STATISTICAL ANALYSIS OF STUDENT PERFORMANCE IN REDESIGNED DEVELOPMENTAL MATHEMATICS COURSES

A thesis presented to the faculty of the Graduate School of
Western Carolina University in partial fulfillment of the requirements
for the degree of Master of Science in Applied Mathematics

By

Malgorzata J. Chockla

Director: Dr. John Wagaman, Mathematics

Committee Member: Dr. Kathy Jaqua, Mathematics
Committee Member: Dr. Jeffrey Lawson, Mathematics

November 19, 2013

ACKNOWLEDGEMENTS

I would like to thank my thesis advisor, Dr. John Wagaman, for his guidance and dedication as well as his valuable suggestions on my research. I am also grateful to my committee members, Dr. Kathy Jaqua and Dr. Jeffrey Lawson, for the support and encouragement I received during my years of studying at WCU.

Thank you to my friends and colleagues at Southwestern Community College for your support and help with the preparation and data collection stages of my research. I would like to express my appreciation to Dr. Barbara Putman and Ms. Mary Bradley for their help and encouragement.

Thank you to Joshua Hiller for your help with my graduate work and Master's thesis. I value and appreciate your advice and friendship. I also appreciate the help and encouragement from WCU faculty and fellow graduate students.

I would like to thank my husband for his unconditional love and support. You and our daughters Natalie and Emily have been a great source of strength for me and the strongest motivation to complete this educational journey. Also, thank you to my brother Pawel for being there for me throughout the years.

CONTENTS

LIST OF FIGURES

ABSTRACT

STATISTICAL ANALYSIS OF STUDENT PERFORMANCE IN REDESIGNED DEVELOPMENTAL MATHEMATICS COURSES

Malgorzata Justyna Chockla, M.S.

Western Carolina University (November 2013)

Director: Dr. John Wagaman, Mathematics

Colleges and universities are focusing their efforts on improving the instruction in developmental mathematics courses. In 2013, community colleges in North Carolina were in the process of implementing the redesigned approach to teaching developmental mathematics with the goal of improving student graduation rates. The purpose of this study is to investigate the differences in academic improvements of developmental mathematics students in order to evaluate the effectiveness of the redesigned MyMathLab (MML) courses. This study investigates the following variables: College Placement Test (CPT) scores in Algebra, Arithmetic, Reading Comprehension, and Sentence Structure, gender, and instructional method. Multiple regression analyses were performed using the statistical computing software, `R`. In Phase I of this study, two linear regression models were developed to predict student academic improvement in MML and Educo developmental mathematics courses using the standardized CPT scores, gender, and methodological indicator as potential predictors. In Phase II, three linear models were analyzed to predict student academic performance in redesigned MML classes. Using the data from Phase II, three additional regression models were developed with the MML post-test as the response variable and the CPT scores, gender, and the MML pre-test as the set of possible predictors to identify "at-risk" students. An out-of-sample prediction method was used to evaluate the misclassification rate in identifying "at-risk" students.

The results of this study suggest that Algebra and Arithmetic CPT scores are significant predictors of student academic improvement. However, for each model, less than 50% of the variability in student improvement is explained by the linear relationship between the variables. Based on the results of this study, students enrolled in module 050 MML classes at Southwestern Community College (SCC) showed greater improvement than Educo students. Furthermore, the predictive models that include CPT scores and gender as the only predictors of learning can be employed to identify "at-risk" students at the beginning of each school semester. In the conclusion of the thesis, the limitations and implications of this study are discussed.

1 INTRODUCTION

Purpose and Significance of the Study

Colleges and universities are focusing their efforts on improving the instruction in developmental mathematics courses and helping students achieve their educational goals. According to the Community College Research Center (CCRC), in 2012 approximately 60% of all community college students were enrolled in at least one developmental education course. It is alarming that in 2012 only about 17% of students referred to the lowest level of developmental mathematics completed the requirements and passed one college-level mathematics course [Jaggars et al., 2013]. In an attempt to remedy this situation, 15 community colleges in 6 states, including Southwestern Community College (SCC) in North Carolina, redesigned their developmental mathematics instruction and piloted redesigned courses.

The academic improvement of lower-performing students in developmental mathematics classes has been a long-standing problem. Extensive research has been conducted to determine the most effective instructional methods to achieve better learning outcomes and improve graduation rates [Scott-Clayton, 2012, Zachary, 2008, Department of Education, 2005]. Some research has been published on using statistical models to predict student academic improvement [Emerson and Taylor, 2004, Huang, 2011, Lowis and Castley, 2008, Spradlin, 2009]. This improvement is affected by many cognitive and non-cognitive variables such as prior mathematical and technology knowledge, prior achievement, learning style, motivation, goals, and student demographics [Hendricks, 2012, Loving, 2007, Lowis and Castley, 2008]. The choice of the predictor variables and the modeling approach are vital in developing the most accurate predictive models, but these models are often limited by the goals and the design of the study.

The purpose of this study is to investigate the differences in the academic improvement of SCC developmental mathematics students in order to evaluate the effectiveness of the redesigned MyMathLab (MML) courses. The redesigned courses consist of eight 4-week modules that replace the three traditional 16-week developmental mathematics classes: MAT060, 070, and 080. In Spring 2012, Pearson, an educational services provider, released a pilot version of the modularized MML courses aligned with the North Carolina Community College System (NCCCS) reform guidelines. In Phase I of this study, redesigned MML courses offered at SCC in Spring 2012 were compared with traditional Educo courses. The main differences between the redesigned MML and traditional Educo courses included changes in the curriculum, course delivery format, and use of technology. In Phase II, developmental students were placed into 16-week "shell" classes: MAT060, 070, and 080, however, the redesigned MML courses were used for all developmental mathematics instruction and the assessment of student learning. This study investigates the following variables that might affect developmental students' learning: College Placement Test (CPT) scores in Algebra, Arithmetic, Reading Comprehension, and Sentence Structure, gender, and instructional method. It is one of the first studies to investigate the effectiveness of the modularized approach to teaching developmental mathematics in North Carolina. The findings of this study have implications for the effective implementation of developmental mathematics redesign at NC community colleges.

Research Methodology

The duration of the study was three semesters: Spring 2012 (Semester 1), Fall 2012 (Semester 2), and Spring 2013 (Semester 3). A total of 308 students enrolled in the courses MAT060: Essential Mathematics and MAT070: Introductory Algebra, participated in this study. The collected data consist of pre-test and post-test scores, CPT scores in Algebra, Arithmetic, Reading Comprehension, and Sentence Structure, gender information, instructional method, and semester indicators. Students' knowledge of the selected algebra concepts such as proportions, ratios, rates, and percents (MAT060 module 030); expressions, linear equations, and linear inequalities (MAT070 module 040); graphs and equations of lines (MAT070 module 050) was tested in both Phases.

In this study, linear regression models were developed to describe the relationship between the dependent variable (student improvement as measured by the difference between pre-test and post-test scores) and independent variables (CPT scores, gender, and instructional method) for all data sets. In Phase I, two linear regression models were developed to predict student academic improvement in developmental mathematics courses using the standardized CPT scores, gender, and methodological indicator as potential predictors. In Phase II, three linear models were analyzed to predict student academic performance in redesigned MML classes. Additionally, the Phase II data was used to identify a subset of students most likely to earn an unsatisfactory post-test grade.

Phase I: Spring 2012

In Phase I of this study, the effectiveness of two learning management systems Educo and MyMathLab (MML) was investigated. In Spring 2012 (Semester 1), the pre-tests and post-tests were administered to 81 students enrolled in one section of MAT070 taught using MML modules and three sections of MAT070 taught using Educo courses. Educo developmental mathematics courses aligned with the M.M. Sharma textbook series *Beginning Algebra* for MAT070 were taught in Spring 2012 by all instructors except for one who piloted MML courses. Educo MAT070 courses consisted of six chapters, online tutorials, homework assignments, and multiple choice and free response quizzes and tests. Modularized MML developmental mathematics courses aligned with the Martin-Gay textbook series *Eight Modules: Correlated with the North Carolina State Standards* were released in October 2011 and piloted in Spring 2012. MAT070 courses consisted of three modules: 010, 040, and 050 with online tutorials, individualized Study Plan, multimedia resources, built-in pre-test and post-test for each module, and interactive homework assignments.

The following research questions guided Phase I of this study:

1. Is there a statistically significant difference in students' learning outcomes between MML and Educo groups associated with the teaching method?

2. Are CPT scores and gender statistically significant predictors of student learning in developmental mathematics courses as measured by the difference between pre-test and post-test scores?

Multiple regression analyses were performed using the statistical computing software, `R`, to determine the relationships among the post-test and pre-test score, mathematics and English CPT scores, gender, and instructional method (MML or Educo). As described earlier, full models in Phase I include 6 independent variables:

$$y' = \beta_0 + \beta_1 x'_1 + \beta_2 x'_2 + \beta_3 x'_3 + \beta_4 x'_4 + \beta_5 x_5 + \beta_6 x_6 + \varepsilon$$

where

- y': Standardized mean improvement from pre-test to post-test
- x'_1: Standardized Algebra CPT score
- x'_2: Standardized Arithmetic CPT score
- x'_3: Standardized Reading CPT score
- x'_4: Standardized Sentence Structure CPT score
- x_5: Gender ($x_5 = 1$ if male; $x_5 = 0$ if female)
- x_6: Methodological indicator ($x_6 = 1$ if MML; $x_6 = 0$ if Educo)
- ε: Model error

Phase II: Fall 2012 and Spring 2013

During Phase II, all developmental mathematics students were taught using MyMathLab software and e-book. Module 030, 040, and 050 pre-tests and post-tests were administered to all MAT060 and MAT070 day classes taught at two different campuses. The classes were taught by one full-time instructor and three developmental mathematics adjunct instructors. Differences in the students' scores were investigated to determine whether CPT scores and gender were significant predictors of student learning. Note that final grades, campus and instructor indicators could be included in the Phase I and Phase II models but the goal of the study was to predict student learning using only CPT scores, gender, and methodological indicator.

The following research questions guided Phase II of this study:

1. Are CPT scores and gender significant predictors of student learning as measured by the difference between pre-test and post-test scores?

2. Which students are most likely to not earn the required grade on the post-test?

To answer question 1 of Phase II, the following comparison model was used:

$$y' = \beta_0 + \beta_1 x_1' + \beta_2 x_2' + \beta_3 x_3' + \beta_4 x_4' + \beta_5 x_5 + \beta_7 x_7 + \varepsilon$$

where y' is the standardized mean improvement calculated as the difference between pre-test and post-test score, x_1' is standardized Algebra CPT score, x_2' is standardized Arithmetic CPT score, x_3' is standardized Reading Comprehension CPT score, x_4' is standardized Sentence Structure CPT score, x_5 is gender ($x_5 = 1$ if male; $x_5 = 0$ if female), x_7 is the semester indicator ($x_7 = 1$ if Spring 2013; $x_7 = 0$ if Fall 2012), and ε is the model error. Note that Phase II models (Model 3, 4 and 5) do not include methodological indicator MML since all the students were enrolled in MML classes. Furthermore, Phase II models combine data from two consecutive semesters, thus the additional categorical variable x_7 is included in models 3, 4, and 5 to account for potential differences in the response variable between Semester 2 and 3.

To answer question 2 of Phase II, the linear regression model was used with the MML post-test as the response variable and the CPT scores, gender, and the MML pre-test as the set of possible predictors. Post-test models were based on the raw data from modules 030, 040, and 050 collected during Semester 2 and 3. The following full model was used:

$$y_2 = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \varepsilon$$

where y_2 is the raw MML post-test score (scale: 0-1 pts.), x_1 is the raw Algebra CPT score (scale: 20-120 pts.), x_2 is the raw Arithmetic CPT score (scale: 20-120 pts.), x_3 is the raw Reading CPT score (scale: 20-120 pts.), x_4 is the raw Sentence Structure CPT score (scale: 20-120 pts.), x_5 is gender ($x_5 = 1$ if male; $x_5 = 0$ if female), x_6 is the raw MML pre-test score (scale: 0-1 pts.), and ε is the model error.

Data Collection and Grading Rubric

Data were collected from developmental mathematics students enrolled in 16-week MAT060 and MAT070 classes that were offered during the day and had a minimum of ten students per class at the end of the semester. Note that subsequent data analysis does not include module 030 model for Phase I because of the circumstances beyond the control of the author of this study. A total of five data sets were used in the multiple regression analyses since student data from Semester 2 and 3 were combined in Phase II.

Phase I Tests

A pre-test and post-test were created by the researcher for each of modules 030, 040, and 050. Module 030 pre-tests were administered to MAT060 students while modules 040 and 050 pre-tests were given to MAT070 students within the first week of classes. Module

030 post-tests were administered as a part of final exam to MAT060 classes. Module 040 post-tests were administered as a part of the midterm exam to MAT070 classes. Module 050 post-test were administered as a part of the final exam to all MAT070 classes. Pre-tests and post-tests for modules 030 and 040 consisted of five questions each, while tests for module 050 consisted of four questions per test. Each test was worth a total of 15 points. Students were required to show all steps, write explanations in complete sentences, label graph axes, and perform calculations. The following 3-point grading rubric adapted from H. Dogan (2001) was used to score students' responses on pre-test and post-test questions:

- 3 points. Complete response to all aspects of the problem which indicates complete mathematical understanding of the problem's concept. Answers are written in complete sentences. Everything is labeled and explained.

- 2.5 points. Complete response to all aspects of the problem but includes minor computational errors or missing explanations.

- 2 points. Incomplete response but student demonstrates understanding of the main idea of the problem. Student shows some deficiencies in understanding aspects or steps of the problem. Incomplete reasoning.

- 1.5 points. Incomplete response. Student shows partial understanding of the problem and/or steps required to solve the problem. Incomplete or incorrectly labeled and/or explained.

- 1 point. Poor attempt. Student fails to answer or complete problem. Very limited or no understanding of problem. Contains words, examples, or diagrams that do not reflect the problem.

- 0 points. No answer. Student made no attempt to respond to the problem or written information was insufficient to allow judgment [Dogan, 2001].

The test questions were chosen by the author of this study based on Developmental Mathematics Modular Curriculum BETA Test Version materials created by the Math Redesign Task Force comprising of 18 math faculty members from participating colleges [NCMATYC, 2012]. The tests were approved as a part of the departmental diagnostic test by SCC's Arts and Sciences Department.

Phase II Tests

In Phase II, computerized MML pre-test and post-test scores were collected from students enrolled in redesigned MML classes during Semesters 2 and 3 (Fall 2012 and Spring 2013, respectively). Each computerized MML test consisted of approximately 20 to 30 free response questions where students were required to enter numerical values for partial credit. The students were allowed multiple attempts on MML pre-tests and post-tests, however, only the scores from the first attempt were recorded and analyzed.

Permissions from Western Carolina University (WCU) and SCC Institutional Review Boards (IRB) were obtained in Fall 2011. Consent forms were collected from all of the participating students.

2 REVIEW OF THE LITERATURE

Within the last ten years, several reports summarizing the results of educational studies were prepared for the U.S. Department of Education. Many of these studies employed multiple regression models to evaluate the impact of various teaching delivery modes (traditional lecture, computer-assisted instruction (CAI), online courses, and accelerated programs) on student success [Department of Education, 2005, Condelli, 2006]. There have been a limited number of rigorous studies conducted in the 21st century that investigated the effectiveness of instructional approaches and determined predictors of students' success in developmental mathematics [Hendricks, 2012, Hodara, 2011, Jaggars, 2011]. The review of research in developmental mathematics, CAI, and applications of multiple regression models is presented in this chapter.

In the report prepared for the Division of Adult Education and Literacy, the U.S. Department of Education (2005) described the results of 15 developmental mathematics studies where "technology-based or technology-enhanced instruction" was compared with traditional lectures [Department of Education, 2005]. Similarly, the American Institutes for Research report included detailed comparisons of 24 studies in Adult Basic Education (ABE) and developmental mathematics. Both reports concluded that there was no consensus on the effectiveness of CAI in comparison to traditional lecture [Condelli, 2006].

According to M. Hodara, Senior Research Assistant at the Community College Research Center (CCRC), redesigned programs "may have the potential to improve the outcomes of developmental math students" [Hodara, 2011, p.2]. She concluded that it is difficult to determine the effectiveness of new pedagogical practices because of the poor internal validity of many studies. The results of these studies were inconclusive regarding the determinants of students' success. However, there was evidence suggesting that students with learning disabilities and lower-achieving students might benefit more than average students from redesigned mathematics programs [Hodara, 2011].

Multiple Regression Models

In terms of methodology, a majority of recent studies used the multiple regression approach to model the relationship between students' test results and independent variables [Hodara, 2011, Jaggars, 2011, Department of Education, 2005, Condelli, 2006]. According to S. Huang, there are many positive effects of predictive modeling using linear regression techniques [Huang, 2011]. Statistical models can be used to help identify academically "at-risk" students and aid the instructors in determining the best course of action to prevent these students from withdrawing or failing a class [Lowis and Castley, 2008, Veenstra et al., 2009, Ware and Galassi, 2006].

The goal of the study conducted by Marsh et al. was to predict student academic improvement (measured by GPA) using age, gender, ACT, SAT, and general psychology exam scores collected from 257 students in an introductory psychology course. The study found that the most effective predictors were the scores from other required psychology courses [Marsh et al., 2008].

Huang conducted a study at the University of Utah using various statistical techniques such as multiple linear regression and neural networks. A total of 24 predictive models were developed using students cumulative GPA, grades in four prerequisite courses, and scores in three dynamics mid-term exams as predictors. Dr. Huang found there were significant differences in internal and external accuracy of different models depending on which modeling technique was used and which subset of predictor variables were included in the model. The research findings from this study implied that neural network models had better prediction accuracy than multiple regression models when using the same predictor variables. However, in comparison to other modeling techniques used to construct predictive models, linear regression is easier to interpret and provides an explicit mathematical equation [Huang, 2011].

Accuracy of Placement Tests

Standardized tests such as the Scholastic Aptitude Test (SAT) and the American College Test (ACT) have been considered potential predictors of student learning, but the results of studies were inconclusive [Harachiewicz et al., 2002, Emerson and Taylor, 2004]. Approximately 92% of community colleges uses ACCUPLACER or COMPASS placement tests to place students into developmental education classes. According to CCRC, 27-33% of community college students are misplaced into developmental education classes based on their CPT scores [Belfield and Crosta, 2012, Hodara et al., 2012, Jaggars et al., 2013, Scott-Clayton, 2012]. Based on most recent research by the CCRC, combining the high school GPA scores with the CPT scores might significantly decrease the placement error and improve student completion rates [Jaggars et al., 2013].

Research on Developmental Mathematics: An Overview

The research on the effectiveness of CAI in developmental mathematics is limited to isolated studies and the results are inconclusive [Hendricks, 2012, Hodara, 2011, Jaggars, 2011, Spradlin, 2009]. A majority of experimental studies found no statistically significant difference in final test scores between the CAI and traditional groups.

A study conducted by Waycaster at five community colleges in Virginia found no statistically significant differences between the pass rates of students taking classes in the lecture format, individualized instruction with tutoring, and CAI format. He analyzed several factors affecting the success of community college developmental students. Waycaster found students' and instructors' gender to be correlated and statistically significant in evaluating the effectiveness of a teaching method. In Waycaster's study, female students taught by a female instructor tended to be more active in class which might have affected their test improvement and the study's results. Therefore, it might be advisable to ensure that both male and female instructors teach both experimental and control groups in educational studies [Waycaster, 2001].

Similarly, Kinney and Robertson found no statistically significant difference in final examination scores between the traditional (lecture) classes and CAI classes. The study was conducted with Elementary Algebra and Intermediate Algebra students at the University of Minnesota. In the CAI class, students worked at their own pace with software providing presentation of the material. The study design allowed all students in both traditional and CAI classes to have access to software, but only CAI students were required to use the learning system in class [Kinney and Robertson, 2003].

Villarreal's study, conducted at a community college in Texas, investigated differences in pass rates of students enrolled in Introductory Algebra and Intermediate Algebra classes. The study compared two modes of delivering instruction: self-paced computer lab and hybrid class. The control group worked in an unstructured, open computer lab with tutors helping the students. The redesigned (structured) hybrid course consisted of three hours of lecture and three hours of required computer lab per week. Students in the redesigned hybrid class had higher scores than students in open computer lab. The results were promising because the pass rate increased by 12% within two years, but these improvements might have been due to other factors not related to the delivery method [Villarreal, 2003].

A study conducted by Teal at a suburban community college in the Middle Atlantic region showed no statistically significant difference between final exam scores in the experimental and control groups. A total of 152 developmental algebra students were enrolled in either traditional lecture classes or CAI classes. The data on students' academic improvement were collected from post-tests given after six weeks and final exams given at the end of the semester. Three instructors each taught one CAI class, where students used Educo software, and one traditional class. Students in CAI classes worked on computer assignments in class with the instructor providing mini-lectures. The methodology and results of Teal's study were compared with other studies described in U.S. Department of Education and CCRC reports (2009) [Hodara, 2011].

A study conducted by Spradlin at a large, private, eastern university provided similar results on the effectiveness of CAI in comparison to traditional lecture. A total of 99 students enrolled in four sections of a developmental Intermediate Algebra course were taught by two full-time instructors. In this study, computer learning system CengageNOW was used to supplement traditional classroom instruction where computer-based learning occurred outside of the classroom. No significant difference in students' scores was found that could be attributed to the use of the computer learning system. Furthermore, Spradlin investigated gender as a potential predictor of success in developmental mathematics courses. There was a significant difference in post-test scores between females and males in both CAI and traditional groups. In Spradlin's study, females outperformed males in both groups; however, the author noted that the gender differences in test scores might be influenced by the fact that the instructor and majority of the students in the classes were females. According to Spradlin, the effectiveness of CAI might be influenced by the human factor, namely, educator's training and students' motivation to incorporate technology effectively. Therefore, improvements in students' academic improvement rely both on the human factor and the quality of the software [Spradlin, 2009, Mejri, 2011]. Furthermore, student learning outcomes might be affected by how well the computer learning system is integrated into the curriculum. Also, Spradlin noted that the results of educational studies cannot be easily compared and their results generalized because these studies differ significantly by the sample size, variables, type of computer learning system and its implementation into curriculum, and students' characteristics.

Developmental Mathematics Redesign: History and Research

The history of modern redesign efforts in developmental mathematics started in 1999 with *The Program in Course Redesign* coordinated by the Center for Academic Transformations directed by Dr. Carol A. Twigg. The program was created to support colleges and universities in their efforts to redesign instruction in order to improve students' learning and reduce the costs [Twigg, 2005].

According to Epper and Baker, the key elements of *The Program in Course Redesign* were "whole course redesign (rather than by section), active learning, computer-based learning resources, mastery learning, on-demand help, and alternative staffing (replacing expensive faculty labor with inexpensive labor or technology where appropriate)" [Epper and Baker, 2009, p.5]. In *An Overview of Current and Emerging Practices*, Epper and Baker concluded that successful redesigned programs were complex models with many factors such as administrative, instructional, and student support strategies affecting the results. Therefore, it might be difficult to evaluate the effectiveness of the modularized approach because of the complexity of the relationship between students' improvement, pedagogical approach, and quality of the software.

Within the last few years, developmental mathematics received increased attention and resources through national initiatives such as the Bill and Melinda Gates Foundation, the Lumina Foundation for Education's Achieving the Dream project, and the Jobs for the Future project [Epper and Baker, 2009, NCCCS, 2011]. Several reports were prepared for the U.S. Department of Education to investigate the effectiveness of redesigned programs and new initiatives geared toward increasing the retention and completion rates of developmental students at community colleges and universities [Hodara, 2011, Jaggars, 2011, Department of Education, 2005, Condelli, 2006].

According to an MDRC report, course redesign produced significant improvements in students' test scores, completion rates, and attitudes toward mathematics. The report described the goals, methodology, and findings of three redesigned pilot programs implemented with the help of the Achieving the Dream: Community Colleges Count initiative at two Virginia and one North Carolina community colleges. At one Virginia community college, a redesigned *Fast Track Math* course had 60% course pass rate in comparison to 27% pass rate in the control group. However, there were major limitations of the evaluation design that raised questions about internal and external validity of the study's results. The studies did not include comparisons of students' characteristics and motivation levels in the treatment and control groups. Furthermore, there was no measure of statistical significance of the results [Zachary, 2008].

Carol Twigg, president of the National Center for Academic Transformation (NCAT), reported dramatic improvements in students' completion rates and reductions in delivery costs for the math emporium model. Twigg's report included results of case studies from 37 colleges and universities that implemented a self-paced, modularized, technology-based approach called *The Emporium Model*. Implementation of redesigned computer-assisted courses improved completion rates by 51% on average and reduced instruction costs by 30%. According to Twigg, the key elements of success of redesigned courses are "interactive computer software, personalized on-demand assistance, and mandatory student participation" [Twigg, 2011, p.26].

However, in the CCRC report, Hodara questioned the internal validity of the case study results reported by Twigg. The effectiveness and quality of the redesigned models depend on many factors, therefore, "outcomes may be due to any number of changes in how course content is delivered, when students can access course content, and the pedagogy utilized in each model" [Hodara, 2011, p.24]. Furthermore, the quality of the software used at various colleges and universities varies greatly which makes it difficult to evaluate the effectiveness of *The Emporium Model*.

According to the MDRC report, redesigned models such as "self-paced, or modularized, courses, which break apart semester-long developmental education classes into smaller, competency-based units" have shown promising results [Zachary, 2008, p.2]. MDRC researchers described redesigned programs as one of the most effective practices in developmental education. It was implied that educational reforms might need to include drastic changes in curriculum, pedagogy, and the use of technology. The authors of the report emphasized that additional rigorous studies should be conducted to evaluate the effectiveness of pedagogical practices, including redesigned programs supported by the Achieving the Dream initiative [Zachary, 2008].

Developmental Mathematics Redesign in North Carolina

North Carolina is one of six states participating in the Developmental Education Initiative (DEI) as part of the Achieving the Dream project. NCCCS focused on redesigning developmental mathematics first because the largest number of developmental students are placed into developmental mathematics courses, and it "represents the greatest stumbling block to students success" [NCCCS, 2011, p.1]. The redesign reform in the state of North Carolina officially started in 2009 when NCCCS president Scott Ralls established the Developmental Education Initiative State Policy Team. In October 2010, the North Carolina EDI State Policy Team approved design principles to initiate the changes in developmental math curriculum across the state. The emphasis has been placed on conceptual and contextual delivery of material, real-life applications, and the use of technology. The goal of redesign has been to allow students "to complete their required developmental mathematics requirements at a pace that is appropriate to their needs and knowledge" [NCCCS, 2011, p.1]

In January 2011, the new Math Task Force, consisting of 18 developmental and curriculum math faculty, started its work to create developmental math modules. In August 2011, the Module Outlines and Notes Beta Test Version was created that included guidelines, sample conceptual test questions and introductory applications.

Research on Computer-Assisted Instruction: A Brief Overview

The theoretical bases of computer-assisted instruction are rooted in behaviorism, constructivism, and online learning theories, and they hold a potential to enhance learning [Moosavi, 2009, Hodara, 2011, Mejri, 2011]. Studies that included different age groups and a broad variety of disciplines provide evidence suggesting that computer-assisted instruction has a positive effect on students' learning outcomes. However, the internal validity of many studies is questionable. There are significant differences in terms of time spent, curriculum, and pedagogy between the experimental and treatment groups that might affect the results. The non-equivalence of instruction and curriculum could account for differences in students' test scores between the CAI and traditional lecture groups [Department of Education, 2005, Hodara, 2011]. Additionally, withdrawal rates might result in misleading comparisons between CAI and traditional instruction groups, and affect the internal validity of many educational studies [Jaggars, 2011, Moosavi, 2009]. Furthermore, the results of studies on the effectiveness of CAI cannot be easily compared because of differences in methodology and quality of learning management systems [Mejri, 2011, Moosavi, 2009, Spradlin, 2009].

Some studies found gender to be a significant predictor of success in developmental classes [Hodara, 2011, Spradlin, 2009]. A few studies investigated students' characteristics (age, race/ethnicity, marital status) and placement test scores as potential predictors of student success as measured by final course grades or post-test scores [Hannafin and Foshay, 2008, Reagan, 2004, Spradlin, 2009, Waycaster, 2001]. The study conducted by Reagan at a rural community college in Texas found significant positive correlation between mathematics and reading scores measured using a computerized placement test. However, no statistically significant difference was found in post-test scores between the traditional and CAI groups. Reagan concluded that more research might be needed to investigate further the correlation between mathematics and reading test scores [Reagan, 2004].

Computer-assisted instruction had a positive effect on low-achieving high school students in a study conducted by Hannafin and Foshay where the Plato Learning System was used. A total of 87 “at-risk” students who scored low on the statewide standardized test in the 8th grade were placed into a remedial 10th grade CAI course to investigate improvements in test scores. Statistically significant improvements in students’ standardized test scores were found. However, the authors noted that the positive results of the study might have been affected by other curriculum changes and professional development efforts geared at improving students’ standardized tests scores. There was no control group, which raised the question of the internal validity of the study’s results [Hannafin and Foshay, 2008].

The difference in the quality of the learning management systems might be a significant factor when evaluating the effectiveness of the redesigned programs in mathematics. No meta-analysis reports were found that compare multiple software packages used in developmental mathematics classes. A few studies compare the effectiveness of two learning management systems with traditional instruction, but the results were inconclusive [Moosavi, 2009]. However, the quality of various learning systems used in statistics is compared in a report by Yung-chen Hsu. Meta-analysis of 25 rigorous studies described by Hsu finds that “different modes of CAI programs produced significantly different effects on students’ achievement in learning statistics” [Hsu, 2003, Abstract]. More research is required to determine whether the differences in the quality of learning management systems used in developmental mathematics produce similar results.

Research on the Effectiveness of MyMathLab

Studies on the effectiveness of MyMathLab (MML) and other learning management systems used in developmental mathematics classes have been compared. This comparison indicates that the effectiveness of MML is limited to a few isolated rigorous studies and Pearson’s annual reports.

Four well-documented studies were conducted recently to investigate differences in students' academic improvement in mathematics that could be attributable to the use of MyMathLab system. These studies employ an experimental design, where instruction is delivered using lecture (control group) and MML computer learning system (treatment group). In terms of methodology, multiple regression is used to analyze the data. These studies address threats to the validity resulting from attrition issues, and students' characteristics are compared and analyzed.

Loving conducted a study at the University of Southern Mississippi and found the difference in students' learning outcomes between CAI and traditional groups to be significant. Loving included analysis of African-American students' academic improvement based on age, gender, and technological proficiency in traditional and MML-assisted courses that lasted from 6 to 12 weeks. Loving's study cannot be easily compared with other studies on the effectiveness of MML-assisted instruction because of the differences in the design, length of treatment, and students' characteristics [Loving, 2007].

Moosavi compared the effectiveness of two computer learning systems: MyMathLab and Thinkwell with traditional instruction in a precalculus class at the University of Alabama. This study found that students performed significantly better in the traditional group than in both CAI groups. Furthermore, Moosavi noted that students using Thinkwell performed better than students in MyMathLab group which implied that one of the learning management systems is better than the other. According to Moosavi, "CAI may be best used to supplement traditional instruction" because of the importance of student-teacher interaction in motivating students and helping them overcome their struggles [Moosavi, 2009, p.56]. Additionally, Moosavi found significant differences in drop-out rates between traditional and CAI groups. Overall, the traditional group performed better on the final tests and had better completion rates in comparison to both Thinkwell and MML-assisted classes [Moosavi, 2009].

Mejri conducted a study on the effectiveness of MyMathLab in comparison to traditional lecture. A total of 100 community college students enrolled in a six-week Basic

Mathematics class received a three-day training course on how to use the software at the beginning of the treatment. The study found that students in MyMathLab-assisted groups performed significantly better than their peers in traditional groups. Students' attitudes towards mathematics and completion rates were analyzed and compared, but no statistically significant differences were found. The researcher was also the instructor in the MML group which raised a question of potential bias resulting from Mejri's dual role. Furthermore, in this study, the passing score was 60% in comparison to redesign guidelines of 80% course passing score and 85% minimum score on module post-test. These differences make it difficult to compare the findings of this study with other studies on the effectiveness of computer-assisted instruction in developmental mathematics [Mejri, 2011].

Hendricks conducted a study at one of the NC community colleges to investigate the predictors of success for developmental mathematics students enrolled in traditional, hybrid, and online courses. Students in hybrid and online courses were using MyMathLab learning system. A total of 130 students enrolled in developmental MAT070: Introductory Algebra and MAT080: Intermediate Algebra classes completed online surveys at the beginning of the semester and final exams at the end of the semester. Students' gender, age, race/ethnicity, marital status, mathematics and technological self-efficacy, and several other characteristics were analyzed as potential predictors of students' success in mathematics classes. Hendricks found that students' mathematics self-efficacy was a significant predictor of success, but technological self-efficacy was insignificant. Mean final exam scores of hybrid and online students were higher than in the traditional group but no statistical analysis on the significance of these results was performed. According to Jaggars, there is evidence that students taking online classes tend to have higher placement scores, which might explain higher mean final exam scores [Jaggars, 2011]. Internal validity of the instructor-created departmental final exam and online survey was analyzed using Cronbach's alpha by comparing the results between students participating in the study and their peers who did not complete the voluntary online survey [Hendricks, 2012].

Pearson, an educational services provider, published annual reports on the effectiveness of its software, including MyMathLab developmental mathematics courses. Pearson's publication *Making the Grade V.5* included 77 case studies of colleges and universities that use MML software as a supplement (CAI classes) or main instruction delivery method (online classes) [Speckler, 2012]. Community colleges and universities provided data showing dramatic improvements in students' scores, completion rates, and retention, but no information on statistical significance, study design, and control over students' characteristics was provided. Therefore, this report might be considered anecdotal evidence of the effectiveness of MML developmental mathematics courses.

3 DATA ANALYSIS

Presented in this chapter are the following steps of the multiple regression analysis: initial diagnostics of the raw and standardized data from Phase I and Phase II of the study, backward and forward selection of predictor variables, identification and removal of influential observations, refitting of the models, and diagnostic tests of reduced models. Additionally, an out-of-sample prediction method is used to evaluate the misclassification rate in identifying “at-risk” students.

Five linear regression models were analyzed based on 308 observations collected during the duration of the study. Each model consisted of six predictors. In order to make effective comparisons, all quantitative variables were standardized to convert them to a common scale.

Building the Regression Models

Verification of Linear Regression Assumptions

The normality of residual distributions were investigated using residual histograms and normality plots for each of the five full and reduced models. The assumption of equal variance of residuals for each model was graphically verified using the plot of residuals versus the fitted values as shown in Appendices 1 and 2. Each of the quantitative predictor variables had a linear relationship with the response variable.

Backward and Forward Selection of Predictor Variables

For each of the five models, backward and forward selection procedures were employed to identify the best subset of independent variables using the adjusted-R^2 value as the selection criterion. At each step, the adjusted-R^2 value, model coefficients, and corresponding p-values were calculated using the statistical computing software, `R`. Then the independent variable with the largest p-value was eliminated. After performing both the forward and backward selection procedure, the model with the highest adjusted-R^2 value was determined. Next, an interaction term was added to the reduced model. In all cases, however, adding the interaction term did not improve model's adjusted-R^2 value.

Identification of Influential Observations

The measure of influence DFBETAS, DFFITS, covariance ratio, Cook's distance, and hat matrix values were used to identify influential observations in the collected data for each of the regression models. The DFFITS measure assesses the influence that case i has on the fitted value when all n cases are used in fitting the regression function. The DFBETAS measure assesses the influence of case i on each regression coefficient. The absolute magnitude of DFBETAS value quantifies the influence of the i^{th} case on the k^{th} regression coefficient relative to the estimated standard deviation of the coefficient. The diagonal elements h_{ii} of the hat matrix were used to directly identify outlying predictor observations. The leverage h_{ii} value is a measure of the distance between predictor values for the i^{th} case and the means of the predictor values for all n cases. Cook's distance is an aggregate measure that assesses the influence of the i^{th} case on all n fitted values. Based on Cook's distance, the i^{th} case can be influential because of the large value of the residual e_i or leverage h_{ii} or large value of both [Neter et al., 1996].

Each of the five reduced models was tested to identify influential observations using `R` and refitted to determine the best subset of predictor variables. The following reduced models had the highest adjusted-R^2 values:

Model 1 $\hat{y}' = -.16 - .49x_1' - .27x_2' + .30x_5$

Model 2 $\hat{y}' = -.16 - .23x_1' + .23x_3' + .27x_4' + .35x_6$

Model 3 $\hat{y}' = .08 - .26x_1' - .34x_2' - .21x_5$

Model 4 $\hat{y}' = -.03 - .29x_1' - .21x_2' + .20x_3' - .27x_4' - .35x_5 + .31x_7$

Model 5 $\hat{y}' = .32 - .43x_1' - .21x_2' - .16x_4' - .59x_7$

where

- $\hat{y}'$: Estimated standardized mean improvement from pre-test to post-test
- x_1': Standardized Algebra CPT score
- x_2': Standardized Arithmetic CPT score
- x_3': Standardized Reading CPT score
- x_4': Standardized Sentence Structure CPT score
- x_5: Gender ($x_5 = 1$ if male; $x_5 = 0$ if female)
- x_6: Methodological indicator ($x_6 = 1$ if MML; $x_6 = 0$ if Educo)
- x_7: Semester indicator ($x_7 = 1$ if Spring 2013; $x_7 = 0$ if Fall 2012)

Diagnostic Tests of Reduced Models

All five reduced models were tested to check whether the assumptions of the normality of residual distribution and equal variance of residuals were satisfied. Repeating the procedures used in the initial diagnostics on full models, the normality of residual distribution was investigated using residual histograms and normality plots. Similarly, the assumption of equal variance of residuals for each model was graphically verified using the plot of residuals versus the fitted values as shown in Appendix 2.

Linear Models

Model 1

The following final reduced model is based on module 040 Phase I data:

$\hat{y}' = -.16 - (.49)\text{zAlgebra} - (.27)\text{zArithmetic} + (.30)\text{Gender}.$

It was developed to compare student improvement between Educo and MML courses. The reduced multiple R^2 value for this model is .3925, therefore, approximately 39% of variability in differences between students' pre-test and post-test scores can be explained by the linear model. The following summary table was generated using R after the refitting of the final model:

	Estimate	Std. Error	t value	Pr(> |t|)	
(Intercept)	-0.1563	.12312	-1.269	.2085	
zAlgebra	-0.4927	0.09109	-5.409	7.87e-07	***
zArithmetic	-0.2729	0.10078	-2.708	0.00845	**
Gender	0.3013	0.17994	1.674	0.09844	.

Figure 1: Model 1 Summary Table

As shown in Figure 1, the reduced Model 1 with the highest adjusted-R^2 value of 0.3672 includes the CPT standardized scores in Algebra (zAlgebra) and Arithmetic (zArithmetic), and gender indicator (Gender). Based on Model 1, CPT scores in Algebra and Arithmetic are both significant at the $\alpha = 0.01$ level. Methodological indicator MML was not included in the reduced Model 1 which suggests that student improvement in module 040 Educo and MML classes in Fall 2012 was not associated with the teaching method. Assuming all the other independent variables are held constant, one standard deviation increase in Algebra CPT score is associated with a reduced increase in student learning. It can be interpreted that larger Algebra CPT scores are associated with reduced values of student improvement. Similarly, one standard deviation increase in Arithmetic CPT score is associated with a 0.27 standard deviation decrease in student improvement. Hence, it can be interpreted that the higher the Arithmetic CPT score, the less improvement in student learning. Based on Model 1, gender's coefficient value of 0.3 implies that male students completing module 040 in Spring 2012 at SCC showed slightly greater improvement than female students.

Model 2

Model 2 is based on module 050 Phase I data. The following comparison model was determined to have the highest adjusted-R^2 value:

$$\hat{y}' = -.16 - (.23)\text{zAlgebra} + (.23)\text{zReading} + (.27)\text{zSentence} + (.34)\text{MML}.$$

As shown in Figure 2, the reduced Model 2 includes standardized CPT scores in Algebra (zAlgebra), Reading Comprehension (zReading), Sentence Structure (zSentence), and methodological indicator (MML). The following Model 2 summary table was generated using `R`:

	Estimate	Std. Error	t value	Pr(> |t|)	
(Intercept)	-0.1603	.1357	-1.181	.2412	
zAlgebra	-0.2298	0.1010	-2.277	.0257	*
zReading	0.2329	0.1192	1.954	0.0545	.
zSentence	0.2689	0.1191	-2.258	0.0269	*
MML	0.3488	0.2005	1.740	0.0861	.

Figure 2: Model 2 Summary Table

The interpretation of the standardized Algebra coefficient is similar for both Phase I models (Model 1 and 2), namely, the larger the Algebra CPT score, the smaller the student academic improvement. Assuming all of the other independent variables are held constant, one standard deviation increase in Reading CPT score is associated with a 0.23 standard deviation increase in student improvement. Similarly, one standard deviation increase in Sentence Structure CPT score is associated with a 0.27 standard deviation increase in student improvement. The methodological indicator (MML) is significant only at the $\alpha = 0.10$ level which can be interpreted that students enrolled in module 050 MyMathLab classes in Spring 2012 showed an improvement of 0.35 standard deviations over Educo students.

However, the multiple R^2 value of 0.2732 implies that only approximately 27% of the variability in differences between students' pre-test and post-test scores can be explained by the linear model. After adjusting for the number of variables, the adjusted-R^2 of reduced Model 2 was 0.234. Thus, it seems reasonable to conclude that there are other potential predictors of student improvement such as previous mathematical knowledge, demographic status, or class size that are not included in Model 2. The analysis of other potential predictors was not the focus of this study since the goal was to model the relationship between student improvement and CPT scores, gender, and instructional method.

Model 3

Model 3 is based on the combined module 030 data from Semesters 2 and 3. It was developed to predict student improvement in the redesigned MML courses. This model has an adjusted-R^2 value of 0.2574.

$$\hat{y}' = .08 - (.26)\text{zAlgebra} - (.34)\text{zArithmetic} - (.21)\text{Gender}$$

The following summary table was generated using `R` after the refitting of the final model:

	Estimate	Std. Error	t value	Pr(> |t|)	
(Intercept)	0.0796	.09384	-1.848	.39819	
zAlgebra	-0.2626	0.08737	-3.006	.003227	**
zArithmetic	-0.3423	0.08582	-3.989	0.000115	***
Gender	-0.2096	0.15755	-1.330	0.186023	

Figure 3: Model 3 Summary Table

As shown in Figure 3, models 1 and 3 include the same subset of predictor variables, namely standardized Algebra and Arithmetic CPT scores, and gender. The Algebra and Arithmetic coefficients for models 1 and 3 have the same order of magnitude and same sign; thus, their interpretations are similar. Note that the smaller the Algebra or Arithmetic CPT score, the greater the student improvement. In Model 3, gender is not a particularly significant predictor. Gender's coefficient implies that in modularized MML classes female students' improvement is slightly greater than male students'.

Model 4

Model 4 is based on module 040 Phase II data and has an adjusted-R^2 value of 0.2946. The following prediction model was determined to have the highest adjusted-R^2 value:

$\hat{y}' = -.03 - (.29)\text{zAlgebra} - (.21)\text{zArithmetic} + (.20)\text{zReading} - (.27)\text{zSentence} - (.35)\text{Gender} + (.31)\text{Semester}.$

As shown in Figure 4, all CPT scores with the exception of Reading CPT score are significant at the $\alpha = 0.10$ level in addition to the categorical variables, gender and semester indicator.

	Estimate	Std. Error	t value	Pr(> |t|)	
(Intercept)	-0.0303	.14765	-0.205	.83811	
zAlgebra	-0.2905	0.09526	-3.050	.00301	**
zArithmetic	-0.2149	0.10502	-2.046	0.04363	*
zReading	0.1959	0.11879	1.649	0.10259	
zSentence	-0.2708	0.11616	-2.331	0.02197	*
Gender	-0.3510	0.18292	-1.919	0.05817	.
Semester	-0.3107	0.18045	1.722	0.08857	.

Figure 4: Model 4 Summary Table

The positive coefficient of semester indicator implies that students enrolled in Spring 2013 improved more on the computerized MML tests than the Fall 2012 students. The negative coefficients for Algebra and Arithmetic CPT scores can be interpreted that larger Algebra and Arithmetic CPT scores are associated with diminished student improvement in MML classes. Furthermore, the negative gender coefficient implies that female students performed slightly better than male students at the $\alpha = 0.10$ significance level.

Model 5

Model 5 was developed to predict student improvement based on combined module 050 data from Semesters 2 and 3. It had the highest adjusted-R^2 value among the five models.

$$\hat{y}' = .32 - (.43)\text{zAlgebra} - (.21)\text{zArithmetic} - (.16)\text{zSentence} - (.59)\text{Semester}.$$

The following summary table was generated using R after the refitting of the final model:

	Estimate	Std. Error	t value	Pr(> |t|)	
(Intercept)	0.3232	.12015	2.690	.008509	**
zAlgebra	-0.4317	0.08721	-4.950	3.39e-06	***
zArithmetic	-0.2098	0.09032	-2.322	0.022441	*
zSentence	-0.1641	0.08170	-2.008	0.047622	*
Semester	-0.5881	0.15800	-3.722	0.000342	***

Figure 5: Model 5 Summary Table

The negative coefficient of semester indicator implies that students enrolled in Spring 2013 improved less on the computerized MML tests than Fall 2012 students. Based on the multiple R^2 value of 0.4583, approximately 46% of the variability in differences between students' pre-test and post-test scores can be explained by the linear model. After adjusting for the number of variables in Model 5, the adjusted-R^2 was found to be 0.4345.

Preliminary Identification of "At-Risk" Students

One of the goals of this study was to identify "at-risk" students, namely students who are most likely not to complete each of the 030, 040, or 050 modules. Early identification of "at-risk" students might allow instructors to take proactive measures and develop strategies to help students pass the developmental mathematics courses. The following full regression model was used in identifying "at-risk" students:

$$y_2 = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \varepsilon$$

where y_2 is the raw post-test score (scale: 0-1 pts.), x_1 is the raw Algebra CPT score (scale: 20-120 pts.), x_2 is the raw Arithmetic CPT score (scale: 20-120 pts.), x_3 is the raw Reading CPT score (scale: 20-120 pts.), x_4 is the raw Sentence Structure CPT score (scale: 20-120 pts.), x_5 is gender ($x_5 = 1$ if male; $x_5 = 0$ if female), x_6 is pre-test score (scale: 0-1 pts.), and ε is model error.

These full models were developed using the raw data collected from students completing modules 030, 040, and 050 during Semesters 2 and 3. For each data set, all of the observations were used to perform a variable selection procedure to identify the predictors significant at the $\alpha = 0.05$. Based on backward and forward selection procedures, the pre-test score was a significant predictor in each reduced post-test model. As shown below, in addition to pre-test score (x_6), Algebra (x_1), Arithmetic (x_2), and Sentence Structure CPT (x_4) scores were significant predictors. The following reduced models were used to identify "at-risk" students based on module 030, 040, and 050 Phase II data, respectively:

$$y_2 = \alpha_0 + \alpha_6 x_6 + \varepsilon$$

$$y_2 = \alpha_0 + \alpha_2 x_2 + \alpha_6 x_6 + \varepsilon$$

$$y_2 = \alpha_0 + \alpha_1 x_1 + \alpha_4 x_4 + \alpha_6 x_6 + \varepsilon$$

The post-test score and selected predictors were used to build classifiers and evaluate the accuracy of these classifiers. An out-of-sample predictions was used to predict post-test scores for each observation as described below.

1. Remove observation i from the data set.

2. Find the linear regression function that models post-test as a function of selected predictors using remaining $n-1$ observations.

3. Use linear regression function to predict post-test score for observation i (omitted observation).

Next, the actual and predicted post-test scores were used to divide all observations into four classes and the graphs of predicted post-test scores versus actual post-test scores were analyzed. Note that the predictive accuracy of the model was limited in that there were several students who had predicted post-test scores greater than 0.90, yet who still did not pass the class. Nevertheless, the predicted post-test score was considered a predictor of the actual post-test score. At SCC passing the class was defined as earning at least 0.85 on the actual computerized MML post-test, regardless of the other grades in the course. Thus, in the process of identifying "at-risk" students, the "true pass" was defined as an actual post-test score of at least 0.85.

As shown in Figure 6, a naive rule for a predicted pass was considered using 0.85 as a naive split point to divide the observations into classes. An observation was considered a predicted pass in cases where the predicted post-test score was at least 0.85. From these classifications, the following table and scatterplot (Figures 6 and 7) were obtained. From Figure 6, it can be seen that the actual pass rate for those with predicted post-test score less than .85 was 46%, while the actual pass rate for those with predicted post-test score at least 0.85 was 81%, which yielded a difference in pass rate of 35%. Additionally, 38/129 = 29% of the observations were misclassified, as 23 students were predicted to fail, but actually passed, and 15 were predicted to pass, but actually failed.

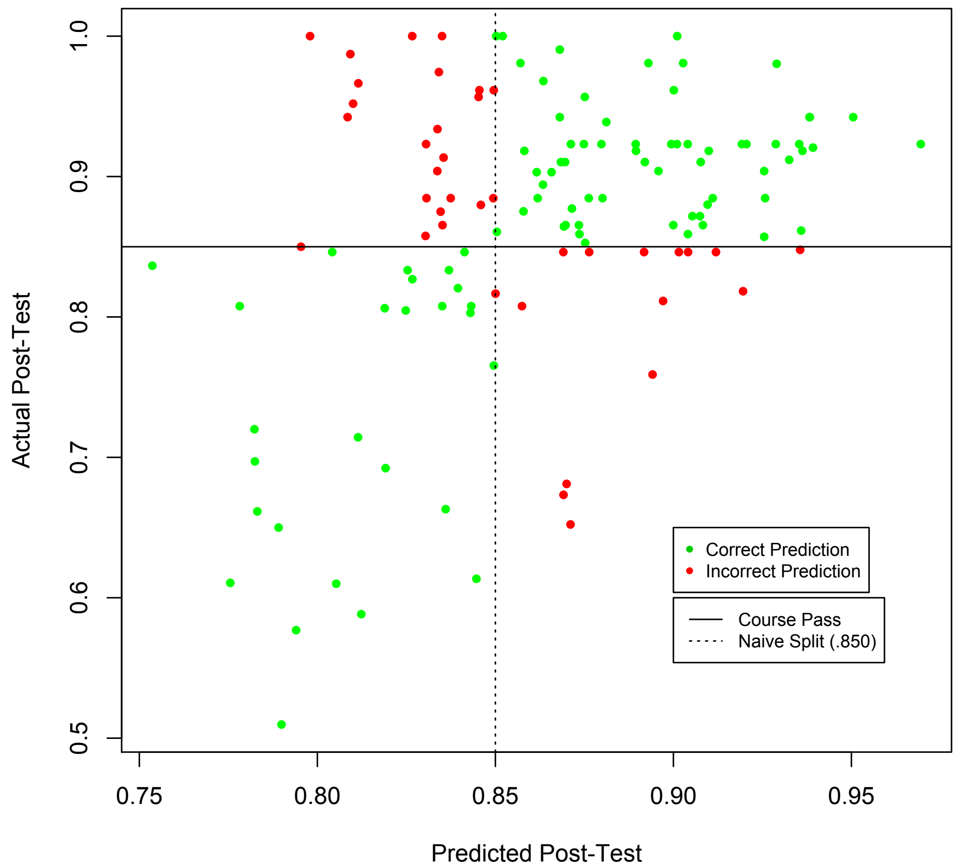


Figure 6: Prediction Summary Scatterplot with Naive Split Point of 0.85

Figure 7 provides information for calculating misclassification rate in identifying "at-risk" students using a naive split point of 0.85 based on Module 030 Phase II data.

		Predicted Passing	
		FALSE	TRUE
Actual Passing	FALSE	27	15
	TRUE	23	64

Figure 7: Prediction Summary Table Based on Naive Split Point of 0.85

After completing these steps, the optimal split point of 0.827 was calculated.This split point minimized the misclassification rate. The predictions were summarized in the Figures 8 and 9.

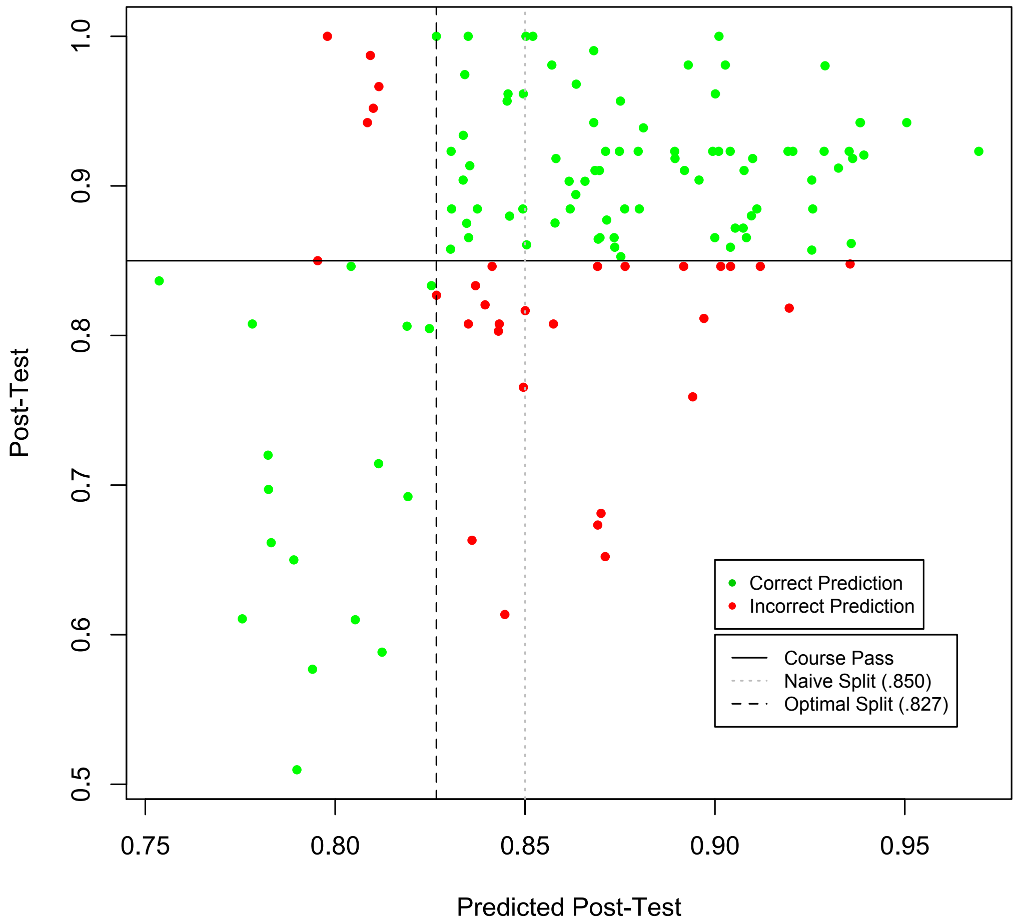


Figure 8: Prediction Summary Scatterplot with Optimal Split Point of 0.827

Figure 9 provides information for calculating misclassification rate in identifying "at-risk" students using an optimal split point of 0.827 based on Module 030 Phase II data.

		Predicted Passing	
		FALSE	TRUE
Actual Passing	FALSE	17	25
	TRUE	6	81

Figure 9: Prediction Summary Table Based on Optimal Split Point of 0.827

From Figure 9, it can be seen that the actual pass rate for those with predicted post-test scores less than 0.827 was 26%, while the actual pass rate for those with predicted post-test score of at least 0.827 was 76%. The misclassification rate has decreased to 31/129 = 24%. While the misclassification rate is not especially remarkable, it does identify a small subset of students (13%) who are very unlikely to pass the class based on measurements available at the beginning of the semester and can be identified early in the course as most "at-risk". It is important to note that a predicted post-test score of at least 0.827 does not guarantee course completion.

Answers to the Research Questions

Research Question 1 (Phase I models): Is there a statistically significant difference in student learning between MML and Educo groups associated with the teaching method?

Model 2 results suggest that there is a statistically significant difference in student learning between MyMathLab and Educo developmental mathematics courses at SCC, however, the p-value was 0.0861 for this predictor. The positive coefficient of the categorical variable MML implies that students enrolled in module 050 MML classes showed greater improvement than Educo students. However, there are several confounding variables associated with teaching method, therefore, improvements in student learning cannot be solely attributed to the teaching method. Model 1 does not include the MML variable which suggests that for module 040 the methodological indicator is not a significant predictor.

Research Question 2 (Phase I and Phase II models): Are CPT scores and gender statistically significant predictors of student learning in developmental mathematics courses as measured by the difference between pre-test and post-test scores?

The results of this study suggest that Algebra and Arithmetic CPT scores are significant predictors of student academic improvement. However, for each model, less than 50% of variability in student improvement is explained by the linear relationship among the variables. Based on the most recent studies conducted by the Community College Research Center (CCRC) to evaluate the effectiveness of college placement tests ACCUPLACER and COMPASS, the placement accuracy of the CPT scores is low. The recommendation proposed by the CCRC researchers is to combine CPT scores and high school GPA to obtain better placement accuracy [Scott-Clayton, 2012, Jaggars et al., 2013]. Further study is needed to determine whether prediction accuracy improves if CPT scores and gender are combined with other variables to predict student improvement.

The results of this study are inconclusive regarding gender as the predictor of student learning since gender coefficients had different signs in different models, and none of those coefficients were significant at $\alpha = .05$. The implication is that further study is needed to determine whether gender is a significant predictor of student academic improvement in redesigned developmental mathematics classes.

Research Question 3 (Phase II): Which students are most likely to not earn the required grade on the post-test?

Based on module 030 Phase II data, 17 out of the total of 129 students are identified as "at-risk" students who are most likely not to earn the required passing grade of at least .85 on the computerized MyMathLab post-test. Similarly, based on module 050 data, approximately 10% of all students are identified as "at-risk." A total of 10 out of 98 students completing module 050 during Semesters 2 and 3 at SCC had the predicted post-test score below the optimal score of .798 and actual post-test score below the required .85. In comparison, only approximately 3% of students completing module 040 in Phase II are identified as "at-risk" as shown in Appendix 3. Note that the misclassification rates when using the optimal split point are between 19% and 24%. Furthermore, adjusted-R^2 values for these models are between .09 and .26 which implies that only between 9% and 26% of the variability in post-test scores can be explained by the linear models with post-test score as response variable and CPT scores and pre-test score as predictors.

4 SUMMARY AND CONCLUSIONS

In this study, five linear regression models were developed to predict student academic improvement in redesigned developmental mathematics courses using CPT scores, gender and methodological indicators as predictors. Different combinations of predictors were identified based on adjusted-R^2 value to model student academic learning in developmental mathematics courses.

As shown in Figure 10, standardized Algebra and Arithmetic CPT scores were significant predictors of student academic improvement (with the exception of Model 2 where Arithmetic CPT score is not included in the model). In all five models, standardized Algebra CPT score was a significant predictor of student academic improvement at the $\alpha = 0.01$ level. Furthermore, all Algebra CPT score coefficients were negative with same order of magnitude which implies that lower-scoring students improved more than the higher-scoring students as measured by the difference between pre-test and post-test scores. Similarly, Arithmetic CPT score was significant with negative coefficients in four out of five models (with exception of Model 2). The results of including the Reading Comprehension and Sentence Structure CPT scores in the models were inconclusive. The Reading CPT score is included as a predictor in models 2 and 4 with positive coefficients. The Sentence Structure CPT score was included in models 2, 4, and 5, but the coefficients have opposite signs making it difficult to interpret. Similarly, gender is included in models 1, 3, and 4, but the coefficients have different signs. Model 1 implies that females performed better than males but the interpretation of Model 4 would be the opposite. The methodological indicator is significant in Model 2 of Phase I at the significance level $\alpha = 0.05$. Phase II models 4 and 5 include the semester indicator but the results are inconclusive since the coefficients have opposite signs. Based on Model 4, students enrolled in Semester 3 improved more than students in Semester 2 at the significance level $\alpha = 0.05$.

Model	Module	n	Adj-R^2	β_0 (Int)	β_1 (alg)	β_2 (arit)	β_3 (read)	β_4 (sent)	β_5 (gen)	β_6 (MML indicator)	β_7 (sem indicator)
M1	40	81	.3672	-.1563	-.4927***	-.2729**			.3013.		X
M2	50	81	.234	-.1603	-.2298*		.2329.	.2689*		.3488.	X
M3	30	129	.2574	.0796	-.2626**	-.3423***			-.2096	X	
M4	40	98	.2946	-.0303	-.2905**	-.2149*	.1959	-.2708*	-.3510.	X	.3107.
M5	50	98	.4345	.3232**	-.4317***	-.2098*		-.1641*		X	-.5881***

Significance codes: 0 ‘***’ .001 ‘**’ .01‘*’ .05 ‘.’ .1‘’ 1

Figure 10: Standardized Coefficients and Adj.-R^2 for Models 1-5

The predictive models that include CPT scores and gender as the only predictors of learning can be employed to identify “at-risk” students at the beginning of each school semester. In this study an out-of-sample cross-validation method is used to evaluate the misclassification rate in identifying “at-risk” students.

Limitations of this Study

This research has several limitations to the validity of study results. Due to the nature of this study, no random assignment of students into classes was employed. Furthermore, no information was collected regarding student demographics such as age, marital status, number of kids, student motivation and attitudes toward studying mathematics, or their proficiency in using technology to complete assignments. Inclusion of additional variables such as course section number, class size, and final course grade might have improved the adjusted-R^2 for each reduced final model. However, these variables were not included in the models since the goal of the study was to predict student improvement using CPT scores and gender, not necessarily to obtain the highest adjusted-R^2 value.

In Phase I, MyMathLab and Educo classes were taught by different instructors at two different campuses where students were using different software and textbooks for tutorials and completion of assignments. Hence, some of the differences in student academic improvement not explained by the multiple regression models might be attributable to the quality of software or textbooks or differences in teaching styles among the instructors. Furthermore, researcher-created pre-test and post-tests for modules 030, 040, and 050 were limited in how well they measured student learning.

Finally, the statistical models were based on the data collected at a rural community college in North Carolina in developmental mathematics classes. Hence, there are limitations to what extent the results can be generalized to predict student academic improvement at other community colleges.

References

[Belfield and Crosta, 2012] Belfield, C. and Crosta, P. M. (2012). Predicting success in college: The importance of placement tests and high school transcripts. Technical Report 42, Community College Research Center.

[Condelli, 2006] Condelli, L. (2006). A Review Of The Literature In Adult Numeracy: Research And Conceptual Issues. Technical report, American Institutes for Research, and Washington, DC. Div. of Adult Education and Literacy. Office of Vocational and Adult Education (ED). US Department Of Education.

[Department of Education, 2005] Department of Education, U. S. (2005). Strengthening mathematics skills at the postsecondary level: Literature and analysis. Technical report, Office of Vocational and Adult Education, Division of Adult Education and Literacy, Washington, DC.

[Dogan, 2001] Dogan, H. (2001). *A Comparison Study between a Traditional and Experimental First Year Linear Algebra Program*. PhD thesis, University of Oklahoma, Norman.

[Emerson and Taylor, 2004] Emerson, T. L. and Taylor, B. A. (2004). Comparing student achievement across experimental and lecture-oriented sections of a principle of microeconomics course. *Southern Economic Journal*, 70(3):672–693.

[Epper and Baker, 2009] Epper, R. M. and Baker, E. D. (2009). Technology solutions for developmental math: An overview of current and emerging practices. `http://www.gatesfoundation.org/learning/Documents/technology-solutions-for-developmental-math-jan-2009.pdf`.

[Faraway, 2002] Faraway, J. J. (2002). *Practical Regression and Anova using R*.

[Fleming, 2002] Fleming, J. (2002). Who will succeed in college? When the SAT predicts black students' performance. *The Review of Higher Education*, 25(3):281–296.

[Hannafin and Foshay, 2008] Hannafin, R. D. and Foshay, W. R. (2008). Computer-based instruction's (CBI) rediscovered role in K-12: An evaluation case study of one high school's use of CBI to improve pass rates on highstakes tests. *Educational Technology Research & Development*, 56(2):147–160.

[Harachiewicz et al., 2002] Harachiewicz, J. M., Barron, K. E., Tauer, J. M., and Elliot, A. J. (2002). Predicting success in college: A longitudinal study of achievement goals and ability measures as predictors of interest and performance from freshman year through graduation. *Journal of Educational Psychology*, 94(3):562–575.

[Hendricks, 2012] Hendricks, G. I. (2012). *Predictors of Success for Community College Developmental Mathematics Students in Online, Hybrid, and Traditional Courses*. PhD thesis, Appalachian State University, Boone, North Carolina.

[Hodara, 2011] Hodara, M. (2011). Improving pedagogy in the developmental mathematics classroom. Technical Report 51, Columbia University, Teacher's College, Community College Research Center, New York, NY.

[Hodara et al., 2012] Hodara, M., Jaggars, S. S., and Karp, M. M. (2012). Improving developmental education assessment and placement: Lessons from community colleges across the country. Technical report, Community College Research Center.

[Hsu, 2003] Hsu, Y. (2003). The effectiveness of computer-assisted instruction in statistics education: A meta-analysis. *Dissertation Abstracts International*, 64:1526.

[Huang, 2011] Huang, S. (2011). *Predictive Modeling and Analysis of Student Academic Performance in an Engineering Dynamics Course*. PhD thesis, Utah State University.

[Jaggars, 2011] Jaggars, S. S. (2011). Online Learning: Does it Help Low-Income or Underprepared Students. Technical Report 26, Community College Research Center.

[Jaggars et al., 2013] Jaggars, S. S., Hodara, M., and Stacey, G. W. (2013). Designing Meaningful Developmental Reform. Technical report, Community College Research Center.

[Kinney and Robertson, 2003] Kinney, D. P. and Robertson, D. F. (2003). Technology makes possible new models for delivering developmental mathematics instruction. *Mathematics and Computer Education*, 37(3):315–328.

[Loving, 2007] Loving, M. L. (2007). *Factors influencing African American students' attitude, comprehension, and course completion in computerized mathematics courses*. PhD thesis, University of Southern Mississippi, Hattiesburg.

[Lowis and Castley, 2008] Lowis, M. and Castley, A. (2008). Factors affecting student progression and achievement: Prediction and intervention. *Innovations in Education and Teaching International*, 45(4):333–343.

[Lykourentzou et al., 2009] Lykourentzou, L., Giannoukos, L., and Mpardis, G. (2009). Early and dynamic student achievement prediction in e-learning courses using neural networks. *Journal of the American Society for Information Science and Technology*, 60(2):372–380.

[Marsh et al., 2008] Marsh, C. M., Vandehey, M. A., and Diekhoff, G. M. (2008). A comparison of an introductory course to SAT/ACT scores in predicting student performance. *The Journal of General Education*, 57(4):244–255.

[Mejri, 2011] Mejri, S. B. (2011). *Will the Use of My Math Lab Influence the Students' Attitude toward Mathematics, Course Completion, and Content Comprehension of Basic College Mathematics?* PhD thesis, University of Toledo, Toledo, Ohio.

[Moosavi, 2009] Moosavi, A. S. (2009). *A comparison of two computer-aided instruction methods with traditional instruction in freshmen college mathematics classes*. PhD thesis, University of Alabama, Tuscaloosa.

[NCCCS, 2011] NCCCS (2011). DEI update: Rethinking developmental math in North Carolina.

[NCMATYC, 2012] NCMATYC, editor (2012). *North Carolina Community College System Developmental Math Modular Curriculum Module Outlines and Notes*. NCMATYC.

[Neter et al., 1996] Neter, J., Kutner, M. H., Nachtsheim, C. J., and Wasserman, W. (1996). *Applied Linear Statistical Models*. McGraw-Hill, 4th edition.

[Reagan, 2004] Reagan, L. (2004). *The effectiveness of computer-assisted instruction in developmental mathematics classes*. PhD thesis, Texas A&M University.

[Scott-Clayton, 2012] Scott-Clayton, J. (2012). Do high-stakes placement exams predict college success? Technical Report 41, Community College Research Center.

[Speckler, 2012] Speckler, M. (2012). *Making the Grade, V.5: 77 data-driven case studies illustrating how the MyMathLab family of products supports student achievement*. Pearson Education, Boston, MA.

[Spradlin, 2009] Spradlin, K. D. (2009). *The effectiveness of computer-assisted instruction in developmental mathematics*. PhD thesis, Liberty University, Lynchburg, VA.

[Teal, 2008] Teal, B. D. (2008). *A comparative analysis of modes of instruction using student test scores in developmental mathematics*. PhD thesis, Morgan State University, Baltimore, MD.

[Twigg, 2005] Twigg, C. A. (2005). Increasing success for underserved students: Redesigning introductory courses. Technical report, National Center for Academic Transformation., Saratoga Springs, NY.

[Twigg, 2011] Twigg, C. A. (2011). The math emporium: A silver bullet for higher education. *Change: The Magazine of Higher Learning*, 43(3):25–34.

[Veenstra et al., 2009] Veenstra, C. P., Dey, E. L., and Herrin, G. D. (2009). A model for freshman engineering retention. `http://advances.asee.org/vol01/issue03/papers/aeevol01-issue03-p07.pdf`.

[Villarreal, 2003] Villarreal, L. M. (2003). A step in the positive direction: Integrating a computer laboratory component into developmental algebra courses. *Mathematics and Computer Education*, 37(1):72–78.

[Ware and Galassi, 2006] Ware, W. B. and Galassi, J. P. (2006). Using correlational and prediction data to enhance student achievement in K-12 schools: A practical application for school counselors. *Professional School Counseling*, 9(5):344–356.

[Waycaster, 2001] Waycaster, P. (2001). Factors impacting success in community college developmental mathematics and subsequent courses. *Community College Journal of Research and Practice*, 25:403–415.

[Zachary, 2008] Zachary, E. (2008). *Promising Instructional Reforms in Developmental Education. A Case Study of Three Achieving the Dream Colleges*. MDRC, New York.

APPENDICES

Appendix 1: Full Models Normality Plots

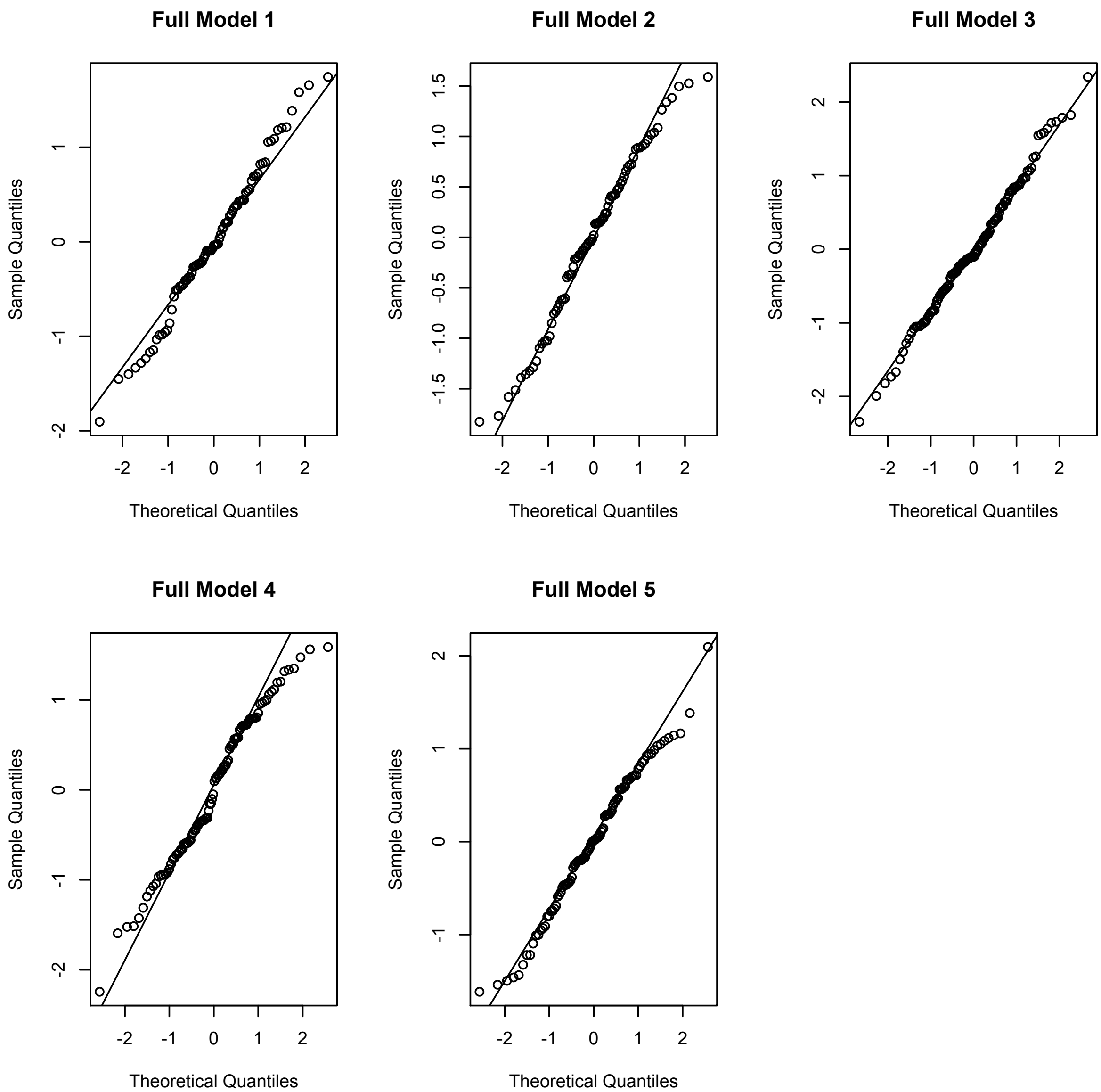


Figure 11: Normality Plots for Full Models 1-5

Appendix 2: Reduced Models Normality Plots

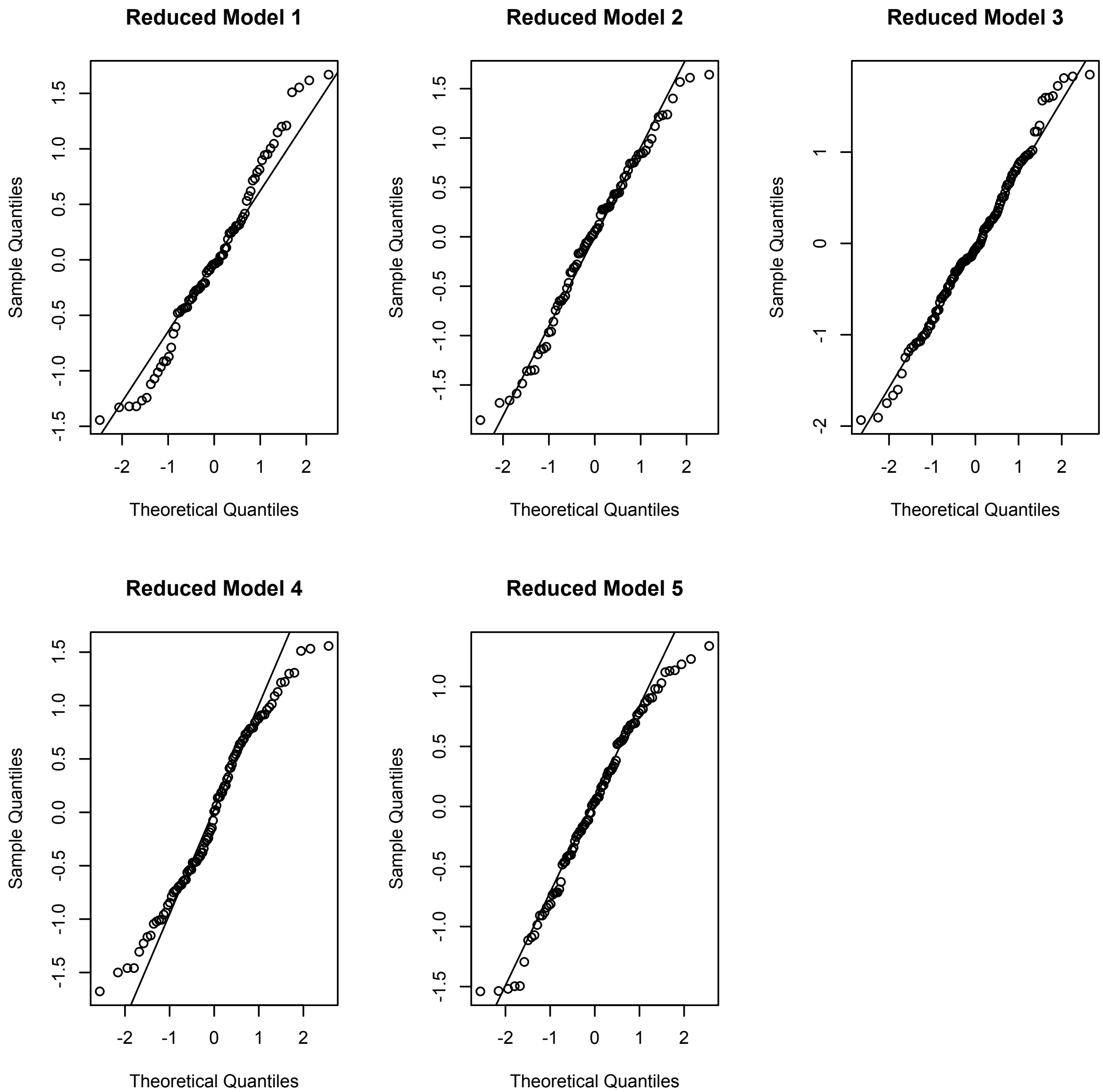
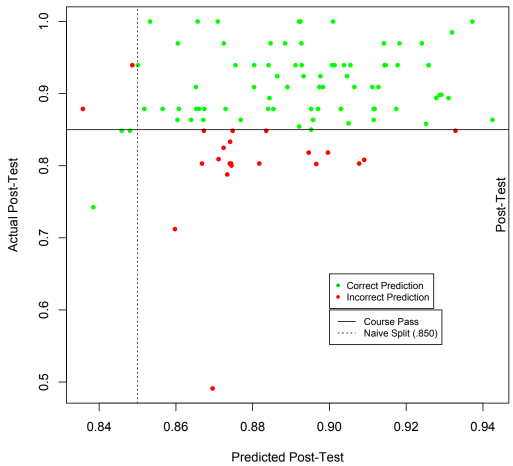
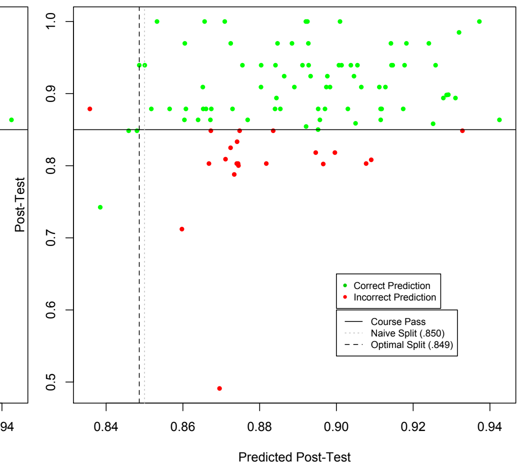


Figure 12: Normality Plots for Reduced Models 1-5

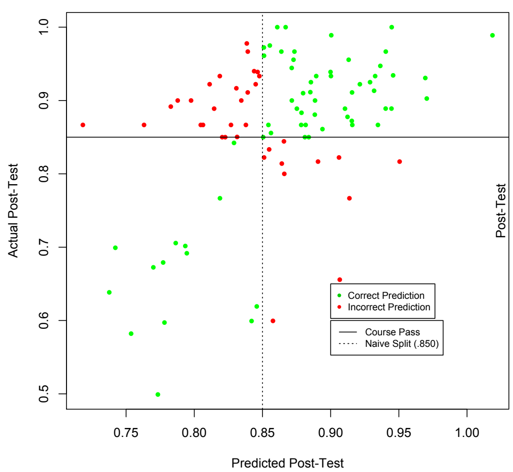
Appendix 3: Prediction Summary Scatterplots



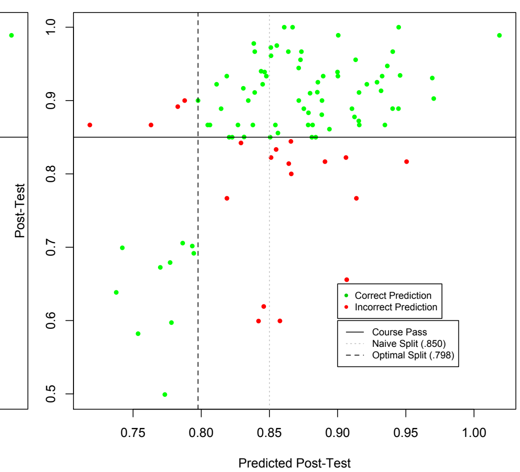
Plot 1



Plot 2



Plot 3



Plot 4

Figure 13: Scatterplots for Post-Test Models Based on Phase II Data: Module 040 Data: Plot 1 and 2; Module 050 Data: Plot 3 and 4

Appendix 4: R code for Model 1

```
#Module 1 R code
#SP12 module 040; standardized data;
#total of 81 students (together in Educo and MyMathLab classes)

#5 number summary of raw data
summary(de45$me4S)
summary(de45$algS)
summary(de45$aritS)
summary(de45$readS)
summary(de45$sentS)

me4S.6 = lm(me4S~algS+aritS+readS+sentS+gen+MML, de45)
 summary(me4S.6)

#plots of histograms, normality plot, fitted vs predicted,
scatterplot matrix
 par(mfrow=c(2,2))
 hist(de45$me4S)
 hist(me4S.6$residuals)
 qqnorm(me4S.6$residuals); qqline(me4S.6$residuals)
 plot(me4S.6$fitted.values,me4S.6$residuals,xlab='Predicted',
 ylab='Residual')
 abline(h=0);
 #scatterplot matrix
 pairs(~algS+aritS+readS+sentS+gen+me4S, de45, main=
 "Matrix SP12070me4S");
```

```
#plots of y vs each x_{i}
 par(mfrow=c(3,2))
 plot(de45$algS, de45$me4S,xlab='CPT Algebra Score',
 ylab='Y value'); abline(lm(de45$me4S~de45$algS));
 plot(de45$aritS, de45$me4S,xlab='CPT Arithmetic Score',
 ylab='Y value'); abline(lm(de45$me4S~de45$aritS));
 plot(de45$readS, de45$me4S,xlab='CPT Reading Score',
 ylab='Y value'); abline(lm(de45$me4S~de45$readS));
 plot(de45$sentS, de45$me4S,xlab='CPT Sentence Structure Score',
 ylab='Y value'); abline(lm(de45$me4S~de45$sentS));

#backward elimination
 #minus MML: me4S.5
 me4S.5 = lm(me4S~algS+aritS+readS+sentS+gen, de45)
 summary(me4S.5)
#minus read: me4S.4
 me4S.4 = lm(me4S~algS+aritS+sentS+gen,de45)
 summary(me4S.4)
#minus sentS: me4S.3
 me4S.3 = lm(me4S~algS+aritS+gen,de45)
 summary(me4S.3)
#minus gen: me4S.2
 me4S.2 = lm(me4S~algS+aritS,de45)
 summary(me4S.2)
 #minus aritS: me4S.1
 me4S.1 = lm(me4S~algS,de45)
```

```
 summary(me4S.1)

#interactions models
 me4S.4.inter = lm(me4S~algS+aritS+readS+algS*MML,de45)
 summary(me4S.4.inter)
me4S.3.inter1 = lm(me4S~algS+aritS+algS*MML,de45)
 summary(me4S.3.inter1)
me4S.3.inter2 = lm(me4S~algS+aritS+gen+algS*aritS,de45)
 summary(me4S.3.inter2)
me4S.3.inter3 = lm(me4S~algS+aritS+gen+algS*aritS+algS*gen,de45)
 summary(me4S.3.inter3)

#forward selection
#add algS: me4S.1
 me4S.1 = lm(me4S~algS,de45)
 summary(me4S.1)
#add aritS: me4S.2
 me4S.2 = lm(me4S~algS+aritS,de45)
 summary(me4S.2)
#add gen: me4S.3
 me4S.3 = lm(me4S~algS+aritS+gen,de45)
 summary(me4S.3)
#add sentS: me4S.4
 me4S.4 = lm(me4S~algS+aritS+sentS+gen,de45)
 summary(me4S.4)
#add readS: me4S.5
 me4S.5 = lm(me4S~algS+aritS+readS+sentS+gen, de45)
```

```
 summary(me4S.5)
#add MML: me4S.6
 me4S.6 = lm(me4S~algS+aritS+readS+sentS+gen+MML, de45)
 summary(me4S.6)

#removing influential observations
 me4S.4 = lm(me4S~algS+aritS+sentS+gen,de45)
 influence.measures(me4S.4)
me4S.4.obs=c(7,22,28,48);
 me4S.4.inf = lm(me4S~algS+aritS+sentS+gen,de45[-me4S.4.obs,]);
 summary(me4S.4.inf)

#diagnotics on final reduced model
hist(me4S.4.inf$residuals);
qqnorm(me4S.4.inf$residuals, main="Reduced Model 1");
qqline(me4S.4.inf$residuals);
#Shapiro-Wilk test
 shapiro.test(me4S.4.inf$residuals)
```

Appendix 5: R code for Post-Test Models in Phase II

```
#FA12 and SP13 module 030 (class MAT060); standardized data;
#total of 129 students
summary(d2)

# backward variable selection
d2m1 = lm(m3post~alg+arit+read+sent+gen+m3pre,d2); summary(d2m1)
d2m2 = lm(m3post~alg+arit+read+gen+m3pre,d2); summary(d2m2)
d2m3 = lm(m3post~alg+read+gen+m3pre,d2); summary(d2m3)
d2m4 = lm(m3post~alg+gen+m3pre,d2); summary(d2m4)
d2m5 = lm(m3post~alg+m3pre,d2); summary(d2m5)
d2m6 = lm(m3post~m3pre,d2); summary(d2m6)

#forward selection:
d2m6f = lm(m3post~m3pre,d2); summary(d2m6f)
d2m5f = lm(m3post~m3pre+alg,d2); summary(d2m5f)
d2m4f = lm(m3post~m3pre+alg+gen,d2); summary(d2m4f)
d2m3f = lm(m3post~m3pre+alg+gen+read,d2); summary(d2m3f)
d2m2f = lm(m3post~m3pre+alg+gen+read+arit,d2); summary(d2m2f)
d2m1f = lm(m3post~m3pre+alg+gen+read+arit+sent,d2); summary(d2m1f)

remove.second = c(1,2,3,4,5)
d3 = d2[,-remove.second]
summary(d3)

# fit regression model to remaining n-1 points and predict post-test
#score for left-out point
```

```
n = nrow(d3)

pred.post = numeric(n)

i=1

for(i in 1:n){

d3m = lm(m3post~m3pre,d3[-i,]); summary(d3m)

pred.post[i] = t(matrix(as.numeric(c(1,d3[i,-1]),nrow=1)))

%*%matrix(as.numeric(d3m$coefficients,nrow=1))}


pred.post.pass = pred.post >= .85

post.pass = d3$m3post >= .85

#Discriminate students based on whether their predicted post-test

#grade is .85 or greater

# Passing defined as actual post-test >= .85

table(post.pass,pred.post.pass)


# graph predpost-naive

plot(pred.post,d3$m3post,ylab='Actual Post-Test',xlab='Predicted

Post-Test',pch='')

points(pred.post[post.pass & pred.post.pass],d3$m3post[post.pass

& pred.post.pass],col='green',pch=20)

points(pred.post[!post.pass & !pred.post.pass],d3$m3post[!post.pass

& !pred.post.pass],col='green',pch=20)

points(pred.post[!post.pass & pred.post.pass],d3$m3post[!post.pass

& pred.post.pass],col='red',pch=20)

points(pred.post[post.pass & !pred.post.pass],d3$m3post[post.pass

& !pred.post.pass],col='red',pch=20)

abline(h=0.85)
```

```
abline(v=0.85,lty=3)
legend(0.90,0.60,legend=c('Course Pass','Naive Split (.850)'),
lty=c(1,3),col=c(1,1),cex=.75)
legend(0.90,0.65,legend=c('Correct Prediction','Incorrect
Prediction'),pch=20,col=c(3,2),cex=.75)

# find optimum split point to minimize the misclassification rate;
misclass.matrix = matrix(NA,nrow=n,ncol=5)
misclass.matrix[,1] = d3$m3post
misclass.matrix[,2] = pred.post
for(i in 1:n){
misclass.matrix[i,3] = length(d3$m3post[ d3$m3post >= .85
& pred.post < pred.post[i] ])
misclass.matrix[i,4] = length(d3$m3post[ d3$m3post < .85
& pred.post >= pred.post[i] ])
}
misclass.matrix[,5] = misclass.matrix[,3]+misclass.matrix[,4]
opt.split = misclass.matrix[which.min(misclass.matrix[,5]),2]
opt.split
misclass.matrix

pred.post.pass.opt = pred.post >= opt.split
#Discriminate students based on whether their predicted post-test
#grade is .827 or greater
# Passing defined as actual post-test >= .85
table(post.pass,pred.post.pass.opt)
```

```
# graph predpost-optimal
plot(pred.post,d3$m3post,ylab='Post-Test',xlab='Predicted
Post-Test',pch='')
points(pred.post[post.pass & pred.post>=opt.split],d3$m3post
[post.pass &pred.post>=opt.split],col='green',pch=20)
points(pred.post[!post.pass & pred.post<opt.split],d3$m3post
[!post.pass & pred.post<opt.split],col='green',pch=20)
points(pred.post[!post.pass & pred.post>=opt.split],d3$m3post
[!post.pass & pred.post>=opt.split],col='red',pch=20)
points(pred.post[post.pass & pred.post<opt.split],d3$m3post
[post.pass & pred.post<opt.split],col='red',pch=20)
abline(h=0.85)
abline(v=0.85,lty=3,col='gray')
abline(v=opt.split,lty=2)
legend(0.90,0.60,legend=c('Course Pass','Naive Split (.850)',
'Optimal Split (.827)'),
lty=c(1,3,2),col=c(1,'gray',1),cex=.75)
legend(0.90,0.65,legend=c('Correct Prediction','Incorrect
Prediction'),pch=20, col=c(3,2),cex=.75)
```