Level Variables

The first model included the student level variables of first time in college status (1=Yes), race, gender (1=Female), student enrollment status (1=Full Time), and age. In this model, also known as the random-coefficient model, the seven fixed effects reflect the seven student level predictors, and the variable for age has been grand mean-centered by subtracting the mean age from the raw age of each respondent. Race was initialized as a binary variable for Black (1=Black, 0=Non-Black), Hispanic, (1=Hispanic, 0=Non-Hispanic), Other (1=Other, 0=Non=Other) and White. White is used as the reference group, and is therefore left out of the model.

The covariance structure for this first model was kept at the default for SAS: variance components. Utilizing the variance components statement ensures that any random effects are uncorrelated with each other. Table 9 shows the output from this model.

Table 9

Estimation of Random Effects for MAT 0018, Model 1a: Student Level Variables

Covariance Parameter	*Estimate*	*Standard Error*	*Z Value*	*Pr Z*
Intercept	0.14	0.06	2.54	.01
Residual	1.76	0.10	17.82	<.01

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female).

The covariance parameter estimates show that the intercept variance is estimated at .14 with a standard error of .06 (p=.01), and the within section variance is significant at p<.01.

In looking at the fixed effects, there are two predictors that differ significantly from zero (race-Black and age). The intercept represents the mean final grade when all predictors are set to ‘0’. The estimated value of -0.65 (p<.01) for the race-Black predictor means that on average, Black students have a lower final grade in MAT 0018 than those compared to the reference group (White). The fact that age is significant means that on average, older students have a higher final grade, as indicated by the observed *t*-score, 4.84 ($p<.01$). Table 10 shows all fixed effects for model 1a.

Table 10

Estimation of Fixed Effects for MAT 0018, Model 1a: Student Level Variables

Fixed Effects	*Estimate*	*Standard Error*	*t Value*	*Pr > |t|*
Intercept	2.61	0.13	19.55	<.01
First Time in College	-0.08	0.11	-0.72	.47
Race-Black	-0.65	0.13	-5.10	<.01
Race-Hispanic	-0.21	0.13	-1.69	.09
Race-Other	-0.25	0.28	-0.89	.38
Student Enrollment Status	-0.15	0.11	-1.41	.16
Student Gender	0.06	0.11	0.61	.54
Student Age	0.03	0.01	4.84	<.01

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female).

MAT 0018 Model 1b: Entering Model 1a Variables Plus PERT Score

In the second model of the multilevel analysis, the variables from Model 1a remained, and the PERT score was added as a predictor variable. The model contains a random intercept and a random slope for the PERT score, which allows the slopes for the PERT score to vary by section. PERT score has been grand-mean centered using the same method as for the age. The PERT score was also included as a random effect to specify that the effect of the mean PERT score on the final grade in the course can vary across sections. Table 11 shows that the estimated variance for the PERT was 0, in which cases SAS does not calculate standard error or *p*-value. This output leads to the conclusion that the PERT slope does not significantly vary by section.

Table 11

Estimation of Random Effects for MAT 0018, Model 1b: PERT Score

Covariance Parameter	*Estimate*	*Standard Error*	*Z Value*	*Pr Z*
Intercept	0.11	0.05	2.23	.01
PERT Final Grade Slope	0	---	---	---
Error Covariance	-0.004	0.002	-1.75	.08
Residual	1.67	0.09	17.80	<.01

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

The fixed effects for model 1b show that there were four significant predictors (p<.01). The student race (Black) predictor has a negative effect on the mean final grade in this course. Race-Black is coded as 1 (0=Non-Black), so the results indicate that Black students have lower final grades in MAT 0018 when compared to White students. The other significant fixed effects imply that mean student age (b=0.04, p<.01) and PERT score (b=0.03, p<.01) have an effect on the overall final grade in MAT 0018. Students with higher PERT scores tended to have higher final grade scores in the class. Table 12 shows the results of the estimation of fixed effects for this model.

Table 12

Estimation of Fixed Effects for MAT 0018, Model 1b: PERT Score

Fixed Effects	*Estimate*	*Standard Error*	*t Value*	*Pr > |t|*
Intercept	2.56	0.13	20.00	<.01
First Time in College	-0.10	0.11	-0.97	.33
Race-Black	-0.60	0.12	-4.84	<.01
Race-Hispanic	-0.17	0.12	-1.36	.18
Race-Other	-0.17	0.27	-0.64	.52
Student Enrollment Status	-0.22	0.10	-2.09	.04
Student Gender	0.13	0.10	1.30	.19
Student Age	0.04	0.01	6.24	<.01
PERT Score	0.03	0.01	6.49	<.01

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

MAT 0018 Model 1c: Entering Model 1a Variables Plus Gender Interaction

For the third model, the Model 1a variables were kept and an interaction effect for gender was added. Gender was coded as a dichotomous variable with Female=1, Male=0. The results of the random effects estimation in Table 13 show that only the within section variance (*b*=1.67, *p*<.01) differs significantly from zero.

Table 13

Estimation of Random Effects for MAT 0018, Model 1c: Gender Interaction

Covariance Parameter	*Estimate*	*Standard Error*	*Z Value*	*Pr Z*
Intercept	0.10	0.05	2.19	.01
Residual	1.67	0.09	17.80	<.01

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

In reviewing the fixed effects for the model (Table 14), only the Race-Black (*b*= -0.59, *p*<.01), student age *b*=0.04, *p*<.01) and PERT score (*b*=0.04, *p*<.01) predictors differ significantly from zero. The gender interaction was not significant, which implies that the relationship between PERT and final grades does not vary by gender.

Table 14

Estimation of Fixed Effects for MAT 0018, Model 1c: Gender Interaction

Fixed Effects	*Estimate*	*Standard Error*	*t Value*	*Pr > |t|*
Intercept	2.54	0.13	19.41	<.01
First Time in College	-0.11	0.11	-1.01	.31
Race-Black	-0.59	0.12	-4.78	<.01
Race-Hispanic	-0.17	0.12	-1.38	.17
Race-Other	-0.20	0.27	-0.73	.47
Student Enrollment Status	-0.20	0.10	-1.97	.05
Student Gender	0.14	0.10	1.34	.18
Student Age	0.04	0.01	6.2	<.01
PERT Score	0.04	0.01	3.81	<.01
PERT Score*Student Gender	-0.001	0.01	-0.13	.90

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

MAT 0018 Model 1d: Entering Model 1a Variables Plus Race Interaction

The fourth model involved keeping Model 1a and adding a race interaction. Interactions were added for races coded as Black, Hispanic and Other. The within section variance is the only random effect that differs significantly from zero.

Table 15

Estimation of Random Effects for MAT 0018, Model 1d: Race Interaction

Covariance Parameter	*Estimate*	*Standard Error*	*Z Value*	*Pr Z*
Intercept	0.10	0.05	2.17	.01
Residual	1.67	0.09	17.78	<.01

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

The fixed effects in Table 16 show that none of the race interaction effects are significant: PERT Score*Race-Black (*b*=0.01, *p*=.24), PERT Score*Race-Hispanic (*b*=0.01, *p*=.39), PERT Score*Race-Other (*b*=0.02, *p*=.43).

Table 16

Estimation of Fixed Effects for MAT 0018, Model 1d: Race Interaction

Fixed Effects	*Estimate*	*Standard Error*	*t Value*	*Pr > |t|*
Intercept	2.57	0.13	19.85	<.01
First Time in College	-0.11	0.11	-1.04	.30
Race-Black	-0.61	0.12	-4.87	<.01
Race-Hispanic	-0.19	0.12	-1.53	.13
Race-Other	-0.21	0.27	-0.77	.44
Student Enrollment Status	-0.20	0.10	-1.94	.05
Student Gender	0.13	0.10	1.23	.22
Student Age	0.04	0.01	6.20	<.01
PERT Score	0.03	0.01	2.90	.004
PERT Score*Race-Black	0.01	0.01	1.18	.24
PERT Score*Race-Hispanic	0.01	0.01	0.86	.39
PERT Score*Race-Other	0.02	0.03	0.78	.43

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

MAT 0018 Model 2: Entering Level 1 Variables and Level 2 Variables

The final model was the inclusion of all level one variables, as well as the level two variables. This is known as the full model, and the results of both the random effects and the fixed effects are shown in Tables 17 and 18.

The random effects show that only the residual variance differs significantly from zero at $p<.01$. This implies that when all variables are included in the model, there is evidence of variability within sections.

Table 17

Estimation of Random Effects for MAT 0018, Model 2: Level 1 and Level 2 Variables

Covariance Parameter	*Estimate*	*Standard Error*	*Z Value*	*Pr Z*
Intercept	0.10	0.05	2.17	.01
Residual	1.66	0.09	17.72	<.01

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

In this full model represented in Table 18, Race-Black (b= -0.57, $p<.01$), student age (b=0.04, $p<.01$), and PERT score (b=0.03, $p<.01$) are statistically significant. No interaction effects were statistically significant.

None of the level two variables were found statistically significant in the model, and it can be concluded that none of these variables have a significant effect on final grade in MAT 0018. Because the only method of delivery for MAT 0018 is on campus, the course delivery predictor was not included as a predictor in this model.

Table 18

Estimation of Fixed Effects for MAT 0018, Model 2: Level 1 and Level 2 Variables

Fixed Effects	*Estimate*	*Standard Error*	*t Value*	*Pr > |t|*
Intercept	2.36	0.19	12.72	<.01
First Time in College	-0.14	0.11	-1.30	.19
Race-Black	-0.56	0.13	-4.51	<.01
Race-Hispanic	-0.17	0.12	-1.37	.17
Race-Other	-0.21	0.27	-0.76	.44
Student Enrollment Status	-0.22	0.11	-2.12	.03
Student Gender	0.14	0.10	1.39	.16
Student Age	0.04	0.01	6.13	<.01
PERT Score	0.03	0.01	6.51	<.01
Instructor Years Teaching	0.01	0.01	1.00	.32
Instructor Employment Status	0.13	0.15	0.84	.41
Course Delivery	---	---	---	---
Time of Day of Course	0.10	0.15	0.61	.54

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered. Instructor Employment Status (0=Part-Time, 1=Full-Time). Time of Day (0=Evening, 1=Day). Course Delivery (0=Online, 1=Campus).

MAT 0028 Models

The ICC run for this unconditional model indicated that 6.8% of the variance was explained by a model with no predictors.

Table 19

Estimates for MAT 0028 Unconditional Model

Fixed Effect	*Estimate*	*Standard Error*	*t Value*	*Pr > |t|*
Intercept	2.48	0.05	45.43	<.01
Variance Components	*Estimate*	*Standard Error*	*Z Value*	*Pr > Z*
Intercept	0.13	0.06	2.44	.01
Residual	1.78	0.09	19.32	<.01

Note. Intraclass correlation coefficient = .13/ (.13 + 1.78) = .068

MAT 0028 Model 1a: Entering Student Level Variables

The random effects for this model reveals that only within section variance (*b*=1.76) is significantly different from 0 at *p*<.01.

Table 20

Estimation of Random Effects for MAT 0028, Model 1a: Student Level Variables

Covariance Parameter	*Estimate*	*Standard Error*	*Z Value*	*Pr Z*
Intercept	0.14	0.06	2.54	.01
Residual	1.76	0.10	17.82	<.01

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female).

After entering all student variables, the results of Table 21 reveal that there are two statistically significant variables for MAT 0028: student gender (*b*=0.39, *p*<.01) and student age (*b*=0.05, *p*<.01).

Table 21

Estimation of Fixed Effects for MAT 0028, Model 1a: Student Level Variables

Fixed Effects	*Estimate*	*Standard Error*	*t Value*	*Pr > |t|*
Intercept	2.45	0.11	22.54	<.01
First Time in College	-0.06	0.10	-0.59	.56
Race-Black	-0.31	0.12	-2.62	.01
Race-Hispanic	0.05	0.11	0.49	.63
Race-Other	0.18	0.20	0.89	.37
Student Enrollment Status	0.19	0.09	2.08	.04
Student Gender	0.39	0.09	4.27	<.01
Student Age	0.05	0.01	7.48	<.01

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female).

MAT 0028 Model 1b: Entering Model 1a Variables Plus PERT Score

Including the PERT score as a random effect revealed that the slope variance for PERT across sections was practically nonexistent. SAS did not report any output due to the lack of variance.

Table 22

Estimation of Random Effects for MAT 0028, Model 1b: PERT Score

Covariance Parameter	*Estimate*	*Standard Error*	*Z Value*	*Pr Z*
Intercept	0.53	0.28	1.89	.03
PERT Final Grade Slope	0	---	---	---
Error Covariance	-0.01	0.01	-1.57	.12
Residual	1.60	0.08	18.87	<.01

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

The table of fixed effects shows that the PERT score differs significantly from zero, which means that this predictor has a positive effect on the final grade in MAT 0028.

Table 23

Estimation of Fixed Effects for MAT 0028, Model 1b: PERT Score

Fixed Effects	*Estimate*	*Standard Error*	*t Value*	*Pr > |t|*
Intercept	1.66	0.23	7.20	<.01
First Time in College	-0.05	0.10	-0.57	.57
Race-Black	-0.29	0.12	-2.48	.01
Race-Hispanic	0.04	0.11	0.38	.70
Race-Other	0.18	0.20	0.89	.38
Student Enrollment Status	0.19	0.09	2.07	.04
Student Gender	0.38	0.09	4.22	<.01
Student Age	0.06	0.01	8.00	<.01
PERT score	0.04	0.01	4.10	<.01

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

MAT 0028 Model 1c: Entering Model 1a Variables Plus Gender Interaction

The fixed effects of this model of gender interaction showed no significance with regard to gender variability.

Table 24

Estimation of Random Effects for MAT 0028, Model 1c: Gender Interaction

Covariance Parameter	*Estimate*	*Standard Error*	*Z Value*	*Pr Z*
Intercept	0.12	0.05	2.31	.01
Residual	1.60	0.08	18.86	<.01

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

Table 25

Estimation of Fixed Effects for MAT 0028, Model 1c: Gender Interaction

Fixed Effects	*Estimate*	*Standard Error*	*t Value*	*Pr > |t|*
Intercept	1.64	0.32	5.06	<.01
First Time in College	-0.04	0.10	-0.48	.63
Race-Black	-0.30	0.12	-2.55	.01
Race-Hispanic	0.04	0.11	0.41	.68
Race-Other	0.16	0.20	0.80	.42
Student Enrollment Status	0.18	0.09	2.01	.04
Student Gender	0.36	0.41	0.88	.38
Student Age	0.06	0.01	7.80	<.01
PERT score	0.04	0.02	2.68	.01
PERT score*Student Gender	0.001	0.02	0.04	.96

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

MAT 0028 Model 1d: Entering Model 1a Variables Plus Race Interaction

Tables 26 and 27 show the effects of the race interactions in MAT 0028. None of these interactions were statistically significant (*p*> .05).

Table 26

Estimation of Random Effects for MAT 0028, Model 1d: Race Interaction

Covariance Parameter	*Estimate*	*Standard Error*	*Z Value*	*Pr Z*
Intercept	0.12	0.05	2.31	.01
Residual	1.60	0.08	18.83	<.01

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

Table 27

Estimation of Fixed Effects for MAT 0028, Model 1d: Race Interaction

Fixed Effects	*Estimate*	*Standard Error*	*t Value*	*Pr > |t|*
Intercept	1.79	0.34	5.28	<.01
First Time in College	-0.04	0.10	-0.45	.65
Race-Black	-0.61	0.53	-1.15	.25
Race-Hispanic	-0.08	0.50	-0.17	.87
Race-Other	-0.79	0.99	-0.79	.43
Student Enrollment Status	0.18	0.09	2.02	.04
Student Gender	0.38	0.09	4.24	<.01
Student Age	0.05	0.01	7.78	<.01
PERT score	0.03	0.02	2.06	.04
Race-Black*PERT score	0.02	0.03	0.61	.54
Race-Hispanic*PERT score	0.01	0.02	0.26	.79
Race-Other*PERT score	0.05	0.05	0.98	.33

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

MAT 0028 Model 2: Entering Level 1 Variables and Level 2 Variables

When entering all level 1 and level 2 variables into the model for MAT 0028, student gender (*b*=0.39, *p*<.01), student age (*b*=0.06, *p*<.01), and PERT score (*b*=0.04, *p*<.01) were significant fixed effects. The predictor of course delivery was entered into the model. Because online sections for this course only made up 2% of the sample, the variation was too small for the standard error to be computed.

Table 28

Estimation of Random Effects for MAT 0028, Model 2: Level 1 and Level 2 Variables

Covariance Parameter	*Estimate*	*Standard Error*	*Z Value*	*Pr Z*
Intercept	0.12	0.05	2.31	.01
Residual	1.60	0.09	18.45	<.01

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

Table 29

Estimation of Fixed Effects for MAT 0028, Model 2: Level 1 Variables Plus Level 2 Variables

Fixed Effects	*Estimate*	*Standard Error*	*t Value*	*Pr > |t|*
Intercept	1.35	0.24	5.57	<.01
First Time in College	-0.01	0.10	-0.14	.89
Race-Black	-0.23	0.12	-1.97	.05
Race-Hispanic	-0.06	0.11	0.52	.60
Race-Other	0.16	0.20	0.80	.43
Student Enrollment Status	0.16	0.09	1.74	.08
Student Gender	0.39	0.09	4.31	<.01
Student Age	0.06	0.01	7.73	<.01
PERT score	0.04	0.01	4.30	<.01
Instructor Years	0.01	0.01	0.91	.37
Instructor Enrollment Status	0.11	0.14	0.81	.42
Course Delivery	0	---	---	---
Time of Day of Course	0.13	0.13	1.00	.32

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered. Instructor Employment Status (0=Part-Time, 1=Full-Time). Time of Day (0=Evening, 1=Day). Course Delivery (0=Online, 1=Campus).

MAT 1033 Models

The unconditional model for the Intermediate Algebra course revealed that about 12% of the model was explained by a model with no predictors. We can conclude that the final grade in this course is 2.45 when all residuals are set to 0 (e.g., male, part time, non-first time in college), and from the variance components, that there is more variability within sections (1.75) than among the different sections (0.23).

Table 30

Estimates for MAT 1033 Unconditional Model

Fixed Effect	*Estimate*	*Standard Error*	*t Value*	*Pr > |t|*
Intercept	2.45	0.06	41.92	<.01
Variance Components	*Estimate*	*Standard Error*	*Z Value*	*Pr > Z*
Intercept	0.23	0.10	2.47	.01
Residual	1.75	0.12	14.62	<.01

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

MAT 1033 Model 1a: Entering Student Level Variables

When all student variables were entered into the model, the predictors student enrollment status (b=0.54, p<.01) and student age (b=0.07, p<.01) are significant.

Table 31

Estimation of Random Effects for MAT 1033, Model 1a: Student Level Variables

Covariance Parameter	*Estimate*	*Standard Error*	*Z Value*	*Pr Z*
Intercept	0.14	0.08	1.95	.02
Residual	1.55	0.11	14.36	<.01

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

Table 32

Estimation of Fixed Effects for MAT 1033, Model 1a: Student Level Variables

Fixed Effects	*Estimate*	*Standard Error*	*t Value*	*Pr > |t|*
Intercept	2.32	0.11	21.39	<.01
First Time in College	-0.20	0.12	-1.62	.11
Race-Black	0.04	0.14	0.32	.75
Race-Hispanic	0.13	0.11	1.17	.24
Race-Other	0.16	0.21	0.77	.44
Student Enrollment Status	0.54	0.10	5.26	<.01
Student Gender	0.27	0.10	2.72	.01
Student Age	0.07	0.01	8.93	<.01

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

<u>MAT 1033 Model 1b: Entering Model 1a Variables Plus PERT Score</u>

The PERT slopes are allowed to vary across sections in this model and show that the variance is not significant, indicating that there is no evidence to suggest that the slopes vary by section.

Table 33

Estimation of Random Effects for MAT 1033, Model 1b: PERT Score

Covariance Parameter	*Estimate*	*Standard Error*	*Z Value*	*Pr Z*
Intercept	0	---	---	---
PERT Final Grade Slope	0.0003	0.001	0.27	.40
Error Covariance	-0.003	0.02	-0.15	.88
Residual	1.56	0.11	14.33	<.01

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

The fixed effects for this model reveal that once again, student enrollment status (b=0.54, p<.01) and student age (b=0.07, p<.01) are the only statistically significant predictors.

Table 34

Estimation of Fixed Effects for MAT 1033, Model 1b: PERT Score

Fixed Effects	*Estimate*	*Standard Error*	*t Value*	*Pr > |t|*
Intercept	1.05	0.70	1.49	.14
First Time in College	-0.18	0.12	-1.49	.14
Race-Black	0.05	0.14	0.38	.71
Race-Hispanic	0.14	0.11	1.81	.23
Race-Other	0.16	0.21	0.77	.44
Student Enrollment Status	0.54	0.10	5.27	<.01
Student Gender	0.27	0.10	2.76	.01
Student Age	0.07	0.01	8.86	<.01
PERT score	0.04	0.02	1.82	.07

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

MAT 1033 Model 1c: Entering Model 1a Variables Plus Gender Interaction

This model showed that, similar to the previous models, only student enrollment status (b=0.04, p<.01) and student age were statistically significant predictors. The gender interaction with PERT score was not statistically significant.

Table 35

Estimation of Random Effects for MAT 1033, Model 1c: Gender Interaction

Covariance Parameter	*Estimate*	*Standard Error*	*Z Value*	*Pr Z*
Intercept	0.14	0.08	1.74	.04
Residual	1.57	0.11	14.35	<.01

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

Table 36

Estimation of Fixed Effects for MAT 1033, Model 1c: Gender Interaction

Fixed Effects	*Estimate*	*Standard Error*	*t Value*	*Pr > |t|*
Intercept	1.32	1.05	1.26	.21
First Time in College	-0.17	0.12	-1.44	.15
Race-Black	0.06	0.14	0.42	.69
Race-Hispanic	0.14	0.11	1.18	.24
Race-Other	0.16	0.2113	0.73	.46
Student Enrollment Status	0.54	0.10	5.23	<.01
Student Gender	-0.23	1.39	-0.17	.87
Student Age	0.07	0.01	8.84	<.01
PERT score	0.03	0.03	0.96	.34
PERT score*Student Gender	0.01	0.04	0.36	.72

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

MAT 1033 Model 1d: Entering Model 1a Variables Plus Race Interaction

Model 1d includes all student level variables, in addition to interaction effects for Black, Hispanic and Other race/ ethnicity statuses. The random effects continue to show that only the variability within sections is statistically significant. The interaction effects themselves are not significant in this model and are shown in Table 38.

Table 37

Estimation of Random Effects for MAT 1033, Model 1d: Race Interaction

Covariance Parameter	*Estimate*	*Standard Error*	*Z Value*	*Pr Z*
Intercept	0.15	0.08	1.77	.04
Residual	1.57	0.11	14.31	<.01

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

Table 38

Estimation of Fixed Effects for MAT 1033, Model 1d: Race Interaction

Fixed Effects	*Estimate*	*Standard Error*	*t Value*	*Pr > |t|*
Intercept	1.00	1.08	0.90	.37
First Time in College	-0.18	0.12	-1.47	.14
Race-Black	1.14	1.95	0.58	.56
Race-Hispanic	-0.54	1.60	-0.34	.73
Race-Other	1.92	2.90	0.66	.51
Student Enrollment Status	0.54	0.10	5.21	<.01
Student Gender	0.27	0.10	2.69	.01
Student Age	0.07	0.01	8.85	<.01
PERT score	0.04	0.03	1.27	.21
PERT score* Race-Black	-0.03	0.06	-0.56	.58
PERT score* Race-Hispanic	0.02	0.05	0.43	.67
PERT score* Race-Other	-0.05	0.08	-0.61	.54

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

<u>MAT 1033 Model 2: Entering Level 1 Variables and Level 2 Variables</u>

In this final model run for MAT 1033, all student level variables and course section variables were entered. The only statistically significant fixed effects that emerged in this model were student enrollment status (b=0.53 p<.01) and student age (b=0.08, p<.01). The results for the random and fixed effects of this model are represented in Tables 39 and 40, respectively.

Table 39

Estimation of Random Effects for MAT 1033, Model 2: Level 1 and Level 2 Variables

Covariance Parameter	*Estimate*	*Standard Error*	*Z Value*	*Pr Z*
Intercept	0.07	0.08	0.87	.19
Residual	1.61	0.12	13.74	<.01

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

Table 40

Estimation of Fixed Effects for MAT 1033, Model 2: Level 1 Variables Plus Level 2 Variables

Fixed Effects	*Estimate*	*Standard Error*	*t Value*	*Pr > |t|*
Intercept	1.62	0.79	2.06	.04
First Time in College	-0.08	0.13	-0.63	.53
Race-Black	0.05	0.15	0.37	.71
Race-Hispanic	0.18	0.12	1.48	.14
Race-Other	0.16	0.22	0.75	.45
Student Enrollment Status	0.54	0.11	4.89	<.01
Student Gender	0.29	0.10	2.77	.01
Student Age	0.08	0.01	9.08	<.01
PERT score	0.03	0.02	1.33	.18
Instructor Years	0.001	0.01	0.07	.95
Instructor Enrollment Status	0.16	0.12	1.32	.19
Course Delivery	-0.42	0.27	-1.57	.12
Time of Day of Course	0.06	0.14	0.41	.68

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered. Instructor Employment Status (0=Part-Time, 1=Full-Time). Time of Day (0=Evening, 1=Day). Course Delivery (0=Online, 1=Campus).

MAC 1105 Models

The ICC for MAC 1105 was calculated at .0048, demonstrating that less than one percent of variance is explained in the unconditional model. The addition of other predictors and their effect on the final grade in this course will be explained in a later section.

Table 41

Estimates for MAC 1105 Unconditional Model

Fixed Effect	*Estimate*	*Standard Error*	*t Value*	*Pr > |t|*
Intercept	2.83	0.09	32.21	<.01
Variance Components	*Estimate*	*Standard Error*	*Z Value*	*Pr > Z*
Intercept	0.01	0.18	0.08	0.47
Residual	2.07	0.25	8.13	<.01

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

MAC 1105 Model 1a: Entering Student Level Variables

Estimating the random effects of the first model, which includes all student variables, shows that only the within section variance differs significantly from zero (p<.01).

Table 42

Estimation of Random Effects for MAC 1105, Model 1a: Student Level Variables

Covariance Parameter	*Estimate*	*Standard Error*	*Z Value*	*Pr Z*
Intercept	0.05	0.19	.24	.41
Residual	1.88	0.25	7.41	<.01

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

The fixed effects of this model (Table 43) showed that none of the student level predictors were statistically significant.

Table 43

Estimation of Fixed Effects for MAC 1105, Model 1a: Student Level Variables

Fixed Effects	*Estimate*	*Standard Error*	*t Value*	*Pr > |t|*
Intercept	2.86	0.18	15.78	<.01
First Time in College	-0.35	0.22	-1.60	.11
Race-Black	-0.66	0.27	-2.39	.02
Race-Hispanic	0.19	0.20	0.95	.34
Race-Other	0.41	0.28	1.46	.15
Student Enrollment Status	-0.02	0.18	-0.11	.91
Student Gender	0.37	0.18	2.05	.04
Student Age	0.04	0.01	2.93	.00

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

<u>MAC 1105 Model 1b: Entering Model 1a Variables Plus PERT Score</u>

An unstructured model was run including the PERT score as a random effect. The results of this model show that the variation with the PERT slopes was not significantly different from zero.

Table 44

Estimation of Random Effects for MAC 1105, Model 1b: PERT Score

Covariance Parameter	*Estimate*	*Standard Error*	*Z Value*	*Pr Z*
Intercept	3.08	5.36	0.57	.28
PERT Final Grade Slope	0	0.002	0.09	.47
Error Covariance	-0.04	0.10	-0.35	.73
Residual	1.67	0.24	6.84	<.01

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

The fixed effects show that Race-Black (*b*= -0.74, *p*<.006) and PERT score (*b*=0.04, *p*<.001) were statistically significant predictors in this model.

Table 45

Estimation of Fixed Effects for MAC 1105, Model 1b: PERT Score

Fixed Effects	*Estimate*	*Standard Error*	*t Value*	*Pr > |t|*
Intercept	1.03	0.60	1.72	.09
First Time in College	-0.25	0.21	-1.18	.24
Race-Black	-0.74	0.26	-2.84	.01
Race-Hispanic	0.19	0.20	0.95	.35
Race-Other	0.23	0.26	0.89	.38
Student Enrollment Status	-0.10	0.17	-0.59	.56
Student Gender	0.38	0.17	2.22	.03
Student Age	0.04	0.01	2.56	.01
PERT score	0.04	0.01	3.33	.001

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

MAC 1105 Model 1c: Entering Model 1a Variables Plus Gender Interaction

The random effects for the within section variance was estimated at 1.76 (*p*<.01), indicating that there is more variance within sections than between sections.

Table 46

Estimation of Random Effects for MAC 1105, Model 1c: Gender Interaction

Covariance Parameter	*Estimate*	*Standard Error*	*Z Value*	*Pr Z*
Intercept	0.09	0.20	0.43	.33
Residual	1.76	0.25	7.12	<.01

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

The gender interaction for this model was estimated at (*b*= -0.01, *p*=.72), which indicates that this interaction does not have a statistically significant effect on final grade in MAC 1105.

Table 47

Estimation of Fixed Effects for MAC 1105, Model 1c: Gender Interaction

Fixed Effects	*Estimate*	*Standard Error*	*t Value*	*Pr > |t|*
Intercept	0.30	0.87	0.35	.73
First Time in College	-0.30	0.21	-1.39	.17
Race-Black	-0.67	0.27	-2.50	.01
Race-Hispanic	0.19	0.20	0.94	.35
Race-Other	0.26	0.28	0.92	.36
Student Enrollment Status	-0.06	0.18	-0.36	.72
Student Gender	0.91	1.42	0.64	.52
Student Age	0.04	0.01	2.82	.01
PERT score	0.05	0.02	3.00	.004
PERT score*Student Gender	-0.01	0.03	-0.35	.72

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

MAC 1105 Model 1d: Entering Model 1a Variables Plus Race Interaction

The random and fixed effects for this model are shown in Tables 48 and 49. None of the race interaction effects were statistically significant.

Table 48

Estimation of Random Effects for MAC 1105, Model 1d: Race Interaction

Covariance Parameter	*Estimate*	*Standard Error*	*Z Value*	*Pr Z*
Intercept	0.09	0.20	0.42	.33
Residual	1.77	0.25	6.97	<.01

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

Table 49

Estimation of Fixed Effects for MAC 1105, Model 1d: Race Interaction

Fixed Effects	*Estimate*	*Standard Error*	*t Value*	*Pr > |t|*
Intercept	-0.09	1.03	-0.90	.93
First Time in College	-0.31	0.22	-1.42	.16
Race-Black	2.05	2.21	0.93	.36
Race-Hispanic	0.32	1.74	0.18	.86
Race-Other	1.41	1.95	0.72	.47
Student Enrollment Status	-0.05	0.18	-0.30	.77
Student Gender	0.39	0.18	2.22	.03
Student Age	0.04	0.01	2.77	.01
PERT score	0.06	0.02	2.89	.01
PERT score* Race-Black	-0.06	0.05	-1.24	.22
PERT score* Race-Hispanic	-0.002	0.04	-0.07	.94
PERT score* Race-Other	-0.02	0.04	-0.61	.55

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

MAC 1105 Model 2: Entering Level 1 Variables and Level 2 Variables

The final model in which all student and course section variables were entered demonstrates that none of the student level nor course section variables were statistically significant.

Table 50

Estimation of Random Effects for MAC 1105, Model 2: Level 1 and Level 2 Variables

Covariance Parameter	*Estimate*	*Standard Error*	*Z Value*	*Pr Z*
Intercept	0.18	0.23	0.80	.21
Residual	1.71	0.27	6.26	<.01

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered.

Table 51

Estimation of Fixed Effects for MAC 1105, Model 2: Level 1 Variables Plus Level 2 Variables

Fixed Effects	*Estimate*	*Standard Error*	*t Value*	*Pr > |t|*
Intercept	-1.75	1.60	-1.09	.28
First Time in College	-0.40	0.23	-1.73	.09
Race-Black	-0.49	0.29	-1.66	.10
Race-Hispanic	0.23	0.23	1.03	.31
Race-Other	0.43	0.31	1.42	.16
Student Enrollment Status	0.04	0.20	0.19	.85
Student Gender	0.47	0.20	2.37	.02
Student Age	0.04	0.02	1.99	.05
PERT score	0.05	0.02	3.13	.003
Instructor Years	-0.003	0.01	-0.30	.76
Instructor Enrollment Status	-0.08	0.21	-0.40	.69
Course Delivery	2.31	1.41	1.64	.10
Time of Day of Course	-0.18	0.26	-0.70	.48

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female). PERT score is grand mean centered. Instructor Employment Status (0=Part-Time, 1=Full-Time). Time of Day (0=Evening, 1=Day). Course Delivery (0=Online, 1=Campus).

Research Question One

Research Question One: To what extent does the predictive relationship of the PERT with final grades in Developmental Math I (MAT 0018) and Developmental Math II (MAT 0028) vary based on student gender?

In reviewing Table 14 for Developmental Math I, the relationship between PERT score and gender was estimated at -0.001 (p=.90). The PERT score-gender interaction was estimated at 0.001 (p=.96) for Developmental Math II. There is not sufficient evidence to support that PERT score with final grade varies by gender in MAT 0018 nor in MAT 0028.

Research Question Two

Research Question Two: To what extent does the predictive relationship of the PERT with final grades in Developmental Math I (MAT 0018) and Developmental Math II (MAT 0028) vary based on student race/ethnicity?

Three races were included in the model, each race coded as a binary variable with '1' = Black, '0'= Non-Black, and this was replicated for Hispanic ('1'=Hispanic, '0'=Non-Hispanic), and Other ('1'=Other, '0'=Non-Other). The White race was used a reference group and was therefore left out of the model. Analysis of the fixed effects for MAT 0018, Table 16, shows that there were two significant predictors: student race-Black and student age. The interaction effect of student race-Black by PERT score was estimated at 0.01 (p=.24), the interaction of student race-Hispanic by PERT score was estimated at 0.01 (p=.39), and the interaction of student race-Other by PERT score was estimated at 0.02 (p=.43). While the main effect of race-Black was significant, none of these interactions were significant, leading one to conclude that the relation of the PERT score with final grade in MAT 0018 does not vary by race/ethnicity.

There were no significant findings with the main effects of race, nor with the interactions of races and the PERT score in the model for MAT 0028. As seen in Table 27, race-Black by PERT score was estimated at 0.02 (p=.54), race-Hispanic by PERT score at 0.01 (p=.79), and race-other at 0.05 (p=.33). Given these results there was not sufficient evidence to conclude that the relationship of PERT score with final grade in this course varies based on student race/ ethnicity.

Research Question Three

Research Question Three: To what extent does the predictive relationship of the PERT with final grades in Intermediate Algebra (MAT 1033) and College Algebra (MAC 1105) vary based on student gender?

The fixed model effects for Intermediate Algebra (MAT 1033) are shown in Table 36. These estimates reveal that the interaction of student gender by PERT score is not significant (b= 0.01, p=.72). Similarly, the estimates for College Algebra (MAC 1105) in Table 47 show that the estimate for the student gender by PERT score interaction is not significant (b= -0.01, p=.72). Given these results, there was not sufficient evidence to conclude that the relationship of PERT score with final grade varied based on student race/ethnicity in either of these credit-bearing courses.

Research Question Four

Research Question Four: To what extent does the predictive relationship of the PERT with final grades in Intermediate Algebra (MAT 1033) and College Algebra (MAC 1105) vary based on race/ethnicity?

Table 38 and Table 49 exhibit the fixed model effects for MAT 1033 and MAC 1105, respectively. No significant effects were found for the interactions with Black, Hispanic nor Other students in either of these credit-bearing courses. The relationship of the PERT score with final grades in each of these courses does not vary based on race/ethnicity.

Research Question Five

Research Question Five: What combined student and course section-level variables (i.e., first time in college status, student race/ethnicity, student gender, student enrollment status, student age, instructor employment status, course time of day, course delivery method) are significant *predictors of student* final course grade *in Developmental Math I (MAT 0018)? In Developmental Math II (MAT 0028)?*

To analyze this question, all student-level variables as well as all the course level variables (Instructor years teaching, Instructor employment status, Time of day of course, and course delivery modality) were entered into the model. For MAT 0018, the only mode of delivery was on campus; therefore, this variable was not included in the model. Table 18 shows the fixed effects for student level variables, in addition to course level variables in MAT 0018. Three student level variables emerged as significant predictors—student race-Black (b= -0.56, p<.01), student age (b= 0.04, p<.01), and PERT score (b=0.03, p<.01). None of the course level variables were significant in this model.

Table 29 reveals three significant student level predictor variable for MAT 0028. Student gender (b=0.39, p<.01), student age was estimated at (b=0.06, p<.01), and PERT score (b=0.04, p<.01). None of the course level variables had a significant effect on final grade in upper Developmental Math II.

Research Question Six

Research Question Six: What combined student and course section-level variables (i.e., first time in college status, student race/ethnicity, student gender, student enrollment status, student age, other admission/placement score, instructor employment status, course time of day, course delivery method) are significant predictors of student final course grade *in Intermediate Algebra (MAT 1033)? In College Algebra (MAC 1105)?*

For Intermediate Algebra, Table 40 shows that student enrollment status (b= 0.53, p<.01) and student age (b= 0.08, p<.01) were found to be significant predictors of final course grade in this course. There were no statistically significant course level variables indicated in this model.

College Algebra revealed no significant student level variables, nor any significant course level variables as seen in Table 51.

CHAPTER FIVE

DISCUSSION

This research study examined the predictive relationships of Florida's Postsecondary Education Readiness Test (PERT) with students' grades in various community college math classes. This study also explored if the PERT-class grade relationships varied across course sections and students' gender and race. This chapter presents a summary of the study and its results, discussion, limitations of this study, and recommendations for future research.

SUMMARY OF THE STUDY

This was an institution-specific validity study for the Florida PERT exam performed at a large community college in Florida. The Educational Testing Services (ETS) recommends conducting institution-specific validity studies on a frequent basis in order to ensure the relevance of measurement policies. These types of studies are necessary because the value of said studies make a stronger case when addressing institutional context when discussing the generalizability of validity studies. A number of studies have documented institutional validity results for placement tests (Cleary, 1966; Eskew, 2013; Mzumara & Shermis, 2000; Young & Kobrin, 2001), which support the *Standards for Educational and Psychological Testing*'s recommendation on test validation processes. Cleary's (1966) study examined test bias between Black and White students taking the SAT. Results of that study revealed that there was slight bias in favor of Black students at one particular institution. Eskew's study (2013) analyzed the relationship between the ACCUPLACER and final grades in credit-level math courses, and found that there were significant relationships between the predictors and the criterion. Mzumara and Shermis (2000) conducted annual assessments on the validity of the COMPASS placement tests at Indiana University Purdue University Indianapolis, and found a direct positive correlation between COMPASS math score and success in subsequent math courses. The present study is aligned with these types of studies that have evaluated the validity of placement test scores at specific institutions.

The PERT is a customized placement test for the state of Florida. The previously used ACCUPLACER placement test, from which the cut scores for the PERT were derived, states in its manual that it is good practice to review the cut scores and the validity of the test every three years. The PERT exam was introduced in 2011 as a new mandatory test used to assess college readiness in postsecondary institutions across the state of Florida. The impetus for the search for a new test to measure college readiness was a collective decision among Florida legislators, college, university and K-12 professionals to find a test that more accurately measured readiness competencies for Florida students, specifically. Given that it was a new test, psychometric analysis was warranted to ascertain its predictive ability with regard to courses affected by its scores. The *Standards for Educational and Psychological Testing* (2014) document that there are a number of sources of validity evidence that can be obtained: the one being analyzed in this research study was validity determined from the relationship with other variables. Following the best practices outlined in the *Standards*, empirical evidence was obtained that analyzed the relationships among the PERT score and final grades in four math courses, as well as the relationships with other various predictors (i.e., gender, race/ethnicity, age, enrollment status, and student first-time-in-college status.

This research study was conducted in order to assess the degree to which the PERT score had a relationship to the final grade in four different math courses: Developmental Math I (MAT 0018), Developmental Math II (MAT 0028), Intermediate Algebra (MAT 1033), and College Algebra (MAC 1105). The relevance of this research study stemmed from the creation of a customized postsecondary placement test for the state of Florida. The initial use of this test by community colleges in Florida came at a critical time—the commonly used ACCUPLACER test's contract was coming to an end, and the procurement process for a new contract resulted in a new test developer being commissioned to create a new placement exam. This developer, McCann and Associates, developed a customized placement test set to the standards of Florida's Department of Education.

At the time of this dissertation, there was no available psychometric information on the PERT test. The lack of information on this test's validity led to this research study. Validity evidence on new placement tests is important for any institution because of the need to assess whether or not tests are accurately measuring what they intend to measure. The PERT math assessment test intends to assess the following competencies: equations, evaluating algebraic expressions, polynomials, dividing monomials and binomials, applying standard algorithms, coordinate planes, and focusing on pairs of simultaneous linear equations in two variables. In this case, the purpose of the research was to use final grade in each of the four math courses as a criterion and the PERT score as a predictor. Students who withdrew from the course were not included in the sample. Because there is no specific curriculum for any given course, instructors may deliver course content in any way, and therefore the resulting grades may vary across instructors. A report by The College Board (2015) stated that when it comes to assessing predictive validity, "…it is important that little or no instructional intervention has occurred between the predictor and the start of the course". The rationale for this can in part, be due to the bias that instruction can impart on the efficacy of the intervention. For example, assume that an instructor wants to implement an intervention that helps students understand how to identify and solve polynomials. Drawing the conclusion that a polynomial intervention improves students' ability to do that is stronger if the student has had no previous instruction on polynomials. If the only exposure to polynomials the student has had is via the intervention, then one can assume that is the reason for improvement in understanding this concept.

The initial sample of students consisted of students who had taken the PERT test in the fall of 2012, and then immediately enrolled in a developmental course, Intermediate Algebra, or

College Algebra as a result of their score. Math was the topic of interest due to a larger percentage of students requiring remediation in math rather than in reading or writing.

This study centered around six primary research questions:

1. To what extent does the predictive relationship of the PERT with final grades in Developmental Math I (MAT 0018) and Developmental Math II (MAT 0028) vary based on student gender?
2. To what extent does the predictive relationship of the PERT with final grades in Developmental Math I (MAT 0018) and Developmental Math II (MAT 0028) vary based on student race/ethnicity?
3. To what extent does the predictive relationship of the PERT with final grades in Intermediate Algebra (MAT 1033) and College Algebra (MAC 1105) vary based on student gender?
4. To what extent does the predictive relationship of the PERT with final grades in Intermediate Algebra (MAT 1033) and College Algebra (MAC 1105) vary based on race/ethnicity?
5. What combined student and course section-level variables (i.e., first-time-in-college status, student race/ethnicity, student gender, student enrollment status, student age, instructor employment status, course time of day, course delivery method) are significant predictors of student final course grade in Developmental Math I (MAT 0018)? In Developmental Math II (MAT 0028)?
6. What combined student and course section-level variables (i.e., first-time-in-college status, student race/ethnicity, student gender, student enrollment status, student age, instructor employment status, course time of day, course delivery method) are significant predictors of student final course grade in Intermediate Algebra (MAT 1033)? In College Algebra (MAC 1105)?

DISCUSSION OF THE RESULTS

The main goal of this study was to look at the predictive relationship of the PERT and final grade scores in four different math courses. The PERT was a significant predictor in three of the four courses that were analyzed with higher scores on the PERT being associated with higher math course grades. PERT was a significant predictor in three of the four courses even after controlling for a range of student (e.g., gender, age, part time vs. fill time status, etc.) and instructor/course (online vs. face-to-face delivery, day vs. evening, part time vs. full time instructor) variables. One factor that may have worked against finding a significant relationship between the PERT scores and course grade in MAT 1033 is the restricted range in the PERT scores for this course (PERT scores ranged from 114 to 122 or 9 points). In the other three courses PERT scores had a much larger range: MAT 0018 (PERT scores ranged from 50 to 95 or 46 points); MAT 0028 (PERT scores ranged from 96 to 113 or 18 points); and MAC 1105 (PERT scores ranged from 123 to 150 or 28 points). Overall, the results from three of the courses provide initial evidence of the predictive validity of the PERT scores at the specific community college examined in this study.

In addition to examining the predictive validity of the PERT, this study evaluated potential differential predictive validity. A number of studies have focused on racial/ethnic differential validity with regard to placement tests (Young & Kobrin, 2001), however, most of these studies were dated prior to 2000. Differential prediction results were also identified in a meta-analysis conducted by Fischer et al. (2013), but these too were more than ten years old. Even so, these previous results highlighted varied outcomes with regard to overprediction or underprediction of certain groups of test takers dependent on the institution. Results from validity studies on the ACCUPLACER are more closely related to the PERT, given that the initial cut scores for the PERT were derived from the ACCUPLACER test. Multilevel models were used to evaluate potential interaction effects with the PERT score with race and gender, respectively. No interaction effects were found to be significant, indicating that the PERT test is not predicting course grades differently based on race or gender (i.e., no evidence of differential predictive validity). Following the analyses focusing on the predictive validity and differential predictive validity of the PERT, I examined the relationship of other student and course-related variables and course grades.

Results for MAT 0018 revealed that overall, Black students had lower final grades than White students and older students tended to have higher final grades. PERT was a significant predictor for this course, which meant that students with a higher PERT score had significantly higher final grades. In MAT 0028, female students had higher final grades than male students, older students did better, and the PERT score had a significant effect on the final grade.

In MAT 1033, students who registered as full-time had higher final grades than part-time students, and older students did better than their younger counterparts. Finally, in MAC 1105, older students fared better in Model 1a, Black students had significantly lower grades than White students in Model 1b, and in the final model, only the PERT score had a significant effect, showing a positive relationship with final grade.

Significant predictor variables varied across the courses, but it seemed that, overall, whether a student was first-time-in-college had no effect on the final grade in any course. The predictors of Race-Hispanic and Race-Other had no effects as predictors at all in any of the courses.

The selection of the variables used in the models was developed in part from a number of research studies on the predictive validity of placement tests. Studies by Eskew (2013) and Sireci and Talento-Miller (2006) evaluated predictive validity using variables that could directly affect student academic performance (i.e., high school GPA, SAT, or ACT scores). However, for this study, including those variables would have greatly diminished the sample size due to the inconsistency with how they are reported and the large amount of missing data for these variables. Disaggregating postsecondary student data by factors such as race/ethnicity, gender, age, or enrollment status is a common occurrence for many postsecondary institutions.

This type of categorization allows policy-makers to determine if there is any population of students that need specific attention, a strategy that is often a high priority among community colleges. With this information, interventions could be put in place that focus on increasing the success of students in that one specific student demographic.

A number of student and course level variables were entered into the models in this research study to determine their relation to the final grade in Developmental Math I, Developmental Math II, College Algebra and Intermediate Algebra. While the relationships of these variables were not the primary focus of this research, results did show that for the Developmental Math I course, the race category of being Black and the age predictor were significant predictors of the grade in the class. For Developmental Math II, the gender and age predictors were significant predictors of final grades in the class. For the College Algebra Course, enrollment status (fulltime vs. part time), gender and age emerged as significant predictors, and in the Intermediate Algebra course the race predictor (Black) and age were significant predictors.

In conclusion, this research study found that the PERT score was a significant predictor three of the four of the courses studied. The only course where PERT score had no significant relationship was in Intermediate Algebra (MAT 1033), where the PERT scores had a restricted range. It is important that more research be conducted on this test and that psychometric information is published based on all colleges and schools that require this test for college placement.

LIMITATIONS OF THE STUDY

As a predictive validity study, one limitation of this research was that no other information was available from other institutions. The researcher had access only to one college in the state of Florida, and no comparative information on PERT performance was available for other colleges in the state of Florida at the time of the publication of this study. Generalizability to other institutions could be compromised because of the lack of existing data on this test. Also, because the baseline cut scores of the PERT were aligned with the previously used placement test, the ACCUPLACER, it is difficult to understand if these results are specific to the PERT or if ACCUPLACER results would be similar. Because it is a customized placement test, there are no national benchmarks with which to make comparisons.

A second limitation was the unavailability of item level information for the PERT. Availability of item level information would have been valuable in reviewing the internal consistency reliability of the item scores and the quality of the items and distractors, which can give a more thorough view of the validity of the PERT test itself. The Standards (2014) also state that assessing validity requires looking at test content, response processes, and internal structure. Test content information, as previously indicated, was not available for analysis. Likewise, student cognitive practices exhibited by the student during responding was not able to be analyzed. The internal structure, which looks at constructs like dimensionality and factor structure, would be a necessary examination of the PERT test. Item analysis and cut score analysis was impossible because these data were not available to the researcher. Because the baseline cut scores of the PERT were aligned with the previously used placement test, the ACCUPLACER, it is difficult to ascertain whether the current results were specific to the PERT or the cut scores that had been developed for the ACCUPLACER but were applied to the PERT scores. For example, the PERT could be fine, but the cut scores could be off. Alternatively, the cut scores could be fine, but the test is not actually measuring the construct it intends to measure.

In addition to measuring the relationships of the student level predictors and course grades, the researcher was also concerned that there may be instructor variability, which may lead to variation in students' final grades. Because there were no significant course-section variables that emerged as significant, one can assume that instructor variability is not significant enough to warrant an effect on the relationship between the PERT and final grade. Regardless of the lack of significant course level variables, an additional area of concern is that there could be problems with the way in which the instructors give course grades. Factors such as extra credit, homework, quizzes, and class attendance are all factors that could contribute to the final course grade. Because this is not consistent across instructors, it could affect the outcome variable, which could in turn, question the validity of the final grade in the course. However, even given these inconsistent grading strategies, the relationship between the PERT and final grade was still positive and significant.

RECOMMENDATIONS FOR FUTURE RESEARCH

The implementation of the PERT came at a critical time in Florida educational history. With the passing of Senate Bill 1720, enrollment in developmental education may have been altered due to the fact that certain students are no longer required to enroll in developmental courses if they 1) entered the 9th grade in a Florida public high school in 2003-04 or after, and graduated with a Florida high school diploma; or 2) are serving as an active duty military from the United States Armed Services. Individual colleges in Florida have already begun the task of researching how SB 1720 has affected success rates in credit bearing courses, and the results are not promising. One such college found that enrollments in college level courses increased 25%, however the success rate (student earning A, B, or C) for those same classes dropped almost 10%. Another college measured success rates for students testing in developmental courses and enrolling in developmental courses versus the success rates of students testing into developmental courses and not enrolling into a developmental course. The results revealed that of this population, only 20% of students who opted to take the college-level course actually passed with a C or better (https://www.insidehighered.com/news/2015/06/25/floridas-remedial-law-leads-decreasing-pass-rates-math-and-english). Future studies on the PERT and its predictive validity could take a view at the statewide community college trends that are occurring in both developmental math and college-level math courses in the wake of Senate Bill 1720.

Because this study only analyzed the relationship between PERT scores and final grades in developmental courses in math and the first two math credit bearing courses, future studies could incorporate the other areas of developmental education to see how the PERT relates to outcomes in reading and writing. While the percentages of students enrolling in developmental reading and writing courses were smaller when compared to those enrolling in developmental math, there could still be effects on these courses.

Another area of interest is in the area of item analysis of the PERT. There was a lack of psychometric data available at time of this study publication, and the test publisher was unwilling to share any item-level data, however, item analysis is an important way of evaluating if items are functioning differently for different groups (DIF). Collaborating with PERT publishers is an important step in understanding this test. A cut score analysis is critical in order to verify that placement into courses is accurate. The consequences to correct course placement means that both students and advisors are more informed to make better decisions when deciding which courses to take. Once item analysis is conducted, a further step in the analysis could be to examine the relationships of PERT scores with other predictors such as other placement scores (ACT, SAT) and high school GPA. Though not always available in community colleges, these variables can provide another source of validity in assessing the PERT. This would also offer the prospect of reviewing concordance tables with the PERT, ACCUPLACER, ACT and SAT tests to set equivalent scores among these frequently used tests, and make comparisons among them.

It is also important to note that the range of scores were restricted in each sample. For Developmental Math I (MAT 0018), the scores ranged from 50 to 95; for Developmental Math II (MAT 0028), the scores ranged from 96 to 113; for Intermediate Algebra (MAT 1033), from 114 to 122; and for College Algebra, from 123 to 150. The fact that the samples were restricted by PERT score at such unequal ranges may have decreased the correlations between the predictors and the criterion. The sample sizes themselves may also have had an effect on the magnitude of the correlations. As noted by Goodwin and Leech (2006), the correlations on smaller samples are more susceptible to change than larger samples. The largest sample size was 900 for MAT 0028, and the smallest was 271 for MAC 1105. Future studies should strive for more equal sample sizes so that power could be increased.

Another area of interest would be to review how closely the course competencies relate to the content areas the PERT purports to measure. Finally, it would be important to investigate how different student level variables or course level variables relate to course outcomes. Enrollment variables collected by institutions vary, and other variables can include high school GPA, success in high school math courses, or scores on other placement exams (I.e., ACT or SAT). Generalizability of the study is an important aspect of in assessing PERT test effects on course outcomes. The more this test is used, and the more research conducted will only deepen the body of knowledge for measurement of this customized placement test.

CLOSING REMARKS

Success in developmental math is a serious issue, not only in Florida, but nationwide. Preparing students for college by ensuring that they have the necessary skills and abilities to be successful in critical courses continues to be a goal of postsecondary institutions across the country. Nationally used tests that gauge college readiness have the benefit of enabling benchmarks, but also are not able to address issues that may be specific to certain demographics. The usage of customized tests, such as the PERT, have both benefits and disadvantages. It is the job of administrators in higher education to weigh the pros and cons of such tests, and make decisions that are most advantageous for its students. The findings revealed in this study can be shared with committees tasked with exploring methods to streamline developmental education, identifying factors affecting course success, or ensuring placement exams are accurately placing students. These discussions can hopefully lead to more data-driven decision making.

Review of test score validity is important in higher education institutions because of the consequences these placement exams have on the students. Placement test outcomes can directly affect student financial aid, time to degree completion, and inform academic advisors on whether or not the student has the course knowledge necessary to be successful. Because these are deemed important factors to consider, the data derived from placement tests are necessary to make informed decisions.

Because of the implications of Senate Bill 1720, students not mandated to take placement tests at postsecondary institutions in Florida may view course placement procedures more as a tool to advise them of their potential success or failure in a college level course. If students test into developmental and opts not to take the developmental course, this could have repercussions on their time to degree if they can never pass the college-level courses. More stringent quantitative research investigating the PERT test is warranted with a larger number of institutions, possibly different student and course variables. Item analysis, reliability and validity are all areas of concern with any new test, and analyses when first establishing a new placement test are extremely important. Educators and students alike will benefit from this research.

REFERENCES

ACCUPLACER Setting Cut Scores. Retrieved from http://media.collegeboard.com/digitalServices/pdf/accuplacer/accuplacer-method-for-setting-cut-scores.pdf

ACT (2007). *ACT Technical Manual.* Iowa City, IA: ACT.

Allen, J. (2005). Grades as valid measures of academic achievement of classroom learning. *The Clearing House, 78*(5), 218-223.

American Association of Community Colleges. (2016). AACC About Community Colleges. Retrieved from http://www.aacc.nche.edu/AboutCC/Pages/default.aspx

American Educational Research Association, American Psychological Association, National Council on Measurement in Education (AERA, APA, NCME) (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association, American Psychological Association, National Council on Measurement in Education.

Attewell, P., Lavin, D., Domina, T. & Levey, T. (2006). New evidence on college remediation. *The Journal of Higher Education, 77*(5), 886-924.

Bailey, T. (2008). *Challenge and Opportunity: Rethinking the role and function of developmental education in community college*. New York, NY: Columbia University, Community College Research Center.

Bailey, T., & Cho, S. (2010). *Issue Brief: Developmental Education in Community Colleges*. Paper prepared for White House Summit on Community Colleges Conference, Washington, D.C. Retrieved from http://ccrc.tc.columbia.edu/media/k2/attachments/developmental-education-community-colleges.pdf

Bailey, T. Jeong, D., & Cho S. (2009). *Referral, enrollment and completion in developmental education sequences in community colleges*. New York, NY: Columbia University, Teachers College, Community College.

Bendig, A. W. (1953). The reliability of letter grades. *Educational and Psychological Measurement, 13*(2), 311-321.

Berk, R. (1986). A Consumer's Guide to Setting Performance Standards on Criterion-Referenced Tests. *Review of Educational Research*, 56(1), 137-172.

Brown, C. & Lancashire, H. (2013). *College readiness: Postsecondary Education Readiness Test (P.E.R.T.) overview and course advisement*. Retrieved from http://georgejenkinshs.com/old/academics/pdf/PERT%20Information%20(PP)%202012.pdf

Burdman, P. (2011). *Testing ground: How Florida schools and colleges are using a new assessment to increase college readiness.* Boston, MA: Jobs for the Future.

Camara, W., & Echternacht, G. (2000). *The SAT I and high school grades: Utility in predicting success in college*. New York: College Entrance Examination Board.

Camara, W., & Millsap, R. (1998). *Using the PSAT/NMSQT and course grades in predicting success in the advanced placement program*. New York: College Entrance Examination Board.

Camilli, G., & Shepard, L. (1994). *Methods for identifying biased test items*. Thousand Oaks, CA: SAGE Publications.

Carolina Journal (2014). Community college remediation rates still problematic. Retrieved from https://www.carolinajournal.com/opinion-article/community-college-remediation-rates-still-problematic/

Casazza, M. (1999). Who are we and where did we come from? *Journal of Developmental Education, 23*(1), 2-4, 6-7.

Chinn, R. (2006). *Considerations in Setting Cut Scores*. Retrieved from http://www.clearhq.org/resources/Cut_Scores_RB_2006.pdf

Cizek, G. (1996). Setting passing scores [NCME instructional module]. *Educational Measurement: Issues and Practice, 15*(2), 20-31.

Clark, B. (1960). The “Cooling-Out” function in higher education. *The American Journal of Sociology, 65*(6), 569-576

Cleary, T.A. (1966). Test bias: *Validity of the Scholastic Aptitude Test for Negro and White students in integrated colleges*. Princeton, NJ: Educational Testing Service.

Cleary, T. A. (1968). Test bias: Prediction of grades of Negro and White student in integrated colleges. *Journal of Educational Measurement, 5*(2), 115-124.

Cohen, A. M., & Brawer, F. B. (2003). *The American Community College* (4th ed.). San Francisco: Jossey-Bass.

College Board (2014). *ACCUPLACER Program Manual*. New York, NY.

Collins, M. (2008). *It’s Not About the Cut Score: Redesigning Placement Assessment Policy to Improve Student Success*. Policy Brief by Achieving the Dream

Conley, D. T. (2011). *Redefining college readiness, Volume 5*. Eugene, OR: Educational Policy Improvement Center.

Crocker, L., & Algina, J. (1986). *Introduction to classical & modern test theory*. Orlando: Harcourt Brace Jovanovich College Publishers.

Day, C. (1997). *A Predictive Validity Study of Computer Adaptive Placement Test for Tennessee Higher Education Institutions* (Doctoral Dissertation). Retrieved from ProQuest (UMI No. 9840290).

Deil-Amen, R., & Rosenbaum, J. (2002). The unintended consequences of stigma-free remediation. *Sociology of Education, 75,* 249-268.

Eskew, R. (2013). ACCUPLACER math placement test scores, SAT-Math scores and high school grade averages as predictors of MA145 and MA200 final grades. Retrieved from https://www.hilbert.edu/docs/default-source/Academics/Institutional-Research-Assessment/oira_predicting-math-course-grades.pdf?sfvrsn=0

Fischer, F., Schult, J., & Hell, B. (2013). Sex-specific differential prediction of college admission tests: A meta-analysis. *Journal of Educational Psychology, 105*(2), 478-488.

Florida College System. *Interim Cut Score Setting for the Postsecondary Education Readiness Test*. Retrieved from http://www.fldoe.org/core/fileparse.php/7724/urlt/0072380-pert.pdf

Florida College System. *What levels of developmental education do students test into? To what extent do students emerge from developmental education?* Retrieved from http://www.fldoe.org/core/fileparse.php/7724/urlt/0083520-fcsdevedlevelstapp.pdf

Florida College System (2014). *How many students enroll in developmental education in the Florida College System and in what academic areas do they study to build their skills?* Retrieved from http://www.fldoe.org/fcs/pdf/q2.pdf.

Florida College System (2010). Florida's Postsecondary Education Readiness Test (2010, November). *Zoom*, 3, p.1-9.

Florida Comprehensive Assessment Test. Retrieved from http://www.fldoe.org/accountability/assessments/k-12-student-assessment/history-of-fls-statewide-assessment/fcat/

Florida Comprehensive Assessment Test (FCAT) Briefing Book 2007. Retrieved from http://fcat.fldoe.org/pdf/BriefingBook07web.pdf

Florida Department of Education Press Office Memo. Retrieved from http://fldoe.org/newsroom/latest-news/204517-with-students-as-top-priority-florida-chooses-replacement-for-fcat.stml

Florida Rule 6A-10.0315. Retrieved from https://www.flrules.org/gateway/ruleno.asp?id=6A-10.0315

Florida Rule 6A-10.0318. Retrieved from https://www.flrules.org/gateway/notice_Files.asp?ID=11410903

Florida Standards Assessments (2014). Commissioner's Decision for New Florida Standards Assessments, Q and A. Retrieved from http://www.fldoe.org/core/fileparse.php/3/urlt/qa-03-17.pdf

Florida Statute 1008.30 (2014). Retrieved from http://www.leg.state.fl.us/statutes/index.cfm?App_mode=Display_Statute&URL=1000-1099/1008/Sections/1008.30.html

Go Higher Florida Task Force Final Recommendations. Retrieved from http://www.flbog.edu/documents_meetings/0003_0072_0505_4b.pdf

Goodwin, L., & Leech, N. (2006). Understanding correlation: Factors that affect the size of r. *Journal of Experimental Education 74*(3), 251-266.

Large Florida Community College Factbook (2013). Retrieved from https://www.hccfl.edu/media/937952/factbook%202013-%20bookmarks.pdf

Hodara, M., Jaggars, S., & Karp, M. (2012). *Improving developmental education assessment and placement: Lessons from community colleges across the country* (CCRC Working Paper No. 51). New York, NY: Columbia University, Teachers College, Community College Research Center.

Hughes, K. L., & Scott-Clayton, J. (2011). *Assessing developmental education assessment in community colleges* (CCRC Working Paper No. 19, Assessment of Evident Series). New York, NY: Columbia University, Teachers College, Community College Research Center.

Hudson, L. M. (2015). *What factors of motivation predict achievement of college readiness? A study of self-determination and college readiness* (Doctoral Dissertation). Retrieved from ProQuest. (UMI No. 3704477).

LaForte, F. (2000). *A Correlational Study Regarding COMPASS' Ability to Predict GPA at a Western Wisconsin Technical College*. Unpublished manuscript, University of Wisconsin-Stout.

Le, C., Rogers, K., & Santos, J. (2011). *Innovations in Developmental Math: Community College Enhance Support for Nontraditional Students*. Retrieved from http://www.jff.org/sites/default/files/MetLife-DevMath-040711.pdf

Lee, V. (2000). Using hierarchical linear modeling to study social contexts: *The case of school effects. Educational Psychologist, 35*(2), 125-141.

Linn, R. (1973). Fair test use in selection. *Review of Educational Research, 43*(2), 139-161.

Linn, R. (1982). Admissions testing on trial. *American Psychologist*, 37(3), 279-291.

Linn, R. (1984). Selection bias: Multiple meanings. *Journal of Educational Measurement*, *21*(1), 33-47.

Linn, R. (1990). Admissions Testing: Recommended uses, validity, differential prediction, and coaching. *Applied Measurement in Education*, *3*(4), 297-318.

Livingston, S. A., & Zieky, M.J. (1982). Passing scores: A manual for setting standards of performance on educational and occupational tests. Princeton, NJ: ETS.

Ma, X., Ma, L., & Bradley, K. (2008). Using multilevel modeling to investigate school effects. In A.A. O'Connell & D.B. McCoach (Eds.), *Multilevel modeling of educational data* (pp. 245-272). Charlotte, NC: Information Age Publishing Inc.

Mattern, K. D. & Packman, S. (2009). *Predictive validity of ACCUPLACER scores for course placement: A meta-analysis* (Research Report No. 2009-2). New York, NY: College Board. Retrieved from http://research.collegeboard.org/sites/default/files/publications/2012/7/researchreport-2009-2-predictive-validity-accuplacer-scores-course-placement.pdf

Meade, A., & Fetzer, M. (2009). Test bias, differential prediction, and a revised approach for determining the suitability of a predictor in a selection context. *Organizational Research Methods, 12*(4), 738-761.

Messick, S. (1989). *Validity*. In R.L. Linn (Ed.), *Educational measurement* (3rd. ed), 13-104. New York: MacMillan.

Mzumara, H., Shermis, M. & Fogel, M. (1998). *Validity of the IUPUI placement test scores for course placement: 1997-1998.* Indianapolis, IN: Indianapolis University Purdue University Indianapolis Testing Center.

Mzumara, H., & Shermis, M. (2001). *Predictive Validity of Placement Test Scores for Course Placement at IUPUI: Summer and Fall 2000*. Indianapolis: Indiana University Purdue University Indianapolis Testing Center.

Morgan, D., & Michaelides, M. (2005). *Setting cut scores for college placement*. New York: The College Board.

National Association for Developmental Education. Retrieved from http://www.nade.net/aboutdeved.html

Noble, J., Crouse, J. & Schulz, M. (1996). *Differential prediction/impact in course placement for ethnic and gender groups.* ACT Research Report Series 96-8. Iowa City, IA: American College Testing.

Palomba, C., & Banta, T. (1999). Assessment Essentials: Planning, Implementing, and Improving Assessment in Higher Education. San Francisco: Jossey-Bass.

Patterson, B., & Mattern, K.D. (2013). *Validity of the SAT for predicting first-year grades: 2011 SAT validity sample* (Statistical Report No. 2013-3). New York, NY: The College Board.

Patterson, B., Mattern, K.D. & Kobrin, J. (2009). *Validity of the SAT for predicting first-year grades: 2007 SAT validity sample* (Statistical Report No. 2009-1). New York, NY: The College Board.

Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction*. London: Wadsworth.

Perin, D. (2006). *Can community colleges protect both access and standards? The problem of remediation*. New York, NY: Columbia University, Teachers College, Community College Research Center.

Perry, M., Bahr, P. R., Rosin, M., & Woodward, K. M. (2010). *Course-taking patterns, policies, and practices in developmental education in the California Community Colleges*. Mountain View, CA: EdSource.

Prather, J. R. (2007). *Hierarchical linear models of teacher effectiveness* (Doctoral Dissertation). Retrieved from ProQuest. (UMI No. 3293579).

Prince, H. (2005). Standardization vs flexibility: State policy options on placement testing for development education in community colleges. Retrieved from http://www.jff.org/sites/default/files/publications/StandardFlex.pdf

Quint, J., Jaggars, S., Byndloss, D. & Magazinnik, A. (2013). *Bringing developmental education to scale: Lessons from the developmental education initiative*. Retrieved from http://www.mdrc.org/sites/default/files/Bringing%20Developmental%20Education%20to%20Scale%20FR.pdf

Raudenbush, S. W. & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods, second edition*. Newbury Park, CA: Sage.

SAT Advising and Admission Handbook. (2015-16). Retrieved from http://media.collegeboard.com/digitalServices/pdf/professionals/sat-advising-and-admission-handbook-2015-16.pdf

Sawyer, R. (2007). Indicators of usefulness of test scores. *Applied Measurement in Education*, 20:3, 255-271.

Sireci, S., & Talento-Miller, E. (2006). Evaluating the predictive validity of graduate management admission test scores. *Educational and Psychological Measurement 66*(2), 305-317.

Texas Public Higher Education Almanac 2015. Retrieved from http://www.uh.edu/provost/houstongps/_documents/THECB2015almanac.pdf

Verbout, M. (2013). *Predictive Validity of a Multiple- Choice Test for Placement in a Community College* (Doctoral Dissertation). Retrieved from ProQuest. (UMI No. 3592228).

Weber, D. (2001, February). *Restriction of range: The truth about consequences and corrections*. Paper presented at the Annual Meeting of the Southwest Educational Research Association, New Orleans, LA.

Young, J., & Kobrin, J. (2001). Differential validity, differential prediction, and college admission testing: A comprehensive review and analysis (Research Report No. 2001-6). New York, NY: College Board.

Appendix A: PERT Subject Area Assessments

Mathematics
• Equations- solving linear equations, linear inequalities, quadratic equations and literal equations
• Evaluating algebraic expressions
• Polynomials – factoring, simplifying, adding, subtracting, multiplying and dividing
• Dividing by monomials and binomials
• Applying standard algorithms or concepts
• Coordinate planes – translating between lines and inspect equations
• Focusing on pairs of simultaneous linear equations in two variables
Reading
• Discerning and summarizing the most important ideas, events, or information
• Supporting or challenging assertions about the text
• Determining the meaning of words and phrases in context
• Analyzing the meaning, word choices, tone and organizational structure of the text
• Determining the author's purpose and the relation of events in the text to one another
• Recognizing relationships within and between sentences
• Analyzing the traits, motivations and thoughts of individuals in fiction and nonfiction
• Analyzing how two or more texts with different styles, points of view or arguments address similar topics or themes
• Distinguishing between facts and opinions
• Evaluating reasoning and rhetoric of an argument or explanation
Writing
• Sustaining focus on a specific topic or argument
• Establishing a topic or thesis
• Demonstrating use of the conventions of standard written English, including grammar, usage and mechanics
• Supporting and illustrating arguments and explanations
• Developing and maintaining a style and tone
• Synthesizing information from multiple relevant sources
• Conveying complex information clearly and coherently
• Representing and accurately citing data, conclusions, and opinions of others
• Establishing a substantive claim and acknowledging competing arguments or information
• Conceptual and Organizational Skills – recognizing effective transitional devices within the context of a passage
• Word Choice Skills – recognizing commonly confused or misused words and phrases
• Sentence Structure Skills – using modifiers correctly; using coordination and subordination effectively; and recognizing parallel structure
• Grammar, Spelling, Capitalization and Punctuation Skills - avoiding inappropriate shifts in verb tense and pronouns; maintaining agreement between pronoun and antecedent; and using proper case forms, adjectives and adverbs

Appendix B: Percentile Ranks for Admission and Placement Tests

National Distributions of Cumulative Percents for ACT Test Scores
ACT-Tested High School Graduates from 2013, 2014 and 2015

Score	ENGLISH	Usage/Mechanics	Rhetorical Skills	MATHEMATICS	Pre-Algebra/Elem.Alg.	Alg./Coord. Geometry	Plane Geometry/Trig.	READING	Soc. Studies/Sciences	Arts/Literature	SCIENCE	COMPOSITE	Score
36	99			99				99			99	99	36
35	99			99				99			99	99	35
34	98			99				98			99	99	34
33	97			98				97			98	99	33
32	95			97				95			97	98	32
31	93			96				92			96	96	31
30	92			95				89			95	95	30
29	90			93				86			94	92	29
28	88			91				84			92	90	28
27	85			88				81			90	87	27
26	82			84				78			87	83	26
25	79			78				75			83	79	25
24	74			73				71			77	74	24
23	69			67				66			70	68	23
22	64			62				61			63	63	22
21	58			57				55			56	56	21
20	52			53				48			48	50	20
19	45			49				42			40	43	19
18	40	99	99	43	99	99	99	36	99	99	33	36	18
17	36	96	98	37	96	99	99	31	97	96	27	30	17
16	32	91	97	27	92	98	99	25	92	91	22	24	16
15	27	87	92	15	88	96	95	21	87	85	17	18	15
14	21	83	86	6	82	90	89	16	81	79	13	12	14
13	16	78	79	2	76	82	82	11	75	73	9	7	13
12	13	72	71	1	67	72	74	7	67	66	6	4	12
11	10	66	61	1	58	63	63	4	57	58	4	1	11
10	7	56	51	1	50	51	52	2	49	49	2	1	10
9	4	46	41	1	42	37	40	1	38	41	1	1	9
8	2	37	31	1	34	24	26	1	29	32	1	1	8
7	1	29	21	1	24	14	15	1	19	24	1	1	7
6	1	21	15	1	11	8	9	1	11	17	1	1	6
5	1	14	10	1	5	4	6	1	5	10	1	1	5
4	1	8	5	1	2	2	3	1	2	5	1	1	4
3	1	4	2	1	1	1	2	1	1	2	1	1	3
2	1	1	1	1	1	1	1	1	1	1	1	1	2
1	1	1	1	1	1	1	1	1	1	1	1	1	1
Mean	20.3	10.1	10.4	20.9	10.7	10.6	10.4	21.3	10.9	10.7	20.8	21.0	
S.D.	6.6	4.0	3.5	5.3	3.6	3.0	3.1	6.4	3.6	3.9	5.4	5.4	

Note: These national norms are the source of U.S. Ranks, for multiple-choice tests, displayed on ACT reports during the 2015-2016 testing year.

These norms with a sample size of 5,569,466, are based on 2013, 2014 and 2015 graduates.

Appendix B, cont'd.

SAT Percentile Ranks

Critical Reading, Mathematics, and Writing

Score	Critical Reading	Mathematics	Writing
800	99	99	99+
790	99	99	99+
780	99	99	99
770	99	99	99
760	99	98	99
750	98	98	99
740	98	97	98
730	97	97	98
720	96	96	97
710	96	95	97
700	95	93	96
690	94	92	95
680	93	91	94
670	92	89	93
660	90	88	92
650	89	86	90
640	87	83	89
630	85	81	87
620	83	79	85
610	82	76	83
600	79	74	81
590	77	71	79
580	74	68	76
570	71	66	73
560	68	63	71
550	65	60	68
540	62	56	64
530	58	53	62
520	55	50	58
510	51	47	54
500	48	43	51
490	44	40	47
480	41	36	44
470	37	33	40
460	34	30	37
450	31	27	33
440	27	25	30
430	25	22	27
420	22	19	23
410	19	16	20
400	17	14	18
390	15	12	15
380	12	11	13
370	10	9	11
360	9	7	9
350	7	6	7
340	6	5	6
330	5	4	5
320	4	4	4
310	4	3	3
300	3	3	3
290	3	2	2
280	2	2	2
270	2	1	1
260	1	1	1
250	1	1	1
240	1	1	1
230	1	1	1
220	1	1	1-
210	1	1-	1-
200	–	–	–
Mean	503	518	497
SD	113	115	109

Appendix C: ACCUPLACER Content Areas: Arithmetic, Elementary Algebra, and College-Level Math

Arithmetic Content Areas
Addition, subtraction, multiplication, and division of whole numbers with 1 to 4 digits
Addition and subtraction of mixed numbers
Multiplication of fractions
Division of fractions
Division of a whole number by a fraction
Division of a fraction by a whole number or another fraction
Applications involving operations on two numbers
Square root and exponent operations
Addition and subtraction of decimals
Multiplication of decimals
Division of decimals
Multiplication and division of decimals
Ordering of decimals, fractions, and percents; rounding
Calculate the percentage of a number
Applications
Fractions, ratios, and proportions
Calculating percentages
Adding and subtracting multiple fractions
Application of the greatest common factor and least common multiple
Calculate the average (mean)
Interpret frequency graphs
Problem solving using whole numbers, fractions, and decimals
Items that have a negative stem
Addition of whole numbers, fractions, and decimals
Subtraction and repeated subtraction of whole numbers, fractions, and decimals
Multiplication of whole numbers, fractions, and decimals
Division of whole numbers, fractions, and decimals
Metric system units
English system units
Currency
Computation with mixed numbers

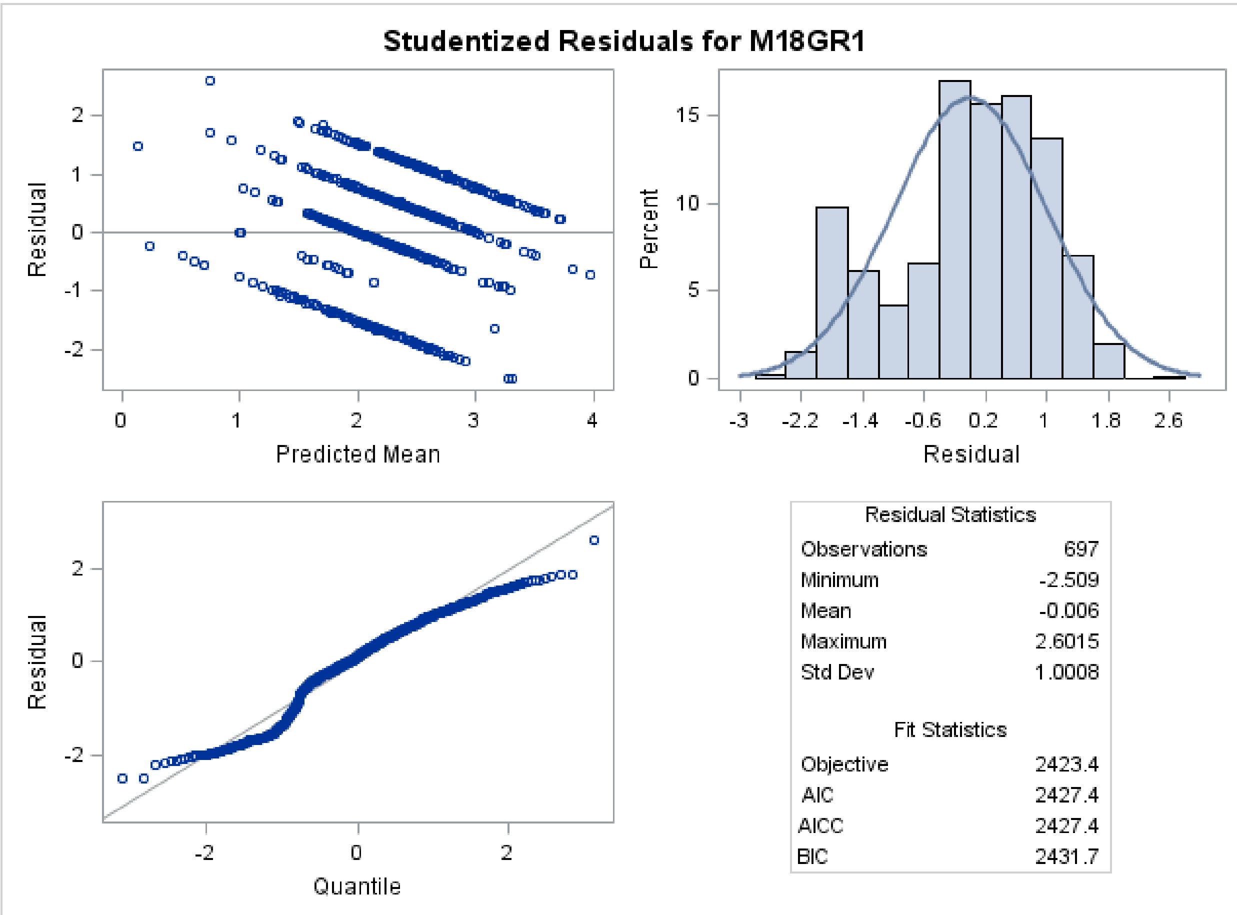
Appendix C, cont'd.

Elementary Algebra Content Area
Integers and Rationals
Ordering
Operations with signed numbers
Absolute value
Algebraic Expressions
Evaluating formulas and other algebraic expressions
Addition and subtraction of monomials and polynomials
Multiplication of monomials and polynomials
Positive rational roots and exponents
Squaring a binomial
Factoring difference of squares
Factoring $ax^2 + bx + c$ over the integers
Factoring polynomials that are not quadratics
Operations with algebraic fractions involving addition, subtraction, multiplication, and division
Division of monomials and polynomials including simplification of algebraic fractions
Equations, Inequalities, and Word Problems
Solving linear equations and inequalities
Systems of linear equations
Solving quadratic equations by factoring
Translating written phrases or sentences into algebraic expressions or equations
Solving verbal problems in an algebraic context including geometric reasoning
Graphing

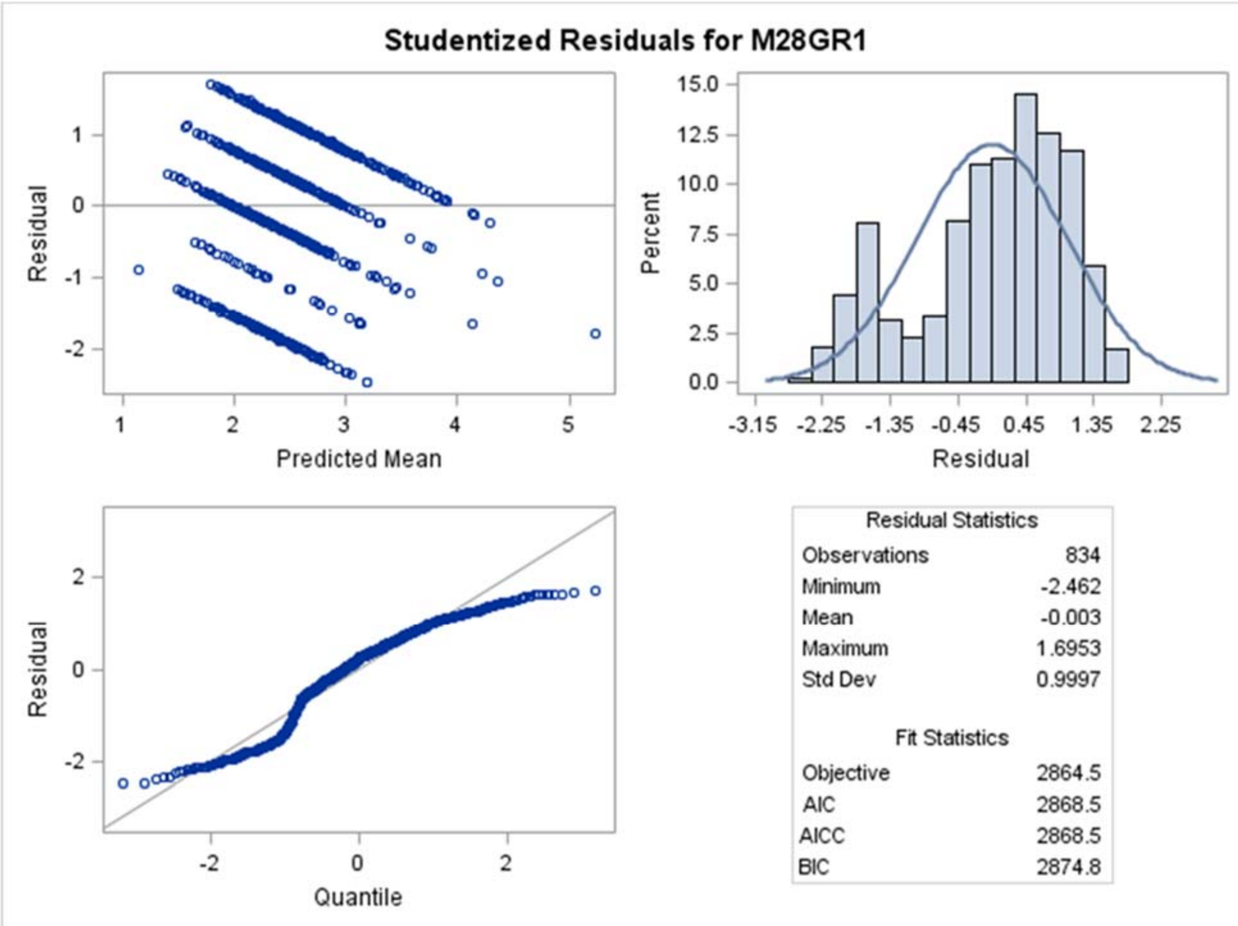
Appendix C, cont'd.

College-Level Math Content Area
Algebraic Operations
Addition of algebraic fractions
Addition and subtraction of expressions involving absolute value
Operations with polynomials
Multiplication, division, and simplification of algebraic fractions
Operations with exponents
Powers, roots, radicals
Factoring quadratic expressions
Solution of Equations and Inequalities
Linear equations and inequalities
Quadratic equations
Systems of equations and inequalities
Exponential equations
Equations of degree greater than 2
Coordinate Geometry
The coordinate plane
Straight lines
Conics
Locus of points
Graphs of algebraic functions
Applications and Other Algebra Topics
Translation
Complex numbers
Series and sequences
Determinants
Permutations and combinations
Factorials
Polygons
Functions
Functions of degree greater than 2
Exponents and logarithms
Graphical properties, exponential and logarithmic functions
Domain and range
Composition of functions
Inverse functions
Computations with simple functions
Periodicity, amplitude, and other properties
Trigonometry
Fundamental definitions of trig functions
Right triangle trigonometry and circular functions
Laws of sines and cosines
Graphs of trigonometric functions
Trigonometric equations and inequalities
Trigonometric identities
Trigonometric functions of two angles
Inverse trigonometric functions

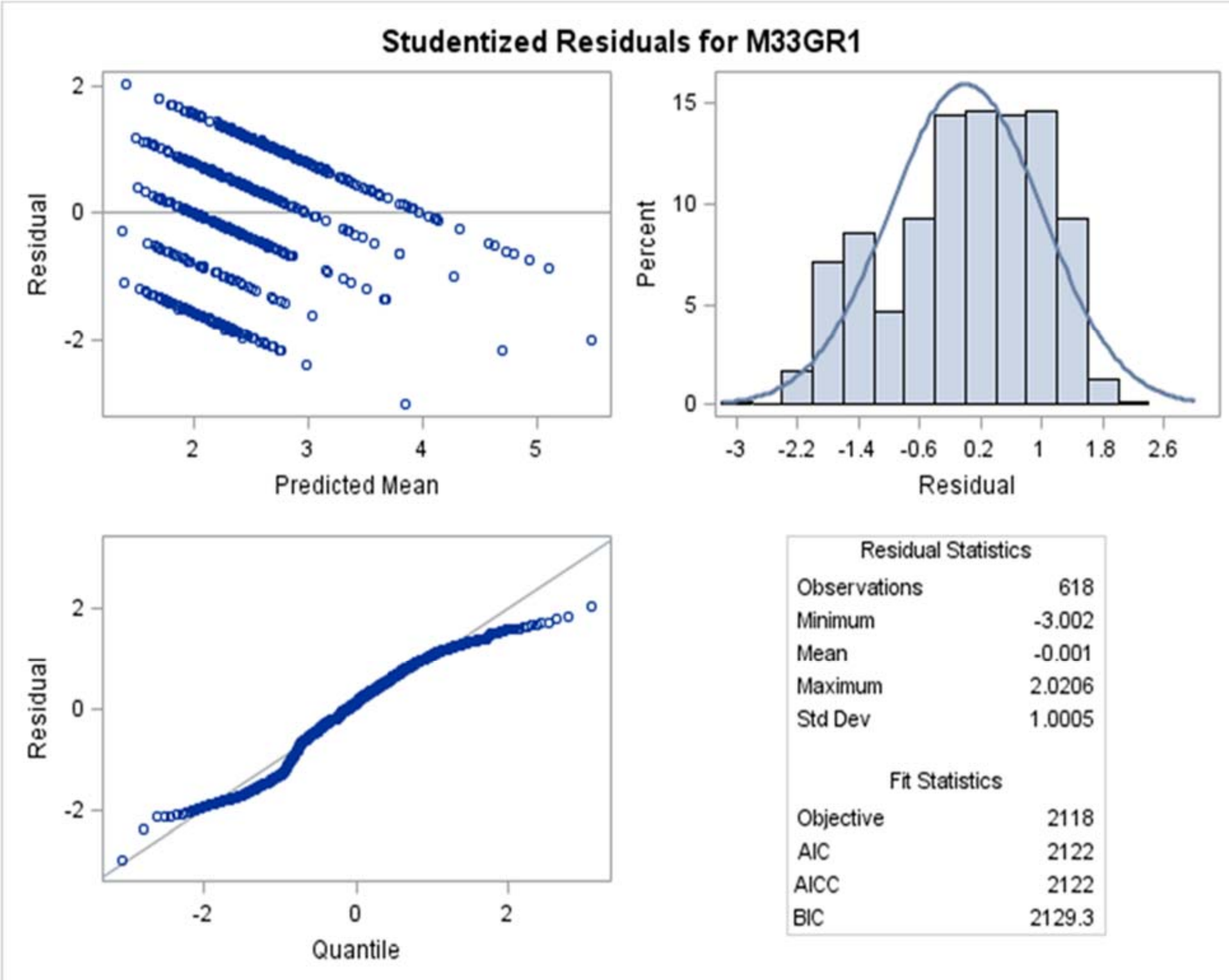
Appendix D: Studentized Residual Graphs for All Courses



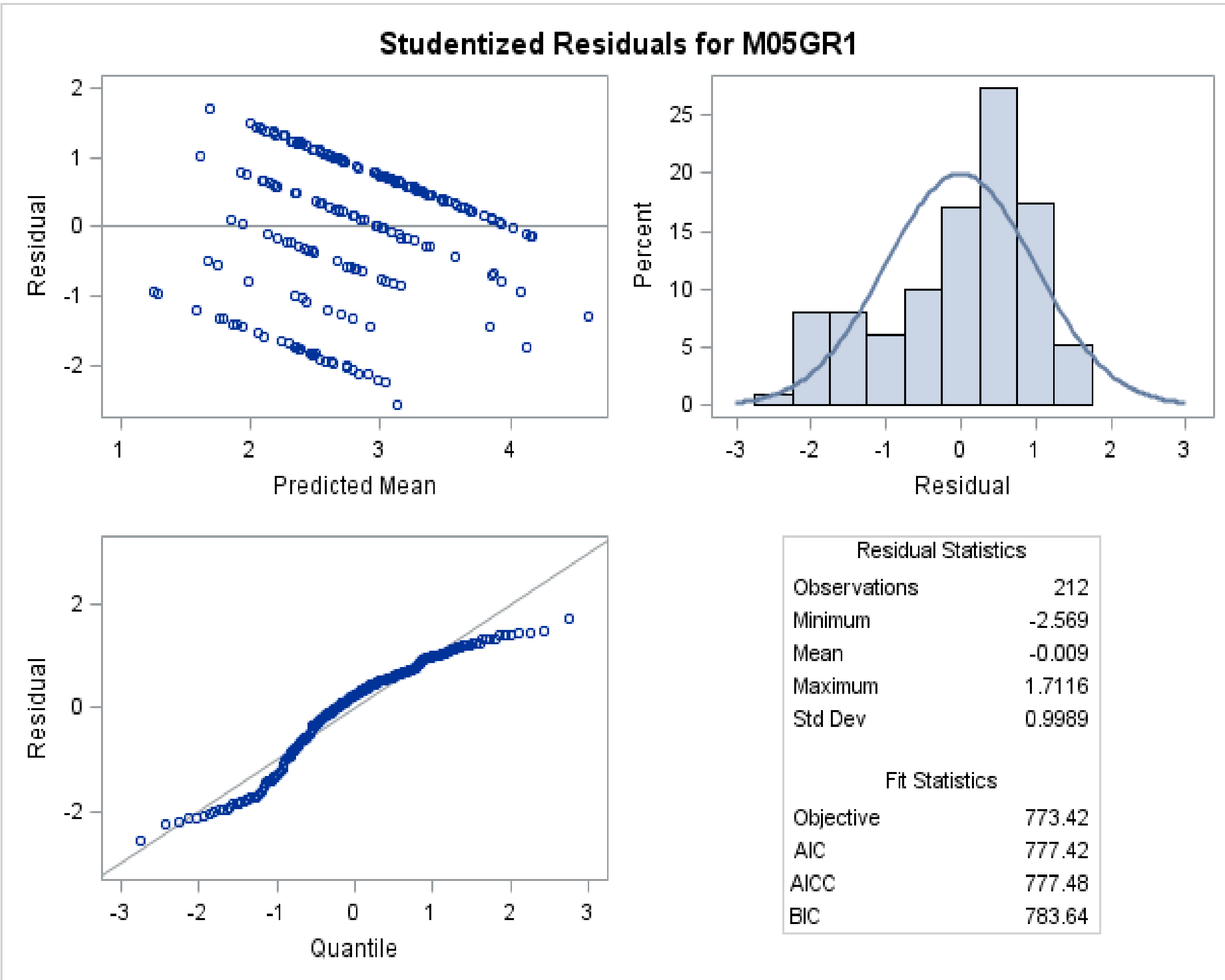
Appendix D, cont'd.



Appendix D, cont'd.



Appendix D, cont'd.



Appendix E: Summary of Hierarchical Regression Variables Predicting Postsecondary Final Course Grade in MAT 0018

MAT0018

	Unconditional Model		Model 1a		Model 1b		Model 1c		Model 1d		Model 2	
Fixed Effects	*b*	*SE b*	*b*	*SE b*	*b*	*SE b*	*b*	*SE b*	*b*	*SE b*	*b*	*SE b*
Intercept	2.27*	0.07	2.61*	0.13	2.56*	0.13	2.54*	0.13	2.57*	0.13	2.36*	0.19
First Time in College Status			-0.08	0.11	-0.10	0.11	-0.11	0.11	-0.11	0.11	-0.14	0.11
Race_Black			-0.65*	0.13	-0.60*	0.12	-0.59*	0.12	-0.61*	0.12	-0.56*	0.12
Race_Hispanic			-0.21	0.13	-0.17	0.12	-0.17	0.12	-0.19	0.12	-0.17	0.12
Race_Other			-0.25	0.28	-0.17	0.27	-0.20	0.27	-0.21	0.27	-0.21	0.27
Enrollment Status			-0.15	0.11	-0.22	0.10	-0.20	0.10	-0.20	0.10	-0.22	0.11
Gender			0.06	0.11	0.13	0.10	0.14	0.10	0.13	0.10	0.14	0.10
Age			0.03*	0.01	0.04*	0.01	0.04*	0.01	0.04*	0.01	0.04*	0.01
PERT Score					0.03*	0.01	0.04*	0.01	0.03*	0.01	0.03*	0.01
PERT*Gender							-0.01	0.01				
PERT*Black									0.01	0.01		
PERT*Hispanic									0.01	0.01		
PERT*Other									0.02	0.03		
Instructor Years											0.01	0.01
Instructor Status											0.13	0.15
Course Delivery Method											---	---
Time of Day of Course											0.09	0.15
Random Effects												
Residual Variance	1.90*	0.10	1.76*	0.10	1.67*	0.09	1.67*	0.09	1.67*	0.09	1.66*	0.09
Intercept Variance	0.12	0.05	0.14	0.06	0.11	0.05	0.10	0.05	0.10	0.05	0.11	0.05
PERT Final Grade Slope					0	---						
Error Covariance					-0.004	0.002						

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female).
******significant at p<.01*

Model 1a: Score = $\beta_0 + \beta_1(Black) + \beta_2(Hispanic) + \beta_3(Other) + \beta_4(Gender) + \beta_5(Enrollment\ Status + \beta_6(Age) + \beta_7(FTIC) + R_{ij}$
Model 1b: Score = Model 1a variables + $\beta_8\ (PERT\ score) +\ R_{ij}$
Model 1c: Score = Model 1a variables + $\beta_9\ (Gender\ x\ PERT\ Score) +\ R_{ij}$
Model 1d: Score = Model 1a variables + $\beta_{10}\ (Black\ x\ PERT\ Score) + \beta_{11}\ (Hispanic\ x\ PERT\ Score) + \beta_{12}\ (Other\ x\ PERT\ Score) +\ R_{ij}$
Model 2: Score = All student variables + course section variables

Appendix F: Summary of Hierarchical Regression Variables Predicting Postsecondary Final Course Grade in MAT 0028

MAT0028

	Unconditional Model		Model 1a		Model 1b		Model 1c		Model 1d		Model 2	
Fixed Effects	*b*	*SE b*	*b*	*SE b*	*b*	*SE b*	*b*	*SE b*	*b*	*SE b*	*b*	*SE b*
Intercept	2.48*	0.05	2.45*	0.11	1.63*	0.23	1.64*	0.32	1.79*	0.34	1.35*	0.24
First Time in College Status			-0.06	0.10	-0.05	0.10	-0.05	0.10	-0.04	0.10	-0.01	0.10
Race_Black			-0.31	0.12	-0.30	0.12	-0.30	0.12	-0.61	0.53	-0.23	0.12
Race_Hispanic			0.05	0.11	0.04	0.11	0.04	0.11	-0.08	0.50	-0.06	0.11
Race_Other			0.18	0.20	0.18	0.20	0.16	0.20	-0.79	0.99	0.16	0.20
Enrollment Status			0.19	0.09	0.19	0.09	0.18	0.10	0.18	0.09	0.16	0.10
Gender			0.39*	0.09	0.38*	0.09	0.36	0.41	0.38*	0.09	0.39*	0.10
Age			0.05	0.01	0.06*	0.01	0.06*	0.01	0.06*	0.01	0.06*	0.01
PERT Score					0.04*	0.01	0.04*	0.02	0.03*	0.02	0.04*	0.01
PERT*Gender							0.001	0.02				
PERT*Black									0.02	0.03		
PERT*Hispanic									0.01	0.02		
PERT*Other									0.05	0.05		
Instructor Years											0.01	0.01
Instructor Status											0.11	0.14
Course Delivery Method											0	---
Time of Day of Course											0.13	0.13
Random Effects												
Residual Variance	1.78*	0.09	1.62*	0.09	1.60*	0.08	1.60*	0.08	1.60*	0.09	1.60*	0.09
Intercept Variance	0.13*	0.06	0.13*	0.05	0.53	0.28	0.12	0.05	0.12*	0.05	0.12	0.05
PERT Final Grade Slope					0	---						
Error Covariance					-0.01	0.01						

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female).
*****_significant at p<.01_

Model 1a: Score = $\beta_0 + \beta_1(Black) + \beta_2(Hispanic) + \beta_3(Other) + \beta_4(Gender) + \beta_5(Enrollment\ Status + \beta_6(Age) + \beta_7(FTIC) + R_{ij}$
Model 1b: Score = Model 1a variables + $\beta_8\,(PERT\ score) +\ R_{ij}$
Model 1c: Score = Model 1a variables + $\beta_9\,(Gender\ x\ PERT\ Score) +\ R_{ij}$
Model 1d: Score = Model 1a variables + $\beta_{10}\,(Black\ x\ PERT\ Score) + \beta_{11}\,(Hispanic\ x\ PERT\ Score) + \beta_{12}\,(Other\ x\ PERT\ Score) +\ R_{ij}$
Model 2: Score = All student variables + course section variables

Appendix G: Summary of Hierarchical Regression Variables Predicting Postsecondary Final Course Grade in MAT 1033

MAT1033

	Unconditional Model		Model 1a		Model 1b		Model 1c		Model 1d		Model 2	
Fixed Effects	*b*	*SE b*	*b*	*SE b*	*b*	*SE b*	*b*	*SE b*	*b*	*SE b*	*b*	*SE b*
Intercept	2.45*	0.06	2.32*	0.11	1.05	0.71	1.32	1.05	0.97	1.08	1.62	0.79
First Time in College Status			-0.20	0.12	-0.18	0.12	-0.17	-0.12	-0.18	0.12	-0.08	0.13
Race_Black			0.05	0.14	0.05	0.14	0.06	0.14	1.14	1.95	0.05	0.15
Race_Hispanic			0.13	0.11	0.14	0.11	0.14	0.11	-0.54	1.60	0.18	0.12
Race_Other			0.16	0.21	0.16	0.21	0.16	0.21	1.92	2.90	0.16	0.22
Enrollment Status			0.54*	0.10	0.54*	0.10	0.54*	0.10	0.54*	0.10	0.54*	0.11
Gender			0.27	0.10	0.27	0.10	-0.23	1.39	0.27*	0.10	0.29	0.10
Age			0.07	0.01	0.07*	0.01	0.07*	0.01	0.07*	0.01	0.08*	0.01
PERT Score					0.04	0.02	0.03	0.03	0.04	0.03	0.03	0.02
PERT*Gender							0.01	0.04				
PERT*Black									-0.03	0.06		
PERT*Hispanic									0.02	0.05		
PERT*Other									-0.05	0.08		
Instructor Years											0.00	0.01
Instructor Status											0.16	0.12
Course Delivery Method											-0.42	0.27
Time of Day of Course											0.06	0.14
Random Effects												
Level One Variance	1.75*	0.12	1.56*	0.11	1.56*	0.11	1.57*	0.11	1.57	0.11	1.61*	0.12
Intercept Variance	0.24*	0.10	0.16	0.08	0.00	---	0.14	0.08	0.15	0.08	0.07	0.08
PERT Final Grade Slope					0.00	0.00						
Error Covariance					0.00	0.00						

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female).
significant at p<.01

Model 1a: Score = $\beta_0 + \beta_1(Black) + \beta_2(Hispanic) + \beta_3(Other) + \beta_4(Gender) + \beta_5(Enrollment\ Status + \beta_6(Age) + \beta_7(FTIC) + R_{ij}$
Model 1b: Score = Model 1a variables + $\beta_8\,(PERT\ score) +\ R_{ij}$
Model 1c: Score = Model 1a variables + $\beta_9\,(Gender\ x\ PERT\ Score) +\ R_{ij}$
Model 1d: Score = Model 1a variables + $\beta_{10}\,(Black\ x\ PERT\ Score) + \beta_{11}\,(Hispanic\ x\ PERT\ Score) + \beta_{12}\,(Other\ x\ PERT\ Score) +\ R_{ij}$
Model 2: Score = All student variables + course section variables

Appendix H: Summary of Hierarchical Regression Variables Predicting Postsecondary Final Course Grade in MAC 1105

MAC1105

	Unconditional Model		Model 1a		Model 1b		Model 1c		Model 1d		Model 2	
Fixed Effects	*b*	*SE b*	*b*	*SE b*	*b*	*SE b*	*b*	*SE b*	*b*	*SE b*	*b*	*SE b*
Intercept	2.83*	0.09	2.86*	0.18	1.03	0.60	0.30	0.87	-0.09	1.03	-1.75	1.60
First Time in College Status			-0.35	0.22	-0.25	0.21	-0.30	0.21	-0.31	0.22	-0.40	0.23
Race_Black			-0.66	0.27	-0.74*	0.26	-0.67	0.27	2.05	2.21	-0.49	0.29
Race_Hispanic			0.19	0.20	0.19	0.20	0.19	0.20	0.32	1.74	0.23	0.23
Race_Other			0.41	0.28	0.23	0.26	0.26	0.28	1.41	1.95	0.43	0.31
Enrollment Status			-0.02	0.18	-0.10	0.17	-0.06	0.18	-0.05	0.18	0.04	0.20
Gender			0.37	0.18	0.38	0.17	0.91	1.42	0.39	0.18	0.47	0.20
Age			0.04*	0.01	0.03	0.01	0.04*	0.01	0.04	0.01	0.03	0.02
PERT Score					0.04*	0.01	0.05*	0.02	0.06	0.02	0.05*	0.02
PERT*Gender							-0.01	0.03				
PERT*Black									-0.06	0.05		
PERT*Hispanic									0.00	0.04		
PERT*Other									-0.02	0.04		
Instructor Years											-0.00	0.01
Instructor Status											-0.09	0.21
Course Delivery Method											2.31	1.41
Time of Day of Course											-0.18	0.26
Random Effects												
Level One Variance	2.07*	0.25	1.88*	0.25	1.67*	0.24	1.76*	0.25	1.77*	0.25	1.71*	0.27
Intercept Variance	0.01	0.18	0.05	0.19	3.08	5.36	0.09	0.20	0.09	0.21	0.18	0.23
PERT Final Grade Slope					0.00	0.00						
Error Covariance					-0.04	0.1						

Note. Student age is grand mean centered. First time in College (0=No, 1=Yes). Student Enrollment Status (0=Part-Time, 1=Full-Time). Student Gender (0=Male, 1=Female).
******significant at p<.01*

Model 1a: Score = $\beta_0 + \beta_1(Black) + \beta_2(Hispanic) + \beta_3(Other) + \beta_4(Gender) + \beta_5(Enrollment\ Status + \beta_6(Age) + \beta_7(FTIC) + R_{ij}$
Model 1b: Score = Model 1a variables + $\beta_8\,(PERT\ score) + \ R_{ij}$
Model 1c: Score = Model 1a variables + $\beta_9\,(Gender\ x\ PERT\ Score) + \ R_{ij}$
Model 1d: Score = Model 1a variables + $\beta_{10}\,(Black\ x\ PERT\ Score) + \beta_{11}\,(Hispanic\ x\ PERT\ Score) + \beta_{12}\,(Other\ x\ PERT\ Score) + \ R_{ij}$
Model 2: Score = All student variables + course section variables

Appendix I: Institution IRB Approval Letter

[redacted]

[redacted]

[redacted]

October 4, 2016

Alisa Murphy Žujović, M.S.
[redacted]

RE: [redacted] IRB #2016_005
TITLE: Evaluation of the Predictive Validity of the Postsecondary Education Readiness Test Using Multilevel Modeling

Dear Ms. Žujović:

On October 4, 2016, the [redacted] Institutional Review Board (IRB) determined that your research meets [redacted] requirements and federal criteria for expedited status which includes activities that (1) present no more than minimal risk to human subjects [21 CFR 56.110], and (2) involve only procedures listed in one or more of the categories listed below:

(6) Collection of data from voice, video, digital, or image recordings made for research purposes.

As the Principal Investigator for this project at [redacted], it is your responsibility to ensure that this research is conducted as detailed in your University of South Florida IRB application and supporting documents and consistent with the ethical principles outlined in the Belmont Report and with [redacted] policies and procedures. Please note that modifications to the research design must be reported to the [redacted] IRB prior to implementing any changes.

The [redacted] IRB will maintain your research proposal and expedited status approval for a period of one year from the date of this letter. If you wish to continue this research beyond one year, you must submit a request for continuing review at least 60 days prior to the expiration date. If you complete the research prior to the end of the one-year period, you must submit a request to close the study. Please note that it is your responsibility to notify the IRB of the status of this study no

Appendix I, cont'd.

later than one year from the date of this letter, or upon completion of the research, whichever is sooner.

If you have any questions concerning this information, please contact me at [redacted] or by email at [redacted].

Best wishes in your research endeavors.

Sincerely,

[redacted]

[redacted], Ph.D., Chairperson
[redacted] Institutional Review Board

Appendix J: USF IRB Approval Letter

USF
UNIVERSITY OF SOUTH FLORIDA

RESEARCH INTEGRITY AND COMPLIANCE
Institutional Review Boards, FWA No. 00001669
12901 Bruce B. Downs Blvd., MDC035 • Tampa, FL 33612-4799
(813) 974-5638 • FAX(813)974-7091

11/2/2016

Alisa Zujovic
Educational and Psychological Studies
[redacted]

RE: **Expedited Approval for Initial Review**
IRB#: Pro00028095
Title: Predictive validity of the Postsecondary Education Readiness Test

Study Approval Period: 11/2/2016 to 11/2/2017

Dear Ms. Zujovic:

On 11/2/2016, the Institutional Review Board (IRB) reviewed and **APPROVED** the above application and all documents contained within, including those outlined below.

Approved Item(s):
Protocol Document(s):
A. ZUJOVIC Proposal 10-18-16 CLEAN.docx

It was the determination of the IRB that your study qualified for expedited review which includes activities that (1) present no more than minimal risk to human subjects, and (2) involve only procedures listed in one or more of the categories outlined below. The IRB may review research through the expedited review procedure authorized by 45CFR46.110 and 21 CFR 56.110. The research proposed in this study is categorized under the following expedited review category:

(5) Research involving materials (data, documents, records, or specimens) that have been collected, or will be collected solely for nonresearch purposes (such as medical treatment or diagnosis).

Your study qualifies for a waiver of the requirements for the informed consent process for a retrospective record review as outlined in the federal regulations at 45CFR46.116 (d) which states that an IRB may approve a consent procedure which does not include, or which alters, some or all of the elements of informed consent, or waive the requirements to obtain informed consent provided the IRB finds and documents that (1) the research involves no more than

Appendix J, cont'd.

minimal risk to the subjects; (2) the waiver or alteration will not adversely affect the rights and welfare of the subjects; (3) the research could not practicably be carried out without the waiver or alteration; and (4) whenever appropriate, the subjects will be provided with additional pertinent information after participation.

As the principal investigator of this study, it is your responsibility to conduct this study in accordance with IRB policies and procedures and as approved by the IRB. Any changes to the approved research must be submitted to the IRB for review and approval via an amendment. Additionally, all unanticipated problems must be reported to the USF IRB within five (5) calendar days.

We appreciate your dedication to the ethical conduct of human subject research at the University of South Florida and your continued commitment to human research protections. If you have any questions regarding this matter, please call 813-974-5638.

Sincerely,

John A Schinka, Ph.D.

John Schinka, Ph.D., Chairperson
USF Institutional Review Board

Appendix K: USF IRB Continuing Review Letter

USF
UNIVERSITY OF SOUTH FLORIDA

RESEARCH INTEGRITY AND COMPLIANCE
Institutional Review Boards, FWA No. 00001669
12901 Bruce B. Downs Blvd., MDC035 • Tampa, FL 33612-4799
(813) 974-5638 • FAX(813)974-7091

10/23/2017

Alisa Zujovic
Educational and Psychological Studies
[redacted]

RE: **Expedited Approval for Continuing Review**
IRB#: CR1_Pro00028095
Title: Predictive validity of the Postsecondary Education Readiness Test

Study Approval Period: 11/2/2017 to 11/2/2018

Dear A. Zujovic:

On 10/21/2017, the Institutional Review Board (IRB) reviewed and **APPROVED** the above application and all documents contained within including those outlined below.

Approved Item(s):
Protocol Document(s):
A. ZUJOVIC Proposal 10-18-16 CLEAN.docx

The IRB determined that your study qualified for expedited review based on federal expedited category number(s):

(5) Research involving materials (data, documents, records, or specimens) that have been collected, or will be collected solely for nonresearch purposes (such as medical treatment or diagnosis).

As the principal investigator of this study, it is your responsibility to conduct this study in accordance with USF HRPP policies and procedures and as approved by the USF IRB. Any changes to the approved research must be submitted to the IRB for review and approval by an amendment. Additionally, all unanticipated problems must be reported to the USF IRB within five (5) calendar days.

We appreciate your dedication to the ethical conduct of human subject research at the University of South Florida and your continued commitment to human research protections. If you have

Appendix K, cont'd.

any questions regarding this matter, please call 813-974-5638.

Sincerely,

Mark Ruiz, PhD, Vice Chairperson
USF Institutional Review Board

ABOUT THE AUTHOR

Alisa Murphy Žujović was born and raised in Texarkana, TX. Upon graduating from Texas Senior High school, she was the recipient of a full track scholarship to Northeast Louisiana University (NLU) in Monroe, LA. In 1999, she received her Bachelor of Arts in Psychology from NLU, and in 2004, her Master of Science degree in Psychology from the newly named University of Louisiana-Monroe. She began working at a large Florida-based Community College in 2004, first as a Research Analyst, then as an Academic Assessment Officer. In 2017, she was promoted to Director of Institutional Research and Grants. She currently serves as chair of the college’s Institutional Review Board, as well as an evaluator on the registry for the Southern Association of Colleges and Schools Commission on Colleges (SACS-COC).

Alisa is a volunteer track and field coach for the Phase 1 Track Club in Tampa, FL, and is a member of USA Track and Field, holding Level I certification in the sport. She is a Diamond Life member of Delta Sigma Theta Sorority, Inc., makes frequent trips to her ‘second home’ in Belgrade, Serbia, and enjoys spending time with her family and friends.