Curriculum-Based Measurement Performance Indicators: A Tool for Undergraduate Calculus Students to Inform and Direct Their Learning Behavior

by

Linda W. Sturges

A dissertation submitted to the Graduate Faculty in Educational Psychology in partial
fulfillment of the requirements for the degree of Doctor of Philosophy,
The City University of New York

2010

This manuscript has been read and accepted for the
Graduate Faculty in Educational Psychology in satisfaction of the
dissertation requirement for the degree of Doctor of Philosophy.

Barry J. Zimmerman Ph.D.

Date

Chair of Examining Committee

Mary Kopala Ph.D.

Date

Executive Officer

David Rindskopf Ph.D.

Hope Hartman Ph.D.

Supervision Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

CURRICULUM-BASED MEASUREMENT PERFORMANCE INDICATORS: A TOOL FOR UNDERGRADUATE CALCULUS STUDENTS TO INFORM AND DIRECT THEIR LEARNING BEHAVIOR

by

Linda Wolchik Sturges

Adviser: Professor Barry J. Zimmerman

The present study investigated the extent to which providing students with individualized performance feedback informed and directed their learning behavior. Individualized performance feedback was delivered to students using curriculum-based measurement progress indicators, either as a visual representation of ongoing performance in the form of a progress graph or as a progress graph supplemented with a qualitative analysis of topic mastery.

Participants were 67 students enrolled in a first course in engineering calculus at a specialized public 4-year college. The College's specialization is within the maritime industry. Intact sections of Calculus I were randomly assigned to each of the feedback conditions. A contrast group of students who did not receive individualized performance feedback was formed.

The impact of individualized performance feedback was examined in terms of measures of calibration accuracy, relearning, and academic performance when contrasted to corresponding measures from students who did not receive individualized performance feedback. Mixed-model analysis of covariance and mixed-model analysis of variance revealed differences between the feedback groups and the no feedback group. Differences for calibration accuracy approached statistical significance; however differences were statistically significant for measures of relearning and academic performance. For each measure, the students who received individualized performance feedback used that information to better judge their calculus capability, relearn topics not mastered more often, achieve at a higher level on course exams.

When comparing the two feedback groups, there was evidence of differential self-monitoring activities. When students used the supplementary information gleaned from the mastery analyses, they studied non-mastered topics more often and had consistent study habits. Additionally, the group that received the supplementary mastery analyses not only reported more positive expressions of the usefulness of the feedback information, they also had strong associations between their perceptions of the usefulness of the feedback and corresponding measures of time engaged in academic activities and instances of relearning.

Implications of the findings of this study suggest performance indicators appear to empower the student with the information to inform and direct one’s learning behavior to become a successful learner.

Acknowledgements

I wish to thank you, Dr. Zimmerman, for your guidance over the years. I am grateful that you were able to incorporate my mathematics teaching background in our research. I consider myself most fortunate to have had the opportunity to work with you and benefit from your experience.

To my dissertation committee members, you have had a profound influence on what I do professionally. Dr. Hartman, I have integrated computer-based learning environments into my teaching practices, with the emphasis on making them meaningful learning experiences for my students. Dr. Rindskopf, I have become a better statistics teacher and have been able to apply our coursework to my institutional assessment activities. Dr. Zimmerman, you have enriched my experiences as a practicing educator. I have been able to extend this research on performance indicators and monitoring into my teaching practices. I thank you all for your professionalism.

The journey to this point would not have been possible without the support of family and colleagues. To my Provost, Joe Hoffman, and my Curriculum Supervisor, Gil Traub, thank you for your encouragement and accommodating my requirements to pursue this goal. To my sister, Doreen, your words of encouragement have sustained my efforts.

To my dearest Craig, over the years you have always encouraged me to see the possibilities in life. Thank you for giving me the room to grow. To our daughter, Heather, you had to share your undergraduate experience with both parents returning to graduate school. Thank you for the “Mommy time” and your support. How lucky am I to have you as my daughter and friend.

To the faculty of the Educational Psychology program, and especially my readers Dr. Chen and Dr. Homer, I thank your for sharing your expertise with me.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

What is a successful learner? In 1997, the American Psychological Association (APA) formulated a set of fourteen learner-centered psychological principles that are primarily under the control of the learner and interact with environmental and contextual conditions. The report, Learner-Centered Psychological Principles: A Framework for School Reform and Redesign, states that "learning in a school environment is most effective when it is an intentional process for students to derive meaning from information, experiences, thoughts and beliefs" (APA, p.3). The fourteen principles are embedded within four themes referring to cognitive and metacognitive, motivational and affective, developmental and social, and individual differences factors that influence learners and their learning. The report encapsulates the essence of a successful learner as one who is "active, goal-directed, self-regulating, and assumes personal responsibility for contributing to one's own learning" (p. 3). The final principle discusses the role of assessment and performance feedback as a source of information to the teacher, to refine teaching when topics have not been mastered, and to the student to set goals, improve self-assessment, raise motivation, and self-direct learning (APA, 1997). This research examines the role of individualized performance feedback, using curriculum-based measurement procedures, as a tool to enable students to make personal changes to become successful learners.

In order to promote a proactive response by students to their learning behaviors, students participating in this study were provided with personalized performance indicators. Performance indicators communicate achievement feedback, in the form of a graphical representation of academic progress and as a continuum of topic mastery. This study uses an established progress monitoring procedure, known as curriculum-based measurement, in a novel way to determine whether a first semester engineering calculus student would make personal instructional decisions based on performance feedback.

Curriculum-Based Measurement operates as an educational framework for evaluating test-based decisions. Measures are drawn from curriculum material, are administered and scored in a standardized manner, depict progress graphically against a performance goal, and have decision-based rules regarding raising a performance goal or adjusting the instruction. This study follows the general CBM criteria of creating testing probes from the curriculum, uniformly administering and scoring the probes, and generating performance indicators of progress. However in this research, it is the student, not the teacher, who makes instructional decisions based on individualized performance feedback. This research investigates the extent to which the performance indicators inform and direct the student's learning behavior. The goal of this study is to inspect the psychological dimensions of a successful learner, in terms of metacognitive factors of monitoring and control of one's learning, setting performance goals, and making changes in one's learning behavior to meet those goals, within a curriculum-based measurement progress monitoring format.

The ensuing discussion elaborates on the APA's learner-centered principles and the relationship between these principles and the psychological dimensions of this study. This chapter also provides an overview of current recommendations for higher education from the Spellings Commission and accreditation agencies. Several indices of student engagement, a summary of data on mathematics proficiency of first year college students, and a brief description of curriculum-based measurement procedures and its evolving applications follow. The chapter concludes with the purpose of this study and its contribution to the current body of knowledge on the instructional utility of curriculum-based measurement as a tool for students to become successful learners. This research also extends the study of monitoring accuracy from a discrete content area of educational psychology to the content area of calculus that is comprised of topics that are integrated and build upon each other.

Learner-Centered Psychological Principles

The principles for learners and learning (APA, 1997) are categorized into four areas: cognitive and metacognitive factors, motivational and affective factors, developmental and social factors, and individual differences factors. In order for the learner to construct meaningful representations of knowledge and to acquire the thinking and learning strategies necessary for continued learning success, the learner must generate and practice personally relevant goals. Schunk (1991) defines a goal as something an individual is trying to accomplish, whereas goal setting involves establishing a goal and modifying it as needed. Successful learners are goal-oriented. In this study, I asked students to set personal goals, and I observed under what conditions learners change their goals. Zimmerman (2000) describes the nature of a high achiever as one who sets more specific learning goals, uses more strategies to learn, self-monitors learning progress more frequently, and more systematically adapts one's efforts on the basis of learning outcomes. Schunk (1991) links attainment of a goal with a greater sense of self-efficacy. Heightened self-efficacy sustains motivation to attain goals and improves skill development. When goals are realistic, students are more apt to monitor progress and make adjustments to their task approach.

Cognitive and metacognitive factors. The third learner-centered principle (APA, 1997) acknowledges that knowledge is constructed when students integrate, through meaningful links, prior knowledge with new information. To encourage the integration of previous and current knowledge, weekly formative assessments given during this study include items that test factual knowledge or the implementation of a single strategy, applied knowledge, and conceptual knowledge of the material within each unit of study.

The learner-centered principles (APA, 1997) describe successful learners as having an extensive repertoire of strategies and more importantly reflect upon their usefulness. Performance feedback facilitates the reflective processes; this study returns to the student both formative and summative feedback. Additionally, successful learners possess a metacognitive awareness of how they think and learn and monitor their progress toward learning goals. A metacognitively aware individual is an efficient decision maker who questions, monitors, and instructs oneself to gain access to relevant information, formulates a plan of action, guides the execution of that plan, and regulates the use of cognitive strategies throughout the decision-making process (Batha & Carroll, 2007, p. 64).

Recent research has focused on the accuracy of metacognitive judgments made by college students over an extended period of time and supports the link between metacognitive skills and performance. These studies highlight the view that accurate monitoring judgments can optimize test preparedness efforts of students. Research provides support for improvement in monitoring accuracy over time (Hacker et al., 2000, Nietfeld et al., 2005, Nietfeld et al., 2006) and the positive relationship between accuracy and performance (Hacker et al., 2000, Nietfeld et al., 2005, Nietfeld et al., 2006). Additional research by Nietfeld, Cao, and Osborne (2006) linked monitoring exercises and accuracy feedback with monitoring accuracy, performance, and self-efficacy.

Everson and Tobias (2001) describe students with effective metacognitive skills as those who "accurately estimate their knowledge in a variety of domains, monitor their on-going learning, update their knowledge, and develop effective plans for new learning" (p. 69). All students in this study were given topic self-monitoring forms to guide their learning decisions. Research by Lan (1996) supports the finding that when college-students are involved in a self-monitoring process, they become alert to the effectiveness of currently used learning strategies and the appropriateness of their learning environment. This alertness helps student select the most effective strategies and arrange their environment to maximize learning.

Motivational and affective factors. Motivational and emotional factors also influence how the learner thinks and processes information, as well as, the individual's motivation to learn. One of the strengths of curriculum-based measurement practices is the ability to see weekly changes in progress. The dynamic nature of the index can serve as a motivational factor to guide one's learning practices. Motivation to learn is often influenced by the type of learning tasks, as well as providing students with a sense of choice and control over these tasks. Kitsantas and Zimmerman (2006) examined how the combination of performance standards, absolute (as measured as success or failure) or graduated (as measured as a continuum of progress toward the ultimate goal) and recording of score (either as a summative score or as a continuum of online progress) affects the accuracy of dart throwing skill. There were main effects of both graduated standards and online graphing of practice score on dart throwing skill, self-efficacy, and perceptions of improvement. This research (Kitsantas and Zimmerman, 2006) lends additional support for the effects of motivation on learning in that graduated standards had a main effect on measures of self-satisfaction with one's performance and the group that graphed measures of online performance contributed their performance to controllable factors. The research also extended the use of a graphed performance indicator from a curriculum content area to a motoric skill area.

Developmental and social factors. The learner-centered principles acknowledge the influence of developmental and social factors on learning. Educators need to be aware of developmental differences in how students learn and provide instruction that is of an appropriate level. Learning occurs in social or collaborative environments and teachers should provide opportunities to students to experience different instruction venues. Regarding individual differences factors, teachers can help students to examine their learning practices and assist students with modifying them, if necessary. Providing students with their performance information promotes opportunities for students to examine their learning practices.

Individual differences factors. The final principle of the framework discusses ongoing assessments as integral to the learning processes. Bandura and Cervone, as cited by Schunk (1991) observed that when goals are combined with progress feedback both performance and self-efficacy for goal attainment were raised. Bransford et al. (2000) elaborate upon learner-centered principle fourteen regarding formative assessments and feedback. Assessments should have the purpose to inform the learner of his or her progress and also to be an instrument that demonstrates progress over time. Curriculum-based measurement performance indicators fulfill both purposes. Bransford et al. also find that “feedback is most valuable when students have opportunities to use it to correct their thinking” (p. 130). Additional assessment occasions have been shown to have a positive benefit on student learning and transfer, as students have the opportunity to revise their learning behavior. This study also provided the learner with ongoing formative assessments and the conditions to relearn material that has not been mastered. This study supplies all students with a weekly progress graph. Some students receive additional feedback as an analysis of mastery level per topic. Individualized performance feedback information equips the student with the tools to be a successful learner (a) by comparing present performance against mastery level performance and with pervious performance, (b) by self-regulating oneself in terms of study topics, and (c) being personally responsible for one’s own learning.

Goals for Higher Education

One of the greatest challenges educators face is not the transmission of knowledge, but rather assessing whether the information was processed and utilized to some end other than repetition by the students. These challenges have been articulated in various reports and identified as outcomes of educational processes. Findings on the current state and future goals of higher education, as reported in the U. S. Department of Education’s “A Test of Leadership”, also known as the report of the Spellings Commission of 2006, include a recommendation that “postsecondary educational institutions … measure and report meaningful student learning outcomes” (p. 23). Learning outcomes include not only assessment of knowledge areas, such as critical thinking skills, written communication, and quantitative reasoning, but also assessments of engagement or “the time and effort students put into educational activities in and out of the classroom” (USDOE, p. 22).

Accreditation agencies, such as the Middle States Commission on Higher Education (MSCHE) have articulated many of the same conclusions as the U. S. Department of Education in their published characteristics of excellence for accreditation. The “Characteristics of Excellence in Higher Education” (MSCHE, 2006) explicate standards for colleges and universities to meet for accreditation. Standard 7 focuses on institutional assessment with an emphasis on support of student learning and Standard 14 addresses student learning and the use of assessment findings to improve teaching and to assist students in improving their learning. More importantly the Standards ask institutions not only to gather data but to show evidence of how this information is used to improve the effectiveness of the institution and the quality of the learning experience as the final step of an assessment cycle. This study adopts the Middle States teaching-learning-assessment cycle of assessing students’ learning and using that information to direct teaching and learning and applies it at a micro or individual level. To address these accreditation requirements, there is an increased availability of information on the nature of the student as a learner and measures of student learning.

The Current Profile of Undergraduate Students. In response to both the report of the Spellings Commission and the “Characteristics for Excellence in Higher Education”, colleges have begun to examine indices of student engagement in an attempt to understand the learning habits and educational impressions of their student population as part of institutional effectiveness. These measures of student engagement serve as a benchmark against which colleges provide services to address shortcomings. The National Survey of Student Engagement (NSSE) and Your First College Year (YFCY) are two national instruments of student engagement. Both of these surveys report annually their most recent findings and concur that in terms of student engagement, students spend on average at most one hour per credit hour outside of class in preparation for each course.

The National Survey of Student Engagement Annual Report of 2008 confirmed that the average number of hours that college students spend studying per week has remained constant since 2001 at 13 to 14 hours per week. In fact, only 33% of freshman respondents report spending more than 15 hours per week preparing for their courses. More dramatic indications of student disengagement from academic activities are enumerated in the findings of the 2007 administration of the “Your First College Year” survey (Higher Education Research Institute, 2007). National aggregates indicate that 29.2% of respondents spend 11 or more hours a week studying or doing homework and 37.5% of respondents spend less than 6 hours on these activities. Other indices from the two instruments support the perception of freshman apathy in that 78.7% of first year students ‘frequently’ or ‘occasionally’ turned in assignments that did not reflect their best work (YFCY, 2007) and 60% of freshman engineering students ‘sometimes’ came to class without completing readings or written assignments (NSSE, 2008). In addition to measures of engagement of freshman college students, their academic preparedness should also be considered. Preparedness is discussed in the domain of mathematics.

The National Assessment of Educational Policy (NAEP) assesses mathematics proficiency at grades 4, 8, and 12. The data for grade 12 provides insight into the level of mathematical proficiency with which students enter college. The most recent grade 12 report of 2005 highlights these findings.

The National Assessment Governing Board sets policy for the National Assessment of Educational Performance or NAEP. To better reflect changes in high school mathematics standards and coursework, the NAEP assessment of mathematics for grade 12 was redesigned in 2005 to include more questions on algebra, data analysis, and probability. This new framework however does not allow for a direct comparison of the results of 2005 with previous mathematics grade 12 results.

Results of the NAEP assessments are reported for both the nation and states in terms of the Governing Board's achievement levels -- basic, proficient, and advanced. Twelfth-grade students performing at the *Basic* level should be able to solve mathematical problems that require the direct application of concepts and procedures in familiar situations. Twelfth-grade students performing at the *Proficient* level should be able to select strategies to solve problems and integrate concepts and procedures. Twelfth-grade students performing at the *Advanced* level should demonstrate in-depth knowledge of the mathematical concepts and procedures represented in the framework. The NAEP framework is a guide to balance content items with items that tap into a student's ability to 'know and do' mathematics The Nation's 2005 Report Card states that 61% of twelfth-graders performed at or above the *Basic* level in 2005 and 23% of twelfth-graders performed at or above the *Proficient* level in 2005 in the area of mathematics. Two percent of twelfth-graders perform at an *Advanced* level and the remaining 39% do not meet *Basic* proficiency. The Spellings Commission report (2006) recommends the NAEP twelfth-grade test be redesigned to also measure college and workforce readiness.

The Mathematics Success Project is a recent initiative, co-sponsored by the National Association of System Heads and the Education Trust, whose primary goal is to increase student success in entry level/high volume undergraduate mathematics courses. The Project is gathering and analyzing data in order to determine which interventions might be used to improve student success, retention, and degree completion. Once sufficient data has been collected, relationship between performance in selected mathematics courses and retention will be examined and as well as, the key characteristics of students enrolled in entry or remedial courses through the first semester of Calculus.

An initial investigation into these relationships was reported in a February 2006 paper on mathematics performance in selected courses offered at the City University of New York (Akst et al, 2006). The report acknowledges these courses: Elementary Algebra, College Algebra, Pre-Calculus, and Calculus, are gateways to many career paths, including the sciences, technology, and engineering. During the Fall 2004 term, data collected across campuses of the City University of New York (CUNY) reveal approximately one third of total university enrollment in mathematics gateway courses was at the elementary algebra course level. NAEP results are reflected here as this number corresponds quite closely to the 39% of twelfth-grades who failed to meet a level of Basic mathematics proficiency. Collectively, these reports (NAEP, 2005, Akst et al. 2006, NSSE, 2008, YFCY, 2007) illustrate the current level of mathematical preparedness of typical college freshmen and their study expectations. These reports indicate that many college freshmen are both underprepared for rigorous courses of study and spend a minimal amount of time engaging in academic activities. Moreover, the findings present a reason to engage students to inform and direct their learning behavior if they are to be successful in gateway mathematics courses and become successful learners in general.

Curriculum-Based Measurement

Since the late 1970's, curriculum-based measurement procedures have furnished teachers with the necessary information to measure student performance growth in basic academic skills in the areas of reading, writing, and mathematics for both special and general education students. "Curriculum-Based Measurement (CBM) is a simple method of repeated measurement toward long-range instructional goals" (Lembke & Espin, 2005, p. 145). CBM is a progress monitoring system that uses a standardized process for obtaining data to monitor student progress in basic skills and to evaluate and improve instructional programs. Helwig, Anderson, and Tindal (2002) describe the characteristics of a CBM probe as follows: tasks are general outcome measures that represent the global content domain, the tasks are technically adequate, relative short in duration, easy to administer and score, and the measures are sensitive to change and conducive to the creation of alternate forms.

Academic progress in basic skills is measured through indicators of 'fluency' or the speed with which a student is able to produce correct answers on an academic task and 'proficiency' in academic skill. These probes are considered to be a static score in that they provide a snapshot of progress at a particular instant. Academic behaviors of fluency and proficiency are displayed as a graphical performance indicator of the learner's progress toward a performance goal over time. This progress graph serves as a tool for teachers to determine the impact of instruction, intervention, and program on student achievement. In this sense, CBM is a formative assessment of progress monitoring both at one point in time and over time. It is the slope of the trend of scores from the repeated administration of CBM tasks that provides information to the practitioner to determine if the student is making progress toward an instructional goal or if an instructional adjustment is needed. L. S. Fuchs (2004) concluded that tasks and depiction of rate of learning together are evidence of progress monitoring. The combination of monitoring learning over time and informing instruction is the unique characteristic of curriculum-based measurement. It is the instructional utility or the use of CBM information by students to make their instructional decisions to improve achievement that is the unique aspect of this study.

The ability of curriculum-based measurement to monitor progress and inform instruction-based decisions has been researched in domains other than basic skills in reading, writing and arithmetic. In order to provide the practitioner with additional information on which to base instructional decisions, the inclusion of a computer software-generated skills analysis in conjunction with the graphed performance indicator was researched by Fuchs and Fuchs (1990) and found to have a positive effect on student achievement in the area of arithmetic skills. CBM procedures have monitored growth in learning in the areas of social studies (Espin, Shin, & Busch, 2005) and critical thinking skills in science (Tindal & Nolet, 1995). Oral reading fluency procedures have been adapted to measure reading growth among English language learners (de Ramirez & Shapiro, 2006). CBM systems have been shown to have predictive ability for middle school mathematics students on statewide tests (Helwig, Anderson, & Tindal, 2002). As CBM procedures have been applied to upper level courses, the probes have evolved to address conceptual understanding and to award partial credit to measure partial growth in learning (Tindal & Nolet, 1995, Foegen et al, 2006). This investigation intends to extend the use of curriculum-based measurement procedures to a general population of undergraduate first semester calculus students and examine its instructional utility to inform students of their progress so they may direct their learning behavior.

Summary and Research Questions

There is strong evidence from both the National Survey of Student Engagement and Your First College Year survey that college students appear to spend a minimal amount of time engaged in academic activities. In the domain of mathematics proficiency, the majority of twelfth graders demonstrate a “partial mastery of the knowledge and skills that are fundamental for proficient work at a given grade” (Kuenzi et al. 2006, p. 3). These traits describe a freshman college student who may not have the tools to succeed in college or be a successful learner.

At the undergraduate level, student achievement is based on summative assessments such as a mid-term and final exam or a term paper. Criterion-referenced measurements are a testing approach in which an individual’s score on a test is interpreted by comparing it to a prescribed standard of performance (Gall & Gall, p. 622). Students have few opportunities to relearn material not mastered, especially when they are uninformed of their progress in specific content areas. The advantage of curriculum-based measurement over criterion-referenced measurement lies in its sensitivity to denote small changes in learning over time and in its ability to inform instruction regarding mastery of the curriculum. Curriculum-based measurement procedures provide immediate feedback in terms of progress and decompose performance into known and unknown elements of the curriculum. By presenting progress as a visual representation of performance supplemented with an analysis of mastery level per topic, students realize academic gains as these indicators are dynamic and respond to changes in performance. The nature of the content of a calculus course offers students ongoing opportunities to develop an in-depth understanding of topics as the academic term progresses. This study provides students with objective feedback to evaluate their understanding of topics through self-monitoring of study habits and metacognitive judgments of test score prediction. The dynamic nature of the feedback alerts students to their progress by graphically depicting changes in weekly performance score or by supplementing the graph with an ongoing assessment of level of understanding per topic.

In her report on curriculum-based measurement research, L. S. Fuchs (2004) described the body of research on curriculum-based measurement in terms of three stages. Stage 1 research focuses on CBM as a measure of performance at one point in time, or a static indicator of progress. Stage 2 examines whether the slope of trend scores is associated with performance. Stage 2 seeks to determine if a change in slope corresponds to an increase in actual performance. The final stage explores the instructional utility of the measures as a tool to make instructional decisions and improve student achievement. This study has the potential to extend the research on the instructional utility of CBM indicators, that is, whether the performance information can be used by the student to direct his or her learning decisions and improve achievement.

Employing a CBM format of weekly probes drawn from a first-term engineering calculus curriculum that are uniformly administered and scored, performance feedback is returned to the student as an analysis of mastery level per topic with a progress graph or solely as a progress graph. Weekly formative assessments give students an opportunity to relearn topics not mastered as these topics may appear on a subsequent probe. The weekly quiz is transformed from a static score to a dynamic indicator of performance over time. CBM practices enable the student to measure progress toward a summative assessment, such as an hourly exam or final exam, as the analysis of mastery level per topic provides a continuum of progress toward the course's learning outcomes. This study extends the research on the instructional utility of progress monitoring as it used course outcome measures, was conducted in an authentic environment, and spanned approximately eleven weeks.

CBM incorporates several of the learner-centered psychological principles - metacognitive factors, motivational factors, and the use of assessments by the student, that are advocated by the APA. For the student who has yet to achieve a level of 'successful learner,' CBM provides performance feedback, occasions to monitor progress, and opportunities to self-assess one's knowledge and make predictions regarding future performance. This study potentially extends the research on monitoring accuracy of Hacker et al. (2000) and Nietfeld et al. (2006) to a new content domain. It examines indices of calibration between predicted performance and actual performance on non-multiple choice tests. Thus, this study examines the extent to which students use individualized performance feedback to inform and direct their learning behaviors. The following research questions are posed:

- Do students who receive individualized performance feedback on weekly curriculum-based measurement assessments exhibit a greater metacognitive awareness of their performance on calculus hourly exams when contrasted to students who do not receive individualized performance feedback as measured by monitoring accuracy indices?
- Do students who receive individualized performance feedback on weekly curriculum-based measurement assessments relearn topics they have not mastered at a higher rate when contrasted to students who do not receive individualized performance feedback?
- Does engagement time in academic activities, such as reviewing notes, completing homework, and studying differ by group for type of feedback received: progress graph supplemented with as analysis of mastery level per topic versus progress graph only?
- Are there differences between feedback groups and reported perceived confidence in one's ability to use feedback information effectively? Are there differences in the relationships between perceptions of usefulness of the feedback information and actual learning behaviors based on the type of individualized performance feedback the student receives?

CHAPTER 2

Literature Review

This chapter describes literature relevant to curriculum-based measurement practices and literature that supports the APA’s learner-centered principles. The chapter is organized into five sections: (1) an overview of curriculum-based measurement (CBM) practices, (2) current trends in curriculum-based measurement research, (3) curriculum-based measurement research in mathematics at the middle and secondary school levels, (4) psychological dimensions of a successful learner, and (5) a summary of the research for this study and the hypotheses this study intends to address.

An Overview of Curriculum-Based Measurement

Curriculum-Based Measurement (CBM) is an established progress monitoring system developed at the University of Minnesota’s Institute for Research on Learning Disabilities (IRLD) from 1977 to 1983 (Deno, 1992). CBM is designed to monitor learning-disabled students’ academic progress through direct assessment of basic academic skills. Researchers at the Institute developed standard simple, short duration fluency measures of basic skills in reading, writing, and mathematics and these curriculum-based measurements serve "as dynamic indicators of basic skills to facilitate timely, formative evaluation for the purpose of improving achievement outcomes" (Shinn & Bamonto, 1998, p. 5). Academic progress in basic skills is measured through indicators of ‘fluency’ or the speed with which a student is able to produce correct answers on an academic task and ‘proficiency’ in academic skills. The dynamic nature of CBM probes lies in its sensitivity to detect differences among individuals with a skill from those without the same skill and differences within individuals over time.

Curriculum-based measurement is a system of assessment that meets the following criteria: the tasks are general outcome measures that represent the global content domain, the tasks are technically adequate, that is reliable and valid, relatively short in duration, easy to administer and score, and the measures are sensitive to change as well as conducive to the creation of alternate forms (Helwig et al., 2002). Curriculum-based measurements convert academic behaviors, such as oral reading, into numbers, as words read per minute. Outcomes are displayed as a visual representation in the form of a graphed performance indicator of the learner's progress toward a performance goal (Wright). The graphed performance indicator displays proficiency scores over time, thus providing an instrument for the teacher or practitioner to examine changes in individual student achievement outcomes based on a particular method of instruction over a specified period of time. Curriculum-based measurement procedures specify how to select materials for inclusion on the curriculum-based probe and how to sample student performance using those stimulus materials in consistent and systematic ways to produce reliable and valid information for instructional decision-making (Deno, 1992).

Frequent monitoring and graphical depiction of student scores are the keystones of CBM. Decisions regarding student progress are made by visually examining the slope of the trendline of a series of scores, usually four to eight data points, against a projected performance goal. In their review of the research on using CBM to improve learner achievement, Skecker, Fuchs, and Fuchs (2005) describe the decision rules succinctly as: to change instruction when the slope of the progress line is less steep than the slope of the goal line and to raise performance goals when the slope of the progress line is steeper than the slope of the goal line.

The original purpose for adopting a curriculum-based measurement progress monitoring approach was for the teacher to monitor and improve student performance by systematically testing alternative approaches to instruction for individual students (Deno, 1992). In the primary grades, and in particular with a special education population, instructional approach decisions to assist the learner in becoming proficient in basic skills are made for the learner based on the graphed image of performance. By citing a meta-analysis of 21 studies by Fuchs and Fuchs (1986) and a summary of nine studies by Shinn and Hubbard (1992) in their research review, Good and Jefferson (1998) present strong support for the effects of progress monitoring using curriculum-based measurement procedures on higher student academic performance when compared to the academic performance of students whose progress is monitored by traditional criterion-referenced testing. Good and Jefferson (1998) document the measurement validity of CBM as an instrument to examine student progress and the extent to which these “assessment results provide information that leads directly to the development of effective intervention plans” (pg. 70). There exists extensive research supporting the utility or the direct contribution of the CBM practices to improved student performance (Good & Jefferson, 1998).

In their review of the effectiveness of CBM as an assessment tool to enhance student achievement, Stecker, Fuchs, and Fuchs (2005) note that progress monitoring alone may not be sufficient to address specific student deficiencies. The effects of additional sources of information on student performance, such as an analysis of skills in terms of level of mastery (Fuchs and Fuchs,1990) and more specific decision rules (Stecker and Fuchs, 2000), are included in this review.

To examine the impact of including a skills analysis of student mastery of topics in guiding decision rules to modify instruction, Fuchs and Fuchs (1990) randomly assigned 30 special education teachers, from 16 schools in a southeastern metropolitan area of the U.S. to one of three groups: graphed performance indicator and skills analysis, graphed performance indicator only, and a control group, which received neither feedback measure. Performance indicator analyses focus on overall rates of growth in the curriculum and skills analyses focus on mastered and nonmastered curriculum skills (Fuchs & Fuchs, 1990, pg. 6). From grades 3 to 9, 100 students, designated as learning disabled or emotionally disturbed according to state regulations, were selected by participating teachers. The distribution of the 100 students by group was 40 : 40 : 20 and no significant differences between groups in terms of age, grade placement, level of mathematics performance, and number of years in special education were discerned. Curriculum-based measurement probes in the domain of mathematics were administered on a computer and scored by a data-management software program, BASIC MATH, for both CBM groups. The control group completed teacher-made tests that serve as a benchmark of standard practice for monitoring student progress within the special education population.

After eight weeks of training on CBM procedures, teachers were instructed to record the number of goal changes made for each student, the number of instructional changes made for each student, and the number of specific math skills referenced by the teacher as part of the information on instructional change. A level of goal ambitiousness was computed as the difference between the final goal level and the baseline goal level. There were no group differences as determined by tests of analysis of variance among teacher characteristics, no differences in treatment fidelity, number of goal changes, level of goal aspiration, or number of instructional changes among group. Significant differences were found in the number of math skills referenced by the teacher in the information on instructional changes, with more math skills noted for the graphed performance indicator with skills analysis group than both the graphed performance group and the control group. In terms of achievement on the Math Computation Test-Revised (MCT-R) in the area of number of digits correct, students in the performance graph with skills analysis feedback group exceeded students in the performance graph indicator feedback group and the control group . The MCT-R is a criterion-referenced measure that samples math problems across grades 1 through 6 from the computational objectives of the Tennessee Basic Skills First Curriculum. The curriculum comprises a statewide set of competencies that are expected for grade promotion (Fuchs & Fuchs, 1990, p. 11).

These results are notable for several reasons. In terms of methodology, support is given to allow for partial credit, by tallying the number of correct digits in an answer, as a measure of partial progress in mathematics learning as opposed to counting the answer as correct or incorrect. The concept of awarding partial credit is developed further in the research of Foegen et al. (Project AAIMS, 2006). Second, the inclusion of a skills analysis resulted in teachers planning more specific instructional adjustments for students, when compared to the instructional adjustments made by teachers for students whose performance was displayed graphically or for students in the control group. Skill analysis was a new attempt to strengthen the effectiveness of CBM as a tool for instructional adjustments in response to student performance. Concurrent studies by

Fuchs and colleagues in spelling and reading have included this qualitative skill information. They found more specific instructional adjustments can effect superior achievement using this supplementary information on skill mastery. The final notable result is the lack of discernible differences between the control group and the graphed performance indicator group. This is attributed the paucity of specific information generated by the score and accompanying graph that the teachers of the graphed performance indicator group received. These teachers did not have access to the responses of their students and therefore lacked information gleaned from those responses.

One of the tenets of CBM is the use of frequent short probes that are administered and scored uniformly. Decision rules for instructional adjustments are made by comparing progress, in terms of the slope of a trendline of increased fluency over time, to a predetermined goal. When adhered to, this tenet leads to reliable scores and valid use of those results. L. S. Fuchs (1998) notes that despite the usefulness of CBM, teachers frequently do not use these procedures due to the significant amount of time required to collect and manage the ongoing assessment information, as well as, difficulty translating the information into instructional changes. To overcome the time imposition to test, score, graph, and apply decision rules, these practices are currently performed on a computer, with data-management software. Previous research by Fuchs and Fuchs (1990) established that students completed CBM computer probes with a high degree of accuracy.

Using computer-based testing and performance indicators, Stecker and L.S. Fuchs (2000) provide further support for the effect of both quantitative and qualitative performance information on student achievement. Their study experimentally contrasted the effects of frequent testing and the effects of frequent instructional adjustments made with and without CBM on student learning in the domain of mathematics, grades 2 to 8. From 12 schools in a southeastern U.S. metropolitan area, twenty-two special education teachers who taught in either resource or self-contained settings participated in the study. Each teacher was asked to identify two CBM candidates and a matched partner who performed at the same mathematics level for each CBM designated student. In total 42 matched pairs completed the 20 week study.

Three study groups were formed. The first group consisted of CBM candidates. Students took CBM tests on a computer, after receiving training on how to do so. Tests were administered twice weekly. Software scored answers in terms of number digits and answers correct and returned that information and a graph of the number of digits correct over time to the student. The software instructed teachers to make instructional adjustments every 3-4 weeks. The program also provided the teacher with a skills analysis twice monthly that summarized the student's level of mastery by problem type (such as, subtraction with regrouping or multiplication of two 2-digit numbers less than 50). On average teachers implemented suggested instructional recommendations 80.9% of the time.

Matched partners were split into two groups. Matched partners in Group 2 took CBM tests at the computer and saw their own performance graph following a test and matched partners who did not take a CBM test were assigned to Group 3. Teachers did not have access to the second group's data. This assignment to group matches the three study groups, graphed performance indicator and skills analysis, graphed performance indicator, and control, from the previous research by Fuchs and Fuchs (1990). Although teachers did not monitor Group 2 students, the software did keep a log of the number of occasions the student worked on the computer and this log revealed Group 2 students had on average 29.8 measurement points compared to 32 measurement points for Group 1 students.

After 20-weeks of data collection, Stecker and L. S. Fuchs (2000) examined whether the achievement effects as measured on the Mathematics Operations Test-Revised (MOT-R) would be the same for matched pairs of students if instructional decisions for each pair was based on the CBM assessments and instructional recommendations of the Group 1 participant in each pair. As paired students practiced skills on the computer, the teacher provided specific feedback to one of the students. Specifically, the researchers examined whether teachers can make instructional decisions for a group based on the CBM data of one student who is representative of the group.

Repeated measures analysis of variance with two within factors (student type: CBM and partner and time: pretreatment and posttreatment) were conducted to examine the effect of CBM instructional decisions and instructional adjustments on achievement. Results showed main effects for both factors; but the interaction between student type and time was greater than the main effects with the CBM group performing significantly better at posttest in terms of mean digits correct on the Mathematics Operation Test-Revised than the partner groups.

A second question posed by the researchers was whether frequent CBM practice enhanced student achievement by itself. One-way analysis of variance on the effects of frequent measurement on achievement revealed a main effect for time, not frequency of measurement, on achievement in that all three groups performed better on posttest.

These findings suggest that instructional changes by themselves may not result in optimal gains for learners with mild to moderate learning disabilities. From the research there is evidence to support the benefits of instructional adjustments made to fill specific learning gaps, in that those students perform at a higher level when compared to the performance of students who have had changes in instructional tactics based primarily on the student not performing at a specific level.

Current Trends in Curriculum-Based Measurement

Curriculum-based measurement methodology provides reliable but simple data to document student growth and to determine the necessity for modifying instructional programs. Curriculum-based measurements have successfully been adapted to screen and identify students in need of additional support services, such as expected reading fluency for Spanish-speaking English language learners (de Ramirez & Shapiro, 2006), and to measure content-domain knowledge in the areas of social studies (Espin, Shin, & Busch, 2005) and undergraduate psychology (Larson & Ward, 2006). CBM has been used to measure the proficiency of written expression at the high school level (Espin, Scierka, & Skare, 1999), demonstrate growth of critical thinking skills for middle school science students (Tindal & Nolet, 1995), and to predict performance of middle school students on statewide mathematics tests (Helwig, Anderson, & Tindal, 2002).

The predictive utility of curriculum-based measurement as an indicator of growth in oral reading fluency for English language learners was studied by de Ramirez and Shapiro (2006). The purpose of the study was to investigate whether established reading curriculum-based measurement (CBM-R) probes could distinguish between Spanish speaking students whose learning difficulties stem from learning another language and Spanish speaking students who require special educational services. The study also examined the oral reading fluency growth rates among Spanish-speaking students in English and in their own language.

Students from grades 1 through 5 were selected for participation through a stratified method that reflected the population of Spanish-speaking English language learners enrolled in bilingual education and monolingual English speaking students from the general school population of a southwestern state. Each group was additionally stratified by reading ability, as determined by state criterion reading measures. Reading fluency probes were administered once in each of the fall, winter, and spring terms during the school year. Rate of reading growth per group within grade was examined. The probes on oral reading fluency were summarized by program: general education, bilingual reading fluency in English, and bilingual reading fluency in Spanish. Mean scores and standard errors of average number of words gained weekly were analyzed by group. Results of repeated measures analysis of variance demonstrated that Spanish-speaking English language learners were making yearly progress in oral reading fluency, although at a slower rate than monolingual students, especially in grades 1 through 4. More importantly, the study established that CBM can be sensitive to measuring the development of language and literacy skills among Spanish-speaking English language learners.

Recent studies have examined growth in subject area learning using CBM procedures. Vocabulary-mapping curriculum based measures have been shown to be an indicator of growth in social studies learning (Espin et al., 2005) and have been studied as a predictor of course grade in psychology (Larson & Ward, 2006). Each of the aforementioned studies has adapted the criteria of a CBM assessment system designed to measure change in student performance, by generating repeated measures within a short period of time, and its documented validity and reliability (Espin et al. 2005) to a novel use.

In the area of social studies, Espin et al. (2005) investigated the validity of vocabulary matching as an indicator of learning within a seventh-grade social studies context. Vocabulary-matching probe items were classified as applied or factual and the ratio of each item type on every probe was 3:1. Applied questions are those in which the student applies social studies concepts and principles to an event or phenomena. An example of an applied question is provided by Espin et al. (p. 356).

> José comes from a working class home. He married Judy who is very wealthy and moves into an upper class neighborhood. José's change in status is an example of
>
> a. mobility
> b. sanctions
> c. mores
> d. primary group

Whereas, factual questions are those in which the student is asked to make a one-to-one relation between concepts and events, such as

> The process by which a member learns the rules of his or her group is called

a. socialization

b. community

c. role play

d. mobility

A novel feature of this study is the use of hierarchical linear models (HLM) to address sensitivity to student improvement over time, sensitivity to interindividual differences in growth rates, and the validity of the growth rates with respect to criterion measures (Espin et al., 2005, pg. 357). Twenty-two probes were administered in a regular social studies classroom setting; the administration alternated weekly between student-read and teacher-read questions. The results support the technical adequacy of a student-read vocabulary probe as an indicator of student learning in social studies. Performance differentiated between students with a leaning disability and those without a learning disability. This study supports the characteristics of CBM as a measure of change that is sensitive to improvement over time and individual differences in the area of social studies.

Larson and Ward (2006) employed vocabulary-matching curriculum based measures to examine whether reading skill correlated with the direction of the trendline of the participant's weekly vocabulary monitoring and whether the direction of the trendline correlated with the final course grade earned. Weekly vocabulary-matching probes, consisting of terms randomly drawn from the glossary of the textbook used in an introductory psychology course, were administered over a period of nine weeks. Sixty-nine students, enrolled in two sections of an introductory undergraduate psychology course at a mid-Atlantic region college, matched glossary terms to their definitions. The probes were scored as number of correct matches in a 5-minute period. The unique feature of this descriptive study is that student pairs graded each other's weekly vocabulary-matching probes and graphed the results. Students constructed line graphs connecting consecutive weekly scores. Students were instructed to observe the overall trend of these segments as increasing or decreasing as a basis for determining progress in learning course material. This format represented a diversion from standard CBM practices in that results of weekly probes were returned directly to the learner and the learner was asked to base an estimate of his or her course progress on the trend of the weekly probes This study extended the vocabulary-matching research of Espin et al. (2005) to an undergraduate group of students enrolled in an introductory psychology class.

Data from this descriptive study was further analyzed using a trendline of progress across all nine measures that was instructor generated. Results did not indicate any relationship between reading score and trendline direction. Results correlating course grade with direction of trendline were not significant in part due to the skewed distribution of grades, where 94% of the grades were either A or B, of which 68% had an ascending trendline and 22% had a descending trendline. Larson and Ward (2006) call this study an initial investigation into the use of CBM procedures as a predictor of course grade for undergraduate students. The researchers suggest these procedures can alert students and their instructors to occasions when the student is not learning the content-area information at a rate that will lead to passing and through progress monitoring, content-area CBM procedures may raise retention by their predictive capability. This is an area for future research.

As students enter secondary school, course content moves from basic skills to the application of those skills and curriculum-based measurements of fluency may not capture underlying growth in knowledge. Espin, Scierka, and Skare (1999) examined the criterion-validity of previously studied indicators of written expression as predictors of general writing ability in grade 10. The study additionally explored which measures were the most useful in predicting general writing ability. Participants were drawn from a senior high school in a suburban mid-western city. Using a stratified random sampling procedure, participants were selected from three English 10 class levels: basic (48 participants), regular (50 participants), and enriched (49 participants). Short writing probes (3 minutes in duration) were administered during the students' English classes. Samples were typed, exactly as written, into a computer by the researcher and two assistants. Using a grammar check feature of a word processing program, samples were scored, for number of words written, number of characters written, number of characters per word, and number of sentences written. Additionally, hand scoring of the number of words spelled correctly, number of correct word sequences, and mean length of correct word sequence strings were computed from a hard copy of the probe.

Correlation coefficients between each of the probe-generated variables and each of the following dependent variables: the Language subtest of the California Achievement Test (CAT), first and second semester English grades, and independent rating of writing samples, revealed the number of sentences written and characters per word were most strongly correlated with writing performance across all groups. Correct word sequences and mean length of correct word sequences were also significantly correlated with each dependent variable. Group differences were significant for characters per word, number of sentences written, and mean length of correct word sequence strings. For all independent variables, there was a consistent pattern for mean group scores, with enriched groups the highest, followed sequentially by regular, basic, and learning disabled groups. Finally, using step-wise regression on the dependent variable Language Total score on the CAT, it was determined that the combination of independent variables: characters per word, number of sentences written, and mean length of correct word sequences accounted for 38% of the variance in the Language Total score.

The contribution of this research to the expanding application of curriculum-based measurement is its recognition that the simple probes of basic skills in reading, writing, and arithmetic do not adequately measure learning or skill ability in middle or high school content areas. The purpose of the Espin et al. (1999) study was not to monitor progress but to establish criterion-related validity of specific curriculum-based measures in writing at the high school level. The strength of this study is its adaptability to small group testing, especially when students write directly into the computer. Computer scoring would greatly reduce administration and scoring time for the teacher.

Time spent administering, scoring, and developing indicators of progress has always been a consideration when developing curriculum-based measures for content-area domains at the upper school level. As noted in the preceding research by Espin et al. (1999), single skill measures may not adequately correlate with criterion measures of performance. In a study by Tindal and Nolet (1995), curriculum-based measurement procedures were adapted to monitor progress in the critical thinking skills of middle and high school science students with mild learning disabilities. In their study, Tindal and

Nolet (1995) define the essential features of progress monitoring as curriculum-based measures that "focus on the response as the critical feature for generating a measurement system that has the potential of being repeated over time, addresses instructional effectiveness, and allows both quantitative and qualitative outcomes to be summarized in determining what to teach and when to modify instruction" (pg. 3). The work describes the development of curriculum-based measures that can be used to monitor changes in critical thinking skills in terms of quantitative and qualitative scoring and whose results are presented as a case study. Samples were generated by an eighth-grade student designated as learning disabled who was instructed in a general science class and received additional instruction support as part of a pull-away program.

Outcomes were measured quantitatively and qualitatively. Quantitative measures were counts of explicit instances of references to concepts within a new context or as an additional attribute or example within a question response. Quantitative measures counted knowledge forms: fact, concept, or principle, in the response. Qualitative scoring was achieved through a series of flowcharts that count intellectual operations. Intellectual operations were defined as identifiable behaviors, such as prediction, evaluation, and explanation that are connected and correct with respect to the response to the question. The intellectual operations interact with the knowledge forms in the question format. For example, in a compare and contrast argument, students engage in explanations using various concepts and principles to support the argument. Twenty-five probes were developed to cover six units of study in the earth science curriculum. Data was analyzed in terms of a split-middle calculation of slope. The researchers attempted to develop measures of critical thinking that would be able to determine whether student performance in science was improving over time, whether the student's critical thinking skills were improving over time, and whether the measurement system would be sensitive to low-performing students. The findings however demonstrated a lack of sensitivity in the measures to denote significant improvement. Slope was examined for each new unit of study and over the entire curriculum. The quantitative measures demonstrated little change both within units and across the year. Qualitative measures showed some growth within units; overall there was a slight positive change in quality of response throughout the year.

The importance of this body of work lies in the development of a curriculum-based measurement system that examines progress in critical thinking skills in terms other than fluency and accuracy. Secondary and undergraduate level content area domains involve the use of skills and information, not just a set of information facts. The measurement system for critical thinking skills is based on a response scored in terms of a set of criteria that address the intellectual operations employed to respond to a task.

Complementing the evolving applications of CBM, measurement techniques have been adapted to capture the key feature of CBM as a metric for small changes in performance. From a visual inspection of trendlines to the inclusion a qualitative skill analysis (Fuchs & Fuchs, 1990, Stecker & Fuchs, 2000) and the utilization of repeated measures of analysis of variance (Stecker & Fuchs, 2000) and hierarchical linear models (Espin et al, 2005), CBM probes have been able to demonstrate change in performance in a variety of domains.

Curriculum-Based Measurement in Mathematics at the Middle and Secondary School Levels

The utility of CBM in the area of mathematics has focused almost exclusively on basic skills. There is an acknowledged lack of research on progress monitoring measures in mathematics at the secondary level. Espin and Tindal (1998) note that the focus of instruction at the secondary school level changes from the acquisition of basic skills to the implementation of those skills to acquire content area knowledge. In their discussion on curriculum for students with mild disabilities, the emphasis placed on assessment of content knowledge is hindered by the low level of proficiency in basic skills. Indicators, such as the number of digits correct in 3 minutes have limited application in mathematics problems on geometry.

Foegen, Jiban, and Deno (2007) present, in their review of the literature on progress monitoring, a discussion on the development of CBM tasks in the domain of mathematics. School accountability mandates, such as No Child Left Behind, are not only to educate all students; but also to measure progress for both learning and non-learning disabled students. CBM has proved itself a reliable and valid indicator of progress monitoring, in particular with a learning disabled population. Foegen et al. (2007) describe two approaches to measuring mathematics learning within a CBM framework. Common practice is to create a pool of test items that are drawn from the year's mathematics curriculum. The advantage of progress monitoring using curriculum sampling is the direct link to curriculum, which facilitates the ability of the teacher to determine if the student's rate of progress toward an educational goal is progressing as expected or if an instructional change must be devised. The limitation of this method of progress monitoring is that the curriculum corresponds to a single year and growth in mathematics skill cannot be interpreted across years.

The second approach to measure mathematics learning is to monitor progress over time (Foegen et al., 2007). Robust indicators are measures that are designed to provide information about student progress on key outcomes over several years. Paralleling oral reading tasks is a robust form of mathematics curriculum-based measurement that assesses growth in core competencies, such as numeracy or problem solving, not for a single academic year but over several years. For example, early numeracy growth, similar to early literacy, can be measured at different levels of developments, grades 1 – 3, grades 4 – 6, etc. Problem solving growth, as with growth in reading skill, can be measured through a body of problems that students should be able to solve as students move through the curriculum at the primary, intermediate, middle, and high school levels.

Concepts or robust indictors of mathematics proficiency were studied for their viability as predictors of achievement on statewide math tests by Helwig, Anderson, and Tindal (2002). Their study examined whether CBM items that measure the quality of conceptual understanding, rather than a production rate at which declarative or procedural knowledge is retrieved, predict success on statewide mathematics achievement tests and whether these items differentiated in their predictive ability for both general and special education populations. In terms of mathematics, Helwig et al (2002) differentiate among production or computational tasks, process or procedural tasks, and conceptual tasks that address the underlying understanding of a property or concept. Conceptual tasks synthesize knowledge. An example of a conceptual task is “Write a number between 6.4 and 6.5”. The authors posit that when procedural tasks are embedded in word problems or presented in a novel format, the likelihood increases that the correct response reflects conceptual understanding rather than rote memorization.

Using CBM procedures, 11 concept test items were developed and embedded within a 48-item mathematics CBM assessment that also included problem solving and computational tasks. The CBM scores were correlated with a computer-adapted test of mathematics achievement, which was developed to be comparable to the statewide achievement test given in grade 8. A total of 199 eighth-grade students participated in this study representing four school districts from a western state. Results of a discriminant function analysis indicated the concept-grounded math CBM, together with educational classification, were useful in predicting performance on a statewide mathematics achievement test. The math CBM together with students' educational classification, general or special education, also predicted whether students would meet state benchmarks.

Recognizing that proficiency in algebra skills is linked to improved outcome in employment and postsecondary education, Project AAIMS (Algebra Assessment and Instruction – Meeting Standards, Foegen, 2003) engaged in a three-year process to identify algebra curriculum, instruction, and assessment for students in both general and special education; to develop and validate a set of measures to assess performance and progress in algebra; and to evaluate whether these measures can be used to support instructional decisions. Working in collaboration with 15 to 20 general and special teachers from three partner school districts in a mid-western state, this research project has developed 3 technically adequate measures in the content area of algebra. The measures are sensitive to student growth and have some predictive correlation with state normed tests in basic skills, computation, and algebraic aptitude (Perkmen, Foegen, & Olson, 2006d). Participating students were drawn from pre-algebra or Algebra 1A, Algebra 1, and 8th grade Algebra. The Project determined the measure with the most sensitivity to growth is the curriculum sampling measurement, *Content Analysis-Multiple Choice Measure.* This measure draws items from eight chapters of the textbook used by all participating districts. This measure awards from 3 to 0 points for both answer and supporting work and subtracts a point for a 'guess' or incorrect answer without any supporting work.

Two robust indicator measures, *Basic Skills* and *Algebra Foundations*, were also developed and found to be both reliable and valid. It was posited that a measure of basic algebraic skills would act as an indicator of overall proficiency in algebra much the same way as basic arithmetic skills operate as a general indicator of proficiency in math skill. Basic algebra skills are skills that should be completed with a high degree of automaticity. The *Algebra Foundations* measure is designed to reflect core algebraic concepts and skills, such as writing variables and expressions, manipulating algebraic expressions, graphing linear equations, solving one-step equations, and identifying and extending patterns and functions. Both the *Basic Skills* and *Algebra Foundations* measures are scored for fluency and proficiency. Both were demonstrated to be sensitive to growth and the validity of the measures appears to vary for different levels of algebra, algebra background, and for ability levels (Perkmen et al., 2006b, 2006c, 2006d).

A Project AAIMS outcome, from the collaboration between the research team and participating teachers, reiterates the need for timeliness in receiving student results if these measures are to be used as a basis for instructional decisions. It is the sense of real-time progress that makes the scores meaningful. It was also noted that students are motivated to do their best work when importance is attached to the task and some form of extra credit points were attached to the completion of the probes during the development of the measures (Foegen et al., in press).

The research into the development of progress monitoring measures for algebra by Project AAIMS (Foegen, 2003) lends further support for the use of multiple measures (Espin et al, 1999) to monitor progress in content area domains. It also supports the work of Espin et al. (2005) by differentiating between different forms of knowledge that are assessed at the secondary level. The awarding of partial credit followed a scoring rubric template as did the awarding of points in the critical thinking skills measures used by Tindal and Nolet (1995).

The recommendations of Foegen, Jiban, and Deno (2007), the findings of Project AAMES (Foegen, 2003), and of Helwig, Anderson, and Tindal (2002) collectively indicate that for middle to upper level mathematics it is a combination of both curriculum sampling and concept-grounded measures that are needed to assess current student learning, predict future performance, and support efforts to design effective curriculum.

Psychological Dimensions of a Successful Learner

A successful learner is one who is "active, goal-directed, self-regulating, and assumes personal responsibility for contributing to one's own learning" (APA, 1997, p. 3). A dimension of the psychological factors that affect the learner and learning involves metacognitive and cognitive abilities. Batha and Carroll (2007) describe a metacognitively aware individual as an efficient decision maker who questions, monitors, and instructs oneself to gain access to relevant information, formulates a plan of action, guides the execution of that plan, and regulates the use of cognitive strategies throughout the decision-making process (p. 64). This research focuses on the regulation of cognitive components of metacognition, as described by Batha and Carroll (2007) and is interpreted as taking control of actions, implementing strategies, and acting on feedback.

Empirical research has provided many illustrations of the benefits of accurate metacognitive judgments. Recent research into measuring metacognitive judgments has moved the discussion from the laboratory to actual classroom settings and from a single instance experiment to an experiment over a sustained period of time (Hacker et al. 2000, Nietfeld, et al. 2005, Nietfel et al, 2006). Monitoring refers to one's on-line awareness of comprehension and task performance (Schraw, 1998) and is viewed as the data driven dimension of metacognition that serves to inform control processes.

Hacker, Bol, Horgan, and Rakow (2000) examined differences in predictive and postdictive monitoring judgments and their relationship to performance among college students (n = 99) enrolled in two sections of an introductory educational psychology course. Three tests were administered as part of regular course practices. The tests were multiple-choice in format and of mixed item difficulty level. The proportion of item difficulty was consistent across all three exams. Correlations between performance and absolute differences in accuracy scores were noted. Prediction accuracy seems to increase over time and the magnitude of the increase is correlated with student performance ability. Associations between study time and predicted confidence in one's knowledge were supported for Exam 2 and between study time and postdicted confidence in one's knowledge with Exam 3.

Nietfeld, Cao, and Osborne (2005) investigated how monitoring accuracy affected performance in the classroom over a one-semester course in educational psychology. Monitoring accuracy or confidence judgments relating to performance in testing was assessed at the local level or by individual test item, and at the global level using overall test performance. Using an intact section of an educational psychology course, these researchers investigated the relationship between monitoring judgments on four exams, how those judgments change over time and affected test performance, and how judgments are affected by item difficulty. Key findings suggest there is an effect of time on both local and global monitoring judgments, with global monitoring judgments more accurate over time than local monitoring judgments. A second finding demonstrated that monitoring accuracy is consistently related to performance. The last result of the study showed that item difficulty is a significant predictor of local accuracy and bias score.

It was noted that monitoring accuracy is stable over time and self-directed feedback did not improve accuracy. The authors concluded that the act of prompting students to think about their performance is probably too passive an attempt to alter monitoring. If the goal is to increase students' metacognitive skills, then more consistent, intensive, and explicit attempts are needed within a course setting that require practice, feedback, and strategies for monitoring on an ongoing basis (Nietfeld, 2005).

In their follow up study, Nietfeld et al. (2006) investigated the effects of monitoring exercises and feedback, on calibration accuracy and test performance over an entire semester in an educational psychology course. This was a quasi-experimental study, where the treatment group received exercises in monitoring comprehension of course content and feedback on their calibration, while the control group received neither monitoring exercises nor feedback on their calibration. Results of repeated measures analysis of variance showed the treatment group became more proficient (one standard deviation higher) at calibrating accuracy of performance than the control group's measures of calibration. By the second exam, the treatment group showed itself to be higher in learning outcomes as well as calibration accuracy. Additionally as determined by a path analysis, there appeared to be direct effects from change in calibration accuracy to final test score and a significant effect from average calibration to a final measure of self-efficacy.

Calibration seems to play a significant role on test performance, integrative knowledge projects, and self-efficacy. One could argue that as the course progresses self-efficacy is transformed into a belief that is more informed by the monitoring processes and experiences (Nietfeld et al., 2006). A recommendation for future research (Nietfeld et al., 2006) is to examine if the effects of monitoring exercises and calibration feedback on performance, monitoring accuracy, and self-efficacy will be different in courses, such as math or science, where concepts and knowledge build upon each other. This study partially addressed that recommendation.

A second area of suggested future inquiry is to investigate how "self-generated feedback, such as monitoring of comprehension and monitoring of performance relate to students' goal setting and revision, assessment of academic progress, and adjustment of study strategies during the learning process" (Nietfeld, 2006, p 176). This study addresses these learner-centered dimensions; specifically how does feedback direct what topics a student will choose to study in preparation for a weekly CBM quiz.

Self-regulatory processes, such as self-monitoring, comprise another dimension of a successful learner. Zimmerman and Paulsen (1995) operationalize academic self-monitoring in terms of keeping a physical record of one's academic performance. The records provide information regarding the quality and outcomes of performance in academic areas. Self-monitoring can provide a wealth of information to the learner by focusing the student's attention to a limited number of outcomes, differentiating between working and non-working strategies, ascertaining when performance is ineffective, facilitating effective management of study time and environment, and fostering self-reflection on the part of the learner. Self-monitoring can be an agent for personal change.

For change to occur, students need to be motivated by the perceived benefits of self-monitoring in order to invest the time and effort necessary to monitor one's progress. Students can be made aware of changes in performance that result from their responses to self-monitoring information (Zimmerman & Paulsen, 1995). Informal monitoring may not result in substantive change. It becomes the teacher's role to formally instruct learners to self-monitor learning behavior and to be cognizant of the changes in performance brought about by changes in one's learning behavior. Research studies by Lan (1996) and Kitsantas and Zimmerman (2006) demonstrate the benefits of self-monitoring in different contexts.

Over the course of four successive college terms, Lan (1996) examined whether the addition of a self-monitoring protocol on statistics topics and their definitions would enhance performance on course exams, use of strategies, and several attributes. The 72 participating graduate students enrolled in his introductory statistics course were assigned to three groups, self-monitoring, instructor-monitoring, and control. Students in the self-monitoring group were asked to record the number of times they spent on a set of study activities (lecture, text assignments, discussion, and help-seeking), how long they spent engaged in study activities, and their self-efficacy for each statistic topic. The instructor-monitoring group monitored the instructor's time on topic, as an indication of its importance. Students in the self-monitoring group performed better on all course exams when compared to the performance of the control group and the instructor-monitoring group as demonstrated by a mixed-design ANOVA. The results of a MANOVA revealed the self-monitoring group also used the strategies of self-evaluation and environmental structuring more frequently than the instructor-monitoring group and the control group.

Performance monitoring can be recorded graphically and the effects of a visual representation of online performance on dart throwing skill were investigated by Kitsantas and Zimmerman (2006). Volunteers were recruited from the psychology pool at a state university and randomly assigned to four groups that combined two types of performance standards with two forms of performance monitoring during a dart throwing session. Performance standards were categorized as absolute, which measured success as 7 points and failure as 0 points in hitting the bull's-eye, or graduated, which awarded points on a scale of 7 to 1, with score decreasing as the dart hit farther away from the bull's-eye. Performance monitoring was recorded either by graphing ongoing performance or recording one's final score from the dart throwing session. A fifth group practiced without performance standards or score recording. All participants completed a post skill level test. In addition to measures of performance, motivational belief measures, such as self-efficacy of one's capability to throw darts, self-satisfaction with one's overall score, self-evaluation regarding their improving abilities, and attributions regarding performance, were also administered.

Data analyses revealed those who graphed their ongoing performance had higher skill levels than participants who did not graph their ongoing performance. The graphing group also demonstrated higher self-efficacy in their capability to throw darts and perceived higher levels of improvement over non-graphing participants. Across all measures the graduated monitoring group had higher skill levels and motivational beliefs than the participants who adopted absolute performance standards.

The benefits of seeing ongoing progress visually and denoting changes in performance have been demonstrated in academic domains through curriculum-based measurement practices and in motoric skill domains (Kitsantas & Zimmerman, 2006). The conclusions drawn by the researchers shift the focus of utilizing a graph to depict outcomes and comparing those outcomes to graduated standards, from motoric skills to students in general who want greater control over their learning. In the classroom, students at times set unrealistic standards for themselves. Kitsantas and Zimmerman (2006) point to instructors to "identify these students and help them readjust their standards to be sensitive to small improvements in skill (p. 211). It is the dynamic nature of graphed performance indicators that allows students to monitor their outcomes and enhance performance very much in the same way curriculum-based measurement practices operate.

Summary and Hypotheses

Research on the tenability of Curriculum-Based Measurement as a progress monitoring tool can be organized about the three stages of research as described by

L. S. Fuchs (2004). Developing probes to measure learning at a single point in time was evidenced through the Stage-1 research on vocabulary-matching measures in social studies (Espin et al., 2005) and psychology (Larson & Ward, 2006), on critical thinking skills in science (Tindal & Nolet, 1995), and in algebraic performance (Foegen et al., 2006).

The second stage of the research focuses on the ability of the measure to denote change and measure progress. The algebra measures developed in Project AAIMS (Foegen, 2003), the social studies vocabulary measures (Espin et al., 2005), and the oral reading fluency measures for English language learners (de Ramirez & Shapiro, 2006) have all demonstrated the adaptability of CBM to monitor progress in domains other than basic skills and with general education populations. The use of hierarchical linear modeling (Espin et al., 2005) to denote sensitivity in progress over time and to denote interindividual differences is more advantageous than the use of slope analysis to measure change in performance. Slope analyses of performance on the critical thinking measures in science (Tindal & Nolet, 1995) failed to denote student progress in concept identification and intellectual operations on earth science topics; although content knowledge was gained. It may not be a deficiency in the measure to denote progress in critical thinking skills, but rather that the construct of critical thinking does not develop linearly.

The last stage of research explores the instructional utility of the measures, that is, can the CBM information be used for instructional decisions and to improve achievement? The use of skills analyses as a component in decision-making (Fuchs & Fuchs, 1990, Stecker & Fuchs, 2000) has been shown to result in greater student achievement when compared to the achievement of students whose instructional decisions were based on a graphed performance indicator of progress, traditional assessment methods, or based upon the performance of others. The predictive validity of CBM has been demonstrated in the area of reading fluency for English language learners (de Ramirez & Shapiro, 2006) and middle school mathematics (Helwig et al., 2002). Collectively, there has been much research on the development of progress monitoring tasks and the assessment of their reliability and validity in content areas beyond basic skills and with general education populations.

Stecker et al. (2005) outlined the characteristics of curriculum-based measurement as assessing progress toward long-term goals, of which a factor of CBM procedures is to draw from a pool of test items that reflect the general outcomes. As researchers extend curriculum-based measurement practices to upper level content domains, the characteristics of the task have extended to measure higher level content. Single skill probes cannot adequately measure knowledge in social studies or algebra. In studies by Espin et al. (2005) and Foegen (2003), tasks included measures of both factual and conceptual knowledge. Writing ability of high school 10th graders (Espin et al., 1999) was best measured using a combination of single skill measures. Tindal and Nolet (1995) extended the essential features of CBM to measure critical thinking tasks in science as response driven measured in terms of both science principles and intellectual operations used in the task response. This format deviates from the accuracy scores of the response judged to be correct or incorrect in particular because the tasks are in a free response format. General CBM procedures remain the same even when specific content or curriculum varies. This study follows the format of Tindal and Nolet (1995) and focuses on the quality of the response in terms of levels of mastery rather than as a measure of accuracy. The current study draws items for weekly assessments from the curricula of a first semester undergraduate engineering calculus course, and these tasks employ multiple skill measures that test basic level knowledge, an applied or proficient level knowledge, and an advanced or conceptual level of knowledge.

Another important facet of CBM is frequent monitoring and displaying student scores graphically. This information provides a better basis for decision making. The data reflects how a student's achievement is changing over time, because the level of difficulty and time allowed for the probes remains constant. The inclusion of skills analyses to supplement performance information (Fuchs & Fuchs, 1990, Stecker & Fuchs, 2000) has been demonstrated to effect more specific instructional adjustments and greater student achievement as compared to traditional forms of performance monitoring or progress monitoring displayed by a graph. This study returns performance feedback to students in two treatment conditions, as a progress graph supplemented with an analysis of mastery level by topic or solely as a progress graph. The current study examines whether instructional adjustments made by students differ by type of feedback received.

In her review of curriculum-based measurement systems, Cusumano (2007) provides an insight into how this assessment system has been used by educators to evaluate and drive instruction. Curriculum-based measurement features bring into focus the areas of instruction that are in need of additional consideration and highlight the effectiveness of instructional strategies. Within a traditional CBM progress monitoring system, the teacher uses this data to inform teaching practices and to develop more specific and realistic learning goals for individual students, and in some instances for the class in general. This study places those decisions in the hands of the student and examines whether students use performance feedback to make their own instructional decisions, in terms of what to study and for how long.

A consistent theme in the research on the psychological dimensions associated with the learner and learning concerns the role of the teacher to provide opportunities for the learner to develop metacognitive skills. The learner-centered principles that concentrate on developmental and social factors advise teachers to help students to examine their learning practices and assist students with modifying them, when necessary. The research of Nietfeld et al. (2006) lends support for the benefits of providing students with monitoring exercises on their understanding of course materials and feedback on their calibrations in terms of calibration accuracy and test performance. Lan (1996) provided his students with a self-monitoring form that directed their attention to study time, help-seeking, and self-efficacy and noted these students had higher achievement levels than students who monitored how long the instructor spent on topic or did not engage in monitoring activities at all. Zimmerman and Paulsen (1995) call upon teachers to formally instruct learners how to self-monitor learning behaviors and to be cognizant of the changes in performance brought about by changes in learning behaviors. This research affords learners the opportunity to examine their learning practices through feedback, both in a graphed form and as a mastery skill analysis. It also provides opportunities for learners to monitor their understanding and their study decisions.

The current study draws upon the framework of CBM in a content area of a first course in engineering calculus and allows the student to design his or her own educational course of action in response to performance feedback. Two research areas are addressed in this study. First, to what extent does performance feedback inform students of their progress when contrasted to students who do not receive individualized performance feedback? Second, will the instructional decisions that students make differ based on the form of feedback received?

I hypothesize the following:

H_1: Students who receive individualized performance feedback on their weekly Calculus-CBM quizzes will have more accurate measures of grade calibration when contrasted to students who do not receive individualized feedback on their weekly Calculus-CBM quiz performance.

H_2: Students who receive individualized performance feedback on their weekly Calculus-CBM quizzes will have a higher rate of relearning topics per unit of study when contrasted to students who do not receive individualized feedback on their weekly Calculus-CBM quiz performance.

H_3: Students who receive individualized performance feedback on their weekly Calculus-CBM quizzes as a progress graph supplemented with an analysis of mastery level per topic will have greater learning efficiency in terms of engagement time in academic activities as compared to students who receive individualized performance feedback on their weekly Calculus-CBM quizzes only as progress graph.

H_4: Students who receive individualized performance feedback as a progress graph supplemented with an analysis of mastery level per topic (a) will have more positive expressions of perceived usefulness of the feedback to guide study time on individual topics, direct the individual to relearn topics not mastered, and to judge which topics the student could answer correctly when compared to students who receive individualized performance feedback only as a progress graph, and (b) will have stronger correlations between perceptions of the usefulness of the feedback to guide study time and reported time engaged in academic activities, as well as, between perceptions of the usefulness of the feedback to direct the student to relearn topics not mastered and measures of relearning, when compared to students who receive individualized performance feedback only as a progress graph.

Chapter 3

Methodology

Overview of the Study

This study drew upon the framework of Curriculum-Based Measurement testing procedures in the content area of a first course in engineering calculus and examined the instructional utility of performance feedback generated from weekly quizzes. Individualized performance feedback in the form of a graphed progress indicator, and for some participants the graph was supplemented with an analysis of mastery level per topic, was returned to the student to be used as a tool to empower the learner to determine one's own educational course of action in response to the weekly individualized feedback. This study followed a quasi-experimental design, with students randomly assigned to mathematics section and the researcher randomly assigning section to treatment condition. The study investigated the extent to which students responded to feedback not only in terms of performance, but also as, reflected in the psychological factors of self-regulatory processes such as goal-setting, self-monitoring, and self-evaluation.

Two research areas were addressed in this study. First, does providing individualized performance feedback enhance the student's metacognitive awareness of calculus capability, as measured by stronger calibration between predicted and actual exam scores, and does this awareness direct students to relearn material not mastered when contrasted to students who did not receive individualized performance feedback on their weekly curriculum-based measurement quiz? Second, when students receive individualized feedback do their perceptions of their ability to use the feedback to make instructional decisions, such as the amount of time the student spent engaged in academic activities, studying topics not yet mastered, and method of studying, differ based on the form of feedback received? Do student perceptions of the usefulness of the type of feedback, progress graph or mastery analysis, correlate with measures of relearning and engagement time

Participants

Students, from a small northeastern public specialized 4-year college who were enrolled in the researcher's first semester engineering calculus course, Math 101, were recruited to participate in this study. The College's specialized mission is within the maritime industry both ashore and afloat and has an active program of licensing Merchant Marine officers. For students electing not to pursue a Merchant Marine license, there are internship opportunities.

All students enrolled in Math 101 have satisfactorily passed a placement exam or have completed a pre-calculus course with a minimum grade of B- or 80%. Math 101 is offered every semester and meets for three 75-minute periods per week. Math 101 is the first of a sequence of three calculus courses required for all Bachelor of Engineering degree candidates and covers differential calculus of one variable. The calculus sequence is taught from Larson, Hostetler, and Edwards' *Calculus, Early Transcendental Functions*. Students must meet a benchmark course grade of C- or 70% to continue into Math 102, Calculus 2. This benchmark serves as a gateway into Physics, which in turn is a gateway into the first course of the engineering curriculum, Statics.

Incoming first year students are randomly assigned to mathematics section through a scheduling algorithm of the Registrar's office. For the Spring 2008 term, in anticipation of this study, the Provost permitted the Registrar to collapse two sections of Math 101 that met at the same time into a single section, with a designation of instructor, ‘To be announced’. The Registrar then randomly assigned students to two sections, one of which was taught by the researcher.

Students enrolled in the researcher’s two Fall 2008 sections of Math 101 were invited to participate in this study on performance feedback indicators following the first hourly exam, at week 4 of a 15 week term. This time frame corresponds to the College’s first report on freshman progress. The researcher described the performance feedback indicators of the study (see Appendix A) and asked students for their consent to release upon completion of the course, the following class information for analyses: self-monitoring of study habits, calibration of exam scores, weekly curriculum-based measurement scores, an exit questionnaire, and course grade information. Each letter of consent (see Appendix B) was returned to the researcher in a sealed envelope and the student’s decision to participate in the study remained unknown to the researcher for the remainder of the term. The researcher randomly assigned the Math 101 sections to the two feedback conditions. The section designated Math 101-02 received individualized performance feedback as a progress graph based on weekly Curriculum-Based Measurement quizzes. This group is denoted as IF for individualized feedback. The other section, Math 101-03 received individualized performance feedback on weekly Curriculum-Based Measurement quizzes as a progress graph supplemented with an analysis of mastery level per topic and is denoted as IFM, individualized feedback plus mastery analysis.

Prior to the implementation of the feedback program, participants received an absolute measure of performance, that is their grade earned on the weekly quiz without any other performance or progress feedback. Commencing in week 5 of the semester, individualized performance feedback, either as a progress graph or as a progress graph supplemented with an analysis of mastery level per topic, became a regular classroom practice for all students. The researcher maintained folders that included performance feedback, self-monitoring of study habits, and exam calibration forms for all students who received individualized performance feedback.

The researcher opened the letters of consent after course grades were recorded by the Registrar. Ninety-six percent of recruited students agreed to participate in the study. Identifying numbers were assigned to forty students who agreed to participate in the study and completed the course. For the study on the effects of form of feedback on study habits and perceptions of usefulness of feedback, this sample size was sufficient to detect a medium effect size with an alpha $\alpha = .05$ (Gall, Gall, & Borg, 2003). The two feedback sections had equivalent instruction and the same course requirements. Every effort was made to have uniform experiences for both feedback groups.

A contrast group of twenty-seven students who did not receive individualized performance feedback on weekly curriculum-based measurement quizzes was formed to examine the effect of individualized performance feedback on grade calibration and rate of relearning topics per unit of study. The contrast group, denoted as no IF, was drawn from the researcher's students who were enrolled in Math 101 during the academic year 2007 – 2008. The contrast group was comprised of an intact section of Math 101 from the Fall 2007 term and first-time calculus students from the Spring 2008 section of Math 101.

For the study on the effect of feedback on grade calibration and instances of relearning, this sample size was sufficient to detect a medium effect size with an alpha $\alpha = .05$ (Gall, Gall, & Borg, 2003). Material was presented in parallel fashion; the researcher used the same note set and textbook for all three groups. The researcher was the course supervisor for all terms involved in this study and wrote the exams for all sections.

Of the 67 participants, 86.4% were male and 13.6% were female. In terms of ethnic origin, participants in the sample were 74.6% White, 9% Hispanic, 6% African American, 4% Asian Pacific Islander, and 4.4% were classified as international students, whose first language is English. These proportions accurately reflect the demographics of the institution. For 85.1% of the participants, this was their first experience with calculus.

Measures

Calculus 1 is divided into four units of study: functions and limits, basic differentiation techniques, advanced differentiation techniques, and applications of differentiation. This study gathered data over units 2 through 4.

- Unit 1 is comprised of functions: algebraic, trigonometric, and transcendental; and the study and determination of limits in three forms: numerically by constructing a table, by the inspection of a graph, and analytically. Unit 1 was completed prior to the implementation of feedback procedures.
- Unit 2 is comprised of basic techniques of differentiation, such as the definition of the derivative, the power rule, the product and quotient rules, the derivative of the exponential function and the trigonometric functions, and their applications.

- Unit 3 includes the topics on advanced differentiation techniques such as, higher derivatives and their applications, derivatives of composite functions, implicit differentiation, related rates, and the derivatives of non-algebraic and inverse functions.
- Unit 4 is comprised of applications of differentiation, such as the first and second derivative tests, curve sketching, and optimization problems.

Self-monitoring protocol. This form was used by students to record the time spent weekly engaged in academic activities for each unit of study and the methods of study the student employed. Time engaged in academic activities was recorded as minutes per week. The protocol was an indicator of which topics the student studied in preparation for a Calculus-curriculum based measurement quiz, and classified according to whether or not the student had attempted to relearn material that was not previously mastered. A topic was coded as ‘1’ if the student indicated that he or she studied a topic not yet mastered or ‘0’ if the student did not indicate he or she studied a topic that was not yet mastered. Methods of study – reviewed notes, read the textbook, reworked homework problems, and used additional resources, such as instructional DVDs or online help sites, were coded ‘1’ if the student selected the method and ‘0’ if no selection was made. The self-monitoring form is adapted from the research of Lan (1996) on the effects of self-monitoring of topic mastery and time spent studying per topic with graduate students enrolled in statistics. This self-monitoring form also asked students to rate their confidence in their ability to use the progress feedback to improve their academic performance and to meet their self-selected course grade goal. The self-monitoring form was distributed in class prior to the weekly Calculus curriculum-based measurement quiz and collected before the quiz was administered. There are three protocols (see Appendices C - E) that correspond to the three course units studied over the research period.

Calculus curriculum-based measurement quizzes (Calculus-CBM). Corresponding to the topics listed on each self-monitoring form are calculus test items. Weekly quizzes are formative assessments of the student's understanding of topics within each unit of study. Since Calculus is an advanced content area, it cannot be measured by a single skill assessment of knowledge. The level of question type within the quiz is varied and follows the research of Espin et al. 2005, Tindal and Nolet, (1995), and Foegen et al., (2006). Quiz items included three types of questions that reflect the NAEP achievement levels (a) a Basic level of achievement that involves solving mathematical problems that require the direct application of concepts and procedures in familiar situations, (b) a Proficient level of achievement that requires students to select strategies to solve problems and integrate concepts and procedures, and (c) an Advanced level of achievement where students demonstrate in-depth knowledge of the mathematical concepts and procedures represented in the framework. An example of a question that tests at a Basic level of achievement is:

> Find the derivative of $f(x) = 4x^3 + 6x - 3$.

An example of a Proficiency level test item is:

> The position of a particle moving along the t-axis is described by the function $s(t) = t^4 - 2t^2$ where position ***s*** is measured in centimeters and time ***t*** is measured in seconds. Find the particle's speed and direction of travel at 0.5 second.

An example of an item testing at an Advanced level of achievement is:

> Sketch the graph of an arbitrary function such that $f'(x) > 0$ over the interval $(-\infty, -4)$, $f'(x) < 0$ over the interval $(-4, 6)$ and $f'(x) > 0$ over the interval $(6, \infty)$. The function is continuous everywhere and *f(x)* is not differentiable at $x = 6$.

Each type of achievement level item may represent a previously taught item, a current curriculum item, or an item that may be encountered in future lessons. The three item types were displayed in random order each week. Following curriculum-based measurement practices, the weekly formative assessment was timed using the common practice rule of a student's time is three times as long as the instructor's time to complete the measure. Therefore, duration of probes varied depending on concepts assessed but averaged ~ 15 minutes. Questions from the curriculum were drawn from *Diploma*, a computerized test generator keyed to the course textbook, Larson, Hostetler and Edward's *Calculus Early Transcendental Functions*, 4th Edition and the instructor's personal test bank of questions. Administration and scoring of probes were uniform.

Feedback instruments. Once the feedback program commenced, students in both groups received their individualized performance feedback when the weekly Calculus-CBM quiz was returned. Both groups were provided with a progress graph for each of the three units of study covered during the research period. The graphed progress indicator is generated by Intervention Central©. This tool plots the weekly Calculus-CBM score (y-axis) in terms of date of assessment (x-axis) and generates a *Trendline* of progress (see Appendix F). For each unit of study, an *Aimline* was drawn by the researcher from the student's current standing based on the previous exam(s) to a designation of mastery at 80%. As weekly assessments were added to the performance indicator, the trendline was automatically recalculated. Students in both feedback conditions were provided with instruction on how to interpret the feedback when they were recruited to participate in the study. Students were directed to note the difference between the slopes of the Trendline and the mastery level Aimline. The ability to see small changes in performance is an outcome of the dynamic nature of the graphed performance indicator.

Students in the IFM group who received performance feedback as a progress graph supplemented with an analysis of mastery level per topic received additional instruction when recruited on how to interpret the analysis of mastery level per topic. A table listed three items: topics as they appear on the weekly self-monitoring form, a percent of points earned per topic, and a designation of level of mastery keyed to the percent of points earned on topic questions as measured by the weekly Calculus-CBM (see Appendix G). The analyses updates mastery information each week, which enables the student to compare previous performance and changes in mastery. Mastery level designations were computed as follows:

- Mastery, at least 80% of the points assigned to this topic are earned.
- Partial Mastery, from 60% through 79% of the points assigned to this topic are earned.
- Not Mastered, less than 60% of the points assigned to this topic are earned.

To interpret the mastery level designations, students in the IFM group received an explanation of the scoring rubric in terms of the mastery levels when they were recruited to participate in the study. The rubrics are based on those developed by the University's discipline panel (SUNY, 2005) for assessment of general education skills in mathematics.

- Mastery level is demonstrated by a full understanding of the problem either through the methods chosen to answer all parts of a question successfully or by giving a complete response with a minor error in calculation, a misstatement, or an almost complete response with a correct answer.
- Partial mastery is demonstrated by an understanding of all aspects of the problem, the selection of an appropriate strategy to solve the problem, and progress toward a correct response, or only correctly responds to a portion of the question. The solution contains some correct aspects but also exhibits conceptual flaw(s).
- Non-mastery is demonstrated by a limited understanding of the problem, a solution to only part of the problem, a response that does not address the problem, or little or no progress toward a correct and complete response.

Students benefited from the dynamic nature of the mastery level feedback in that they received information on changes in level of mastery as topics were retested on the weekly Calculus - CBM. This feature offered the participant an opportunity to see incremental changes in understanding of topics within the unit. No additional instruction was given to students on how to interpret the feedback, unless a student requested it. Thus the primary role of the teacher was to convey the curriculum and was identical across all groups.

Measures of metacognitive judgments. Summative assessments were administered as unit hourly exams. To assess the student's metacognitive judgments regarding their learning and performance, prospective decisions were made by the student. Prior to the distribution of each hourly exam, students were directed to submit monitoring judgments as a prediction of their confidence to answer questions on each topic within the unit of study, as well as, an overall prediction of expected performance on the exam. The individual topics were listed as they appeared on the weekly self-monitoring form. This format used self-efficacy judgments as a tool for the student to "assess task complexity and estimate performance difficulty" (Schraw, 2008). The combination of a fine-grained approach of self-efficacy judgments and a larger-grain approach of prediction of score returned more information on the test and also provided the learner with a greater opportunity to make deeper metacognitive evaluations. This combination of levels of monitoring judgments is supported by the research of Nietfeld et al. (2005). On this self-efficacy form, students were additionally asked to select their personal course grade goal (see Appendices H – J for units 2 through 4).

This form was returned to students with performance scores corresponding to each topic and overall score. Students were prompted to note instances of accuracy and inaccuracy between their prediction of their capability and their actual score. No other direction was given to the students that would unduly influence the role of the performance feedback for one group over the others.

Exit Questionnaire. A self-report questionnaire was administered at the end of term. Students were asked to rate the usefulness of the individualized performance feedback to: reach their course grade goal, improve academic performance, guide study time, direct learning, and rate their self-efficacy in calculus topic capability. Both feedback groups were asked to rate five questions on the utility of the progress graph (see Appendix K), while the IFM group rated an additional three questions (see Appendix L) on the usefulness of the analysis of mastery level per topic to: guide study time, direct learning, and rate their self-efficacy in calculus topic capability. Responses use a likert scale that range across the attitudes: Very much, Quite a bit, Somewhat, Very little, and Not at all.

For the five items that were common to both forms of feedback the Cronbach alpha coefficient was .80, which is an acceptable measure of internal consistency reliability for this scale. To assess the degree to which the responses of the group that received individualized feedback as a progress graph supplemented with an analysis of mastery level per topic were consistent across the three additional questions regarding the usefulness of the analysis of mastery level, a second reliability analysis yielded the Cronbach alpha coefficient of .85, which is a measure of good internal consistency reliability for this scale.

Chapter Four

Results

This study investigated the effectiveness of Curriculum-Based Measurement testing and individualized performance feedback procedures with students enrolled in an initial course in engineering calculus. The study departs from traditional teacher-led responses to student performance indicators to allow the student to regulate his or her own educational course of action in response to performance feedback. The overarching questions of this study are to determine to what extent individualized performance feedback informs students of their progress and directs the instructional decisions students make.

The results of this study are presented in two sections. The first section presents the findings on the hypotheses concerning differences among the three groups: no individualized feedback (no IF), feedback as a progress graph (IF), and feedback as a progress graph supplemented with an analysis of mastery level per topic (IFM), in the areas of calibration accuracy and rate of relearning. The second section of this chapter addresses differences in study habits and perceptions of the usefulness of the feedback instruments between the two feedback conditions, IF and IFM.

Placement into mathematics course for incoming students is based primarily on a placement test, in conjunction with SAT – Mathematics score. The average score of study participants on the placement exam was 41.4 based on 60 points, and there were significant group differences on placement test score, $F(2,58) = 7.251$, $p = .002$. Post hoc comparisons revealed significant differences on placement test score were between the contrast group (no IF) and the group that received feedback as a progress graph (IF),

$t(58) = 3.328$, $p = .002$, and between the contrast group (no IF) and the group that received feedback as a progress graph supplemented with an analysis of mastery level (IFM), $t(58) = 3.151$, $p = .003$. There were no significant differences on placement test score between the two feedback groups, IF and IFM, $t(58) = .348$, $p = .729$. Other pre-existing factors were determined not to be significantly different among the three groups. Average SAT-mathematics score of the participants was 583; there were no significant differences among the three groups, $F(2,61) = 2.233$, $p = .116$. There were no significant differences between groups regarding the number of students repeating the course, Math 101: $\chi^2(2, n = 67) = 1.928$, $p = .381$. Placement test score was selected as the covariate for this study for analyses involving all three groups. The placement test is a standardized test from the College Entrance Examination Board (CEEB) that assesses student proficiency in both intermediate algebra and functions.

Section I

Hypothesis 1

Students who receive individualized performance feedback will have more accurate measures of grade calibration when contrasted to students who do not receive individualized feedback on their performance.

Descriptive statistics. In Table 1 descriptive statistics (using raw scores) are presented for the study variables of Hypothesis 1.

Table 1
Means and standard deviations of major variables

Variable	IFM group		IF group		Contrast group		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Placement Score	43.83	7.81	44.71	9.09	36.74	5.67	40.41	8.23
Exam 1 (Pre-test)	78.36	13.11	77.11	14.68	74.54	13.71	76.49	13.68
Exam 2 score	75.32	13.44	76.58	13.17	68.65	17.53	73.04	15.33
Exam 2 calibration	11.96	9.36	17.81	10.55	16.58	15.95	15.37	12.71
Exam 2 confidence	77.83	12.88	74.89	15.39	82.58	9.10	78.90	12.55
Exam 3 score	71.18	16.68	75.00	14.31	66.30	17.03	70.24	16.38
Exam 3 calibration	10.00	7.42	14.33	12.30	17.54	12.19	14.10	11.15
Exam 3 confidence	67.91	11.71	67.67	11.75	71.00	14.07	69.03	12.59
Exam 4 score	67.74	16.68	65.53	14.36	58.80	17.35	63.48	16.70
Exam 4 calibration	11.41	9.05	12.19	8.84	17.33	13.37	13.99	11.13
Exam 4 confidence	71.62	10.23	69.67	15.39	68.35	16.12	69.77	14.11
Final Exam	67.05	15.00	70.18	14.31	61.33	17.49	65.58	16.07

Placement scores correspond from 0 to 60. Scores for exam performance correspond to percent correct. Scores for calibration range from 0 to 100, with scores closer to zero being more accurate. Scores for confidence range from 0 to 100.

A measure of *absolute accuracy*, as defined by Schraw (2008), is the precision of a confidence judgment compared to performance on a criterion task. For each student an index of absolute accuracy or calibration was computed as the absolute value of the difference between the predicted or confidence score on each hourly exam during the research period, C_i for $i = 1 - 3$, and the performance score on each hourly exam, P_i for $i = 1 – 3$, or $|C_i - P_i|$. Predicted score and performance score are continuous measures.

Data was screened for outliers and abnormality. The dependent measures, calibration score for each exam, were transformed by the square root function to meet the assumption of normality as confirmed by the Kolmogorov-Smirnov statistic. The transformed calibration scores were converted to standardized *z* scores, variables Z1, Z2, and Z3, to control for differences in difficulty across exams. Higher *z* scores indicate more accurate calibration.

To assess whether exam score calibration is more accurate when students receive individualized performance feedback compared to students who did not receive individualized performance feedback on weekly Calculus-CBM quizzes, calibration scores were analyzed by a 3 x 3 analysis of covariance (ANCOVA) with standardized measure of calibration (Z1, Z2, and Z3) as a within-subject factor and group (no IF, IF, and IFM) as a between subjects factor. The placement test to assess readiness to enter the calculus sequence was used as a covariate.

A preliminary check was conducted to ensure that there were no violations of a mixed-model ANCOVA. Homogeneity of inter-correlations among the levels of the within-subjects variable, standardized calibration, was the same for each level of the between-subjects variable, group (Box's M: $p = .569$). Conditions for homogeneity of variance in scores for the three groups were met as determined by Levene's Test of Equality of Error Variance for variable Z1: $F(2,55) = 1.649$, $p = .202$; variable Z2: $F(2,55) = 1.427$, $p = .249$; and variable Z3: $F(2,55) = .103$, $p = .902$. Additionally, the covariate, placement test score, was pre-screened for linearity between it and each dependent variable, standardized calibration for each exam, and for homogeneity of regression slopes, that is, there was no interaction between the covariate placement test score and the dependent measures.

The Sphericity assumption that requires the variance of the population difference scores for any two conditions to be the same as the variance of the population difference scores for any other two conditions was attained with $p = .351$. There was no significant interaction between group and time on the dependent measure, calibration accuracy, $F(4,108) = .653$, $p = .623$, partial eta squared = .024 or between placement test score and time, $F(2,108) = .031$, $p = .970$, partial eta squared = .001. There was no significant main effect for time on calibration accuracy, $F(2,108) = .066$, $p = .936$, partial eta squared is .001.

The main effect for comparing groups on the dependent measure, calibration accuracy was marginally significant, $F(2,54) = 2.713$, $p = .075$, partial eta squared = .091, which is an effect size of medium strength. As the hypothesis compares the effects of receiving individualized performance feedback on calibration accuracy when contrasted to calibration accuracy of students who do not receive performance feedback, analyses showed differences approaching significance on calibration accuracy between the two feedback groups, IF and IFM, and the no IF group, $F(1,54) = 3.875$, $p = .054$, partial eta squared = .066, which is a effect size of medium strength. Individual comparisons revealed significantly higher calibration accuracy ($p = .025$) by the IFM group compared to the no IF group. Eta-squared was calculated as .11.

Figure 1 presents the means of the measures of calibration for each exam, as standardized z scores, adjusted by the placement test score. Students in the IFM group were on average .52 standard deviations higher in standardized calibration accuracy than the no IF group. The magnitude of the differences in the means is an effect size of medium strength.

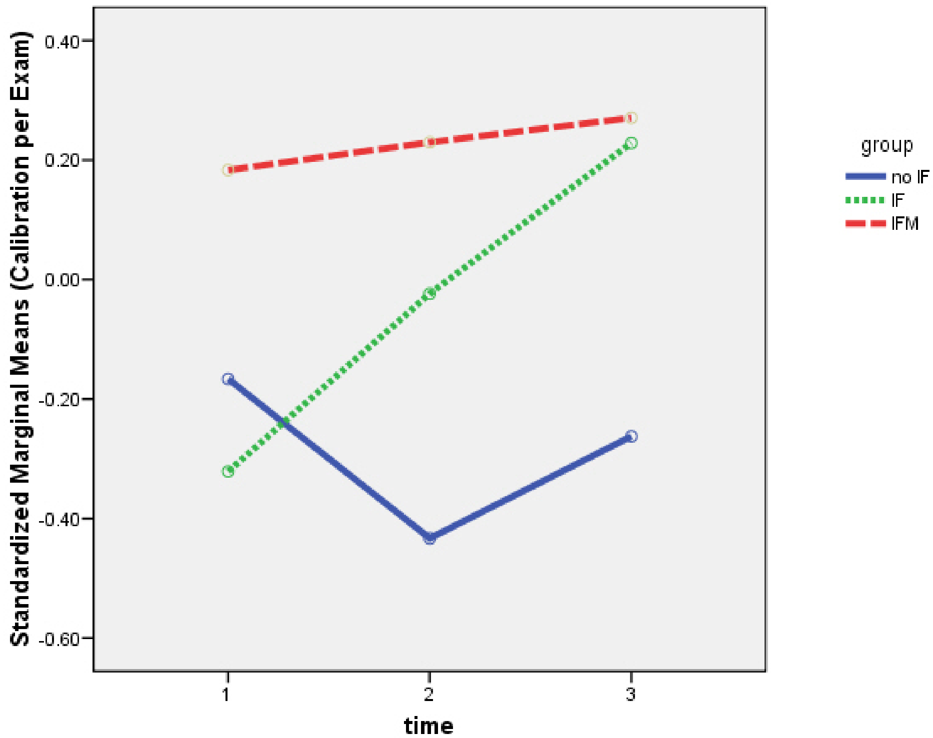


Figure 1: Standardized means of calibration scores for each exam, adjusted by the covariate placement test score

The IFM group that received feedback as a progress graph supplemented with an analysis of mastery level per topic consistently calibrated exam performance more accurately than the other two groups. The group (no IF) that did not receive individualized performance feedback and the group (IF) that received individualized feedback as a progress graph were less accurate in their calibration accuracy that the IFM group by over 5 points at the initial calibration point, Time 1.

The IF group improved its calibration accuracy over time 1 to time 3, approximately 5.5 points and its final measure of calibration is equivalent to the IFM group. The contrast group (no IF) had a net decrease of 1 point in their grade calibration accuracy over time. When compared to the calibration accuracy of the IFM group, the no

IF group was on average 6 points less accurate in their calibration accuracy over the three exams.

When students are given individualized feedback on weekly Calculus-CBM quiz performance, they gain insight into their calculus capability as demonstrated by their measures of calibration accuracy. In particular, when students receive the benefit of individualized performance feedback as a progress graph supplemented with an analysis of mastery level per topic (IFM) they have significantly greater metacognitive awareness of their ability to assess what skills they bring to an exam when compared to students who do not receive individualized performance feedback (no IF) on weekly quizzes. These findings lend support to Hypothesis 1, "Students who receive individualized performance feedback will have more accurate measures of grade calibration when contrasted to students who do not receive individualized feedback on their performance".

Hypothesis 2

Students who receive individualized performance feedback on their weekly Calculus-CBM quizzes will have a higher rate of relearning topics per unit of study when contrasted to students who do not receive individualized feedback on their weekly Calculus-CBM quiz performance.

Each weekly Calculus-CBM quiz contained questions that tested current curriculum topics, topics within the unit previously taught, and an item not yet encountered. To calculate the rate of relearning topics, the researcher recorded points earned per topic per week throughout the study. For each unit of study, a summary of topics was prepared that computed a cumulative score for each item as it was retested. For example, suppose on topic A, a student earned 2 of 5 points or 40% on quiz 1, earned

3.5 of 5 points or 70% on quiz 2, and earned 8 of 11 points or 73% on the hourly exam. The student's summary for topic A was recorded as follows: time 1 = 40%, time 2 = 55% (5.5/10 points), and time 3 = 64% (13.5/21 points). The researcher aggregated points per retest to account for differences in the type of question achievement level for the same topic.

A point was assigned for each instance that resulted in a ten-point improvement within either non-mastery or partial mastery levels, a change in mastery level, or a 5 point improvement at the 'mastered' level. The student in this example initially had a performance score of 40% on topic A and the next assessment resulted in a performance score of 55% on topic A. The student would be awarded one point as an indication of relearning because there was an improvement of 10 points within the 'not mastered' level. The student would be awarded another point from time 2 to time 3 because the student had a change in mastery level from' not mastered' to 'partially mastered'. For each unit of study, the points awarded were summed per student. To ensure coding fidelity, a sample of performance data from the groups was coded by another researcher at 2 points in time: week 7 and week 10 of the semester. The second researcher was provided with scoring training; there was 94% coding agreement at week 7 and 98% coding agreement at week 10 of the semester.

To assess whether students who receive individualized performance feedback have a higher rate of relearning topics per unit of study within the Calculus-CBM quiz format, a 3 x 3 analysis of covariance (ANCOVA) with instances of relearning (Time 1, Time 2, and Time 3) as a within-subject factor and group (no IF, IF, and IFM) as a between subjects factor was performed. For each unit of study, the tally of awarded points was converted into a percentage as instances of relearning out of total opportunities of relearning as a measure of rate of relearning per unit of study. The placement test score that assesses a student's readiness to enter the calculus sequence was used as a covariate. In Table 2 descriptive statistics (using raw scores) are presented for the dependent measure, rate of relearning, for all groups.

Table 2
Means and standard deviations for the rate of relearning

	Time 1		Time 2		Time 3		
	M	*SD*	*M*	*SD*	*M*	*SD*	*n*
No IF group	41.25	30.87	29.17	25.47	33.02	23.35	24
IF group	51.38	29.44	55.29	17.72	48.24	17.41	17
IFM group	53.62	22.36	50.95	18.14	50.48	19.62	21
Total	48.22	27.97	43.71	23.95	43.10	21.84	62

Scores represent the percent of instances of relearning out of total opportunities for relearning

Data were screened for outliers and abnormality. The assumption of normality was confirmed by the Kolmogorov-Smirnov statistic. A preliminary check was conducted to ensure that there were no violations of a mixed-model ANCOVA. Box's test for homogeneity of inter-correlations is supported with $p = .541$. Conditions for homogeneity of variance in scores for the three groups were met as determined by Levene's Test of Equality of Error Variance at Times 1 and 3; Time 1: $F(2,59) = 2.125$, $p = .129$; Time 2: $F(2,59) = 7.776$, $p = .001$; and Time 3: $F(2,59) = 1.609$, $p = .209$. The Sphericity assumption that requires the variance of the population difference scores for any two conditions to be the same as the variance of the population difference scores for any other two conditions was attained with $p = .695$. The covariate, placement test score, was pre-screened for linearity between it and each dependent variable, rate of relearning per unit, and for homogeneity of regression slopes, that is, there was no interaction between the covariate placement test score and the dependent measures.

There was no significant interaction between group and time on the dependent measure, rate of relearning, $F(4,116) = .804$, $p = .525$ or between placement score and time, $F(4,116) = .268$, $p = .795$. There was no significant main effect for time, $F(2,118) = .121$, $p = .887$ on the dependent measure, rate of relearning.

The main effect for comparing groups on the dependent measure, rate of relearning was significant, $F(2, 58) = 4.298$, $p = .018$, partial eta squared = .129, which is a large effect size. Comparisons to the no IF group, revealed significant differences between the no IF group with each of the feedback groups, the IF group ($p = .018$) and the IFM group ($p = .009$). In both cases the inclusion of feedback increased the rate of relearning topics on subsequent assessments. There were no significant differences between the IF and the IFM group.

Details of differences in rate of relearning in the three groups over three time periods, given as mean score, adjusted by the covariate placement test score are presented in Figure 2. Average rates of relearning over the study are 36.8% for the no IF group, 50.0% for the IF group, and 50.4% for the IFM group. The groups that received individualized performance feedback on weekly Calculus-CBM quizzes have an overall increase of more than 13 percentage points in relearning over the study period when compared to the no IF group that did not receive weekly individualized performance feedback

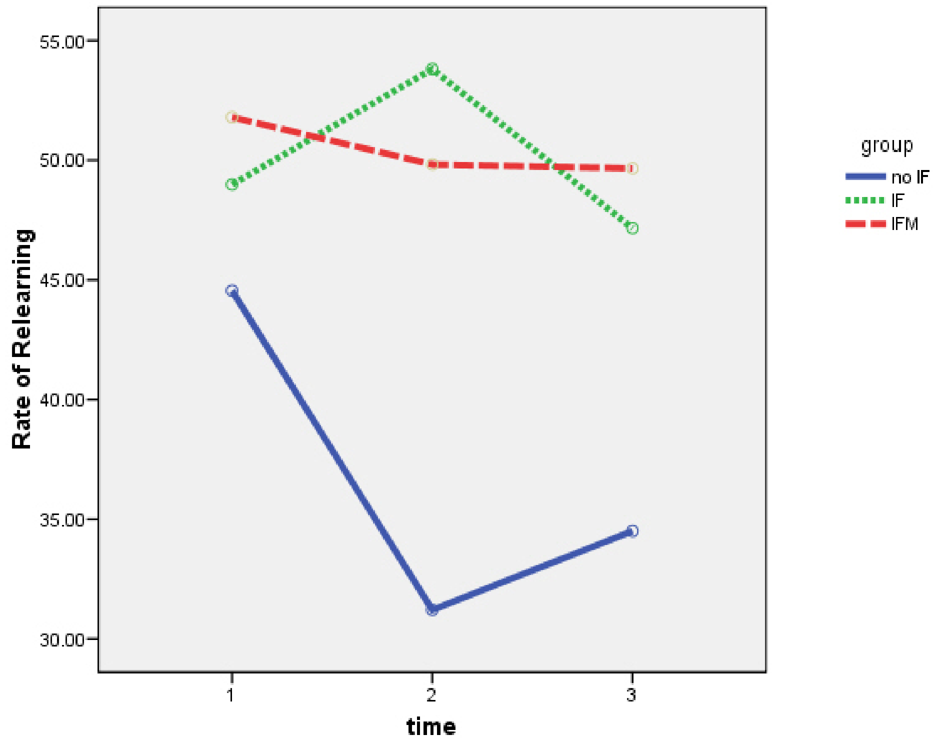


Figure 2 . Mean rate of relearning for each group at each time period, adjusted by the covariate, placement test score.

The results of the analyses suggest that although all three groups were quizzed by Curriculum Based Measurement procedures, the two groups that received weekly individualized performance feedback had a higher rate of relearning over the three units of study than the group that did not receive weekly individualized performance feedback. Therefore Hypothesis 2 “Students who receive individualized performance feedback on their weekly Calculus-CBM quizzes will have a higher rate of relearning topics per unit of study when contrasted to students who do not receive individualized feedback on their weekly Calculus-CBM quiz performance” is supported.

Supplementary Analyses

The finding of Hypotheses 1 and 2 suggest that providing students with individualized performance feedback on weekly Calculus-Curriculum Based

Measurement quizzes empowers the student to develop the qualities of a successful learner, in particular to have a greater metacognitive awareness of one's calculus capabilities and to relearn topics that were not mastered. These findings raise the question: Are there differential growth rates in learning across exams for the three study groups?

Students enrolled in calculus are deemed calculus ready either having demonstrated pre-calculus proficiency on the placement test or having passed a pre-calculus course with a minimum grade of B-. Prior to examining any treatment effects on the growth rate in learning across exams, a one-way analysis of variance (ANOVA) was performed to determine whether the three groups had performance differences on the first exam before the feedback program began. The dependent variable, Exam 1 score, is continuous and normally distributed. No statistical differences were found on Exam 1 performance between the three groups: $F(2, 64) = .493$, $p = .613$. Therefore, students from the three groups did not differ significantly in exam performance before the feedback procedures were implemented.

To examine whether exam performance of the three groups was different across the four subsequent testing periods, a 4 by 3 analysis of variance (ANOVA) was performed with testing period for each subsequent exam (Exam2, Exam 3, Exam 4, and Final Exam) as a within-subject factor and group (no IF, IF, and IFM) as a between subjects factor.

Exam scores were standardized to control for differences in difficulty across exams. The assumptions for a mixed-model ANOVA were met. Box's test for homogeneity of inter-correlations is supported with $p = .990$. Conditions for homogeneity of variance in scores for the three groups were met as determined by Levene's Test of Equality of Error Variance for variable ZX2: $F(2,60) = 1.526$, $p = .226$; variable ZX3: $F(2,60) = .755$, $p = .474$; variable ZX4: $F(2,60) = .910$ $p = .408$; and variable ZFX: $F(2,60) = 1.417$, $p = .250$. The Sphericity assumption that requires the variance of the population difference scores for any two conditions to be the same as the variance of the population difference scores for any other two conditions was attained with $p = .067$.

There was no significant main effect for time, $F(3,180) = .013$, $p = .998$ or significant interaction between group and time, $F(6,180) = .503$, $p = .806$ on exam performance. However, the main effect for comparing groups on exam performance was marginally significant, $F(2,60) = 2.646$ $p = .079$, partial eta squared = .081, which is an effect size of medium strength.

Comparisons to the no IF group were included to further examine any differential performance between the group that did not receive individualized feedback and the two feedback groups. Findings reveal differences in the dependent measure, subsequent exam performance for the IF group, $p = .032$ and for the IFM group, $p = .061$ when each is compared to the no IF group. A second analysis was performed to additionally examine any differential performance between the no IF group and the combined feedback groups. Mean differences on performance on exams between the no individualized feedback group and the combined individualized feedback groups were significant, $F(1,61) = 4.997$, $p = .029$.

Table 3, presented below, displays performance data (raw scores) for the combined feedback groups.

Table 3
Means and standard deviations for exam performance

	Feedback Groups		Contrast Group	
	M	*SD*	*M*	*SD*
Exam 2	75.89	13.17	68.65	17.53
Exam 3	72.90	15.58	66.30	17.03
Exam 4	66.72	15.64	58.80	17.35
Final Exam	68.41	14.60	61.33	17.49
Average	70.86	12.27	63.41	13.94

Providing students with individualized performance feedback on weekly Calculus-CBM quizzes resulted in an average 7.5 point increase in performance on subsequent exams when compared to the subsequent exam performance for students who did not receive individualized performance feedback. The supplementary analyses support the supposition that providing students with individualized performance feedback empowers students not only to make instructional decisions to relearn material, but also results in improved student performance.

Section II

Hypothesis 3

Students who receive individualized performance feedback on their weekly Calculus-CBM quizzes as a progress graph supplemented with an analysis of mastery level per topic will have greater learning efficiency in terms of engagement time in academic activities as compared to students who received individualized performance feedback on their weekly Calculus-CBM quizzes only as a progress graph.

Summary of responses. Prior to the administration of the weekly Calculus-CBM assessment, the self-monitoring protocol was distributed. Students checked which topics they studied, which methods of study they used to prepare for the weekly Calculus-CBM quiz, and indicated the total time spent engaged in academic activities for the week.

Reported topics studied. Topics covered within the unit of study were listed on the self-monitoring protocol. If a topic checked by the student was previously taught and not yet mastered, it was coded as '1'. A topic that was not yet mastered and the student did not study was coded as '0'. A tally of topics not yet mastered and reported as 'studied' was summed over the study period, as well as a sum of instances a topic not yet mastered was not reported as 'studied'.

A chi-square test of association was calculated comparing the frequency of studying topics not yet mastered as reported by students in the two feedback groups. A significant association between feedback group and topics studied was found, $\chi^2(1, n = 256) = 4.518$, $p = .034$. The IFM group that received individualized performance feedback on their weekly Calculus-CBM quiz as a progress graph supplemented with an analysis of mastery level per topic reported studying 'items not yet mastered' more frequently (56.5%) when compared to the IF group that received individualized performance feedback as a progress graph each week (43.2%).

Methods of study. To assess the fidelity of the responses reported by students on which methods of study were used; correlations between the methods of study were calculated for each group using Spearman's *rho* and are given below in Table 4.

Table 4
Correlations between methods of study for each feedback group.

	IFM Group				IF Group			
Measure	1	2	3	4	1	2	3	4
1 Reworking homework	1				1			
2 Reviewing notes	.514*	1			.318	1		
3 Reading the text	.351	.248	1		.142	.057	1	
4 Using additional resources	.577**	.040	.319	1	.038	.201	.-316	1

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

For the IFM group, ‘reworking homework’ significantly correlates with the study methods ‘reviewing notes’ and ‘using additional resources’. There are no significant correlations between study methods for the IF group. The correlations suggest the addition of the mastery level component to the individualized feedback progress graph provides the student with additional information to make judgments about how best to prepare for weekly Calculus-CBM quizzes.

The following methods of study were analyzed: reviewed notes, read the textbook, reworked homework, and used additional resources, such as online help sites, end of chapter problems, and viewing instructional videos in the library. A tally of each method of reported use weekly was summed over the three units of the study. The method of study was coded as ‘0’ if the student did not use the method, ‘1’ if the student used the method once per unit, and as ‘2’ if the student used the method at least twice per unit.

Chi-square tests of association between feedback group and reported method of preparing for weekly Calculus-CBM quizzes were performed. Results, given in Table 5, show no significant association between groups and how they prepare for weekly Calculus-CBM assessments. When the tally of frequencies for method of study was recoded as ‘0’ for did not use the method and ‘1’ as used the method at least once per unit of study, a marginally significant association between groups and the method of study ‘reading the textbook’ emerged ($F(1, n = 34) = 3.265$, $p = .071$). The IF group reported ‘reading the textbook’ more frequently (75%) as compared to the IFM group (44.4%).

Table 5
Tests of association between feedback type and method of study.

	Frequency coded as: not at all; once per unit; twice per unit	Frequency coded as: not at all; at least once per unit
Reworked homework	χ 2 (2) = .971, p > .05	χ2 (1) = .283, p > .05
Reviewed notes	χ 2 (1) = .384, p > .05	χ2 (1) = .000, p > .05
Read the text	χ 2 (2) = 3.418, p > .05	χ2 (1) = 3.265, p < .10
Used additional resources	χ2 (2) = 4.125, p > .05	χ2 (1) = .389, p > .05

Time spent on academic activities. Weekly time spent engaged in academic activities was aggregated into the average number of minutes per week engaged in academic activities, for each unit of study. To determine the effect of form of feedback on learning efficiency of academic engagement time, a 3 x 2 analysis of variance (ANOVA), with measures of average time per week spent engaged in academic activities per unit of study across three time periods (Time 1, Time 2, and Time 3) as a within-subject factor and feedback group (IF vs. IFM) as a between subjects factor was performed.

Data was screened for outliers and abnormality. Average academic engagement time per week was transformed by a square root function to attain a more normal distribution of scores, as ascertained by a Kolmogorov-Smirnov test. The assumptions for the mixed-model ANOVA were met. Box's test for homogeneity of inter-correlations is supported with $p = .848$. Conditions for homogeneity of variance in scores for the three groups were met as determined by Levene's Test of Equality of Error Variance for variable Time1: $F(1,32) = .100$, $p = .754$; variable Time 2: $F(1.32) = .241$, $p = .627$; and variable Time 3: $F(1,32) = .165$, $p = .687$; The Sphericity assumption that requires the variance of the population difference scores for any two conditions to be the same as the variance of the population difference scores for any other two conditions was met, $p = .051$.

There was a significant interaction between the unit of study and the feedback group in engagement time in academic activities, $F(2,64) = 3.486$, $p = .034$, partial eta squared = .098, which is an effect size of medium strength. This interaction suggests that the impact of the form of feedback on engagement time in academic activities is influenced by the unit of study. There was a significant main effect for time, $F(2,64) = 19.111$, $p < .001$, partial eta squared =.374, with both groups becoming more efficient in their use of engagement time in academic activities over the term. Form of performance feedback explained an additional 27.6% of the variance in average weekly engagement time in academic activities over that explained by the interaction between unit of study and type of performance feedback. Both groups use the feedback to guide their study behaviors over time.

The main effect for comparing the two forms of feedback on engagement time in academic activities was not significant, $F(1,32) = 1.664$, $p = .206$, partial eta squared is .049.

Table 6 displays average weekly time as a function of unit of study for each feedback group. Time is measured in minutes spent weekly engaged in academic activities.

Table 6
Descriptive Statistics (raw data) of time, minutes per week, spent engaged in academic activities

	IFM Group		IF Group	
Time period: Unit of Study	*M*	*SD*	*M*	*SD*
1: Basic Differentiation	104.24	17.89	174.24	20.88
2: Advanced Differentiation	76.21	10.11	83.72	11.97
3: Applications	72.59	11.49	92.54	13.84

At the first time period, an independent samples t-test was performed to examine group differences in engagement time in academic activities. Marginal support for significant differences between groups on average weekly engagement time in academic activities at the first time period were revealed, $t(32) = 1.979$, $p = .057$. The IFM group (M = 104 minutes per week) demonstrated greater learning efficiency in terms of engagement time in academic activities as compared to the time engaged in academic activities (M = 174 minutes per week) by the IF group.

Overall, the IFM group had consistent engagement time in academic activity behavior across the research period, with an overall decrease in time of 32 minutes per week. The IF group exhibits an inconsistent pattern of engagement time in academic activities. From time 1 to time 2, the IF group decreased average weekly engagement time in academic activities over 90 minutes per week. From time 2 to time 3, the IF group increased average weekly engagement time in academic activities almost 9 minutes per week, for an overall net decrease of approximately 82 minutes per week.

The IFM group that received individualized performance feedback as a progress graph supplemented with an analysis of mastery level per topic appears to use the feedback to guide study time efficiently. The addition of the analysis of mastery level seems to facilitate the process of guiding time spent engaged in academic activities, that is, the IFM group appears to learn more in a shorter period of time. Initially, there is a 70 minute per week difference between the feedback groups in engagement time in academic activities. At time 2 the two groups are within 7.5 minutes per week of each other. The final difference at time 3 is approximately 20 minutes per week engaged in academic activities.

Therefore, Hypothesis 3, “Students who receive individualized performance feedback on their weekly Calculus-CBM quizzes as a progress graph supplemented with an analysis of mastery level per topic will have greater learning efficiency in terms of engagement time in academic activities as compared to students who received individualized performance feedback on their weekly Calculus-CBM quizzes only as progress graph” is not fully confirmed. It is only at the beginning of the study that there is marginal support for significant differences between groups on learning efficiency in engagement time in academic activities. Providing students with performance feedback appears to afford the learner insight into how best to prepare for weekly Calculus-CBM quizzes. The feedback also facilitates within the student the development of a process of selectivity on what to study and for how long. Both groups become more efficient in their academic activities as the term progressed.

Hypothesis 4

Students who receive individualized performance feedback as a progress graph supplemented with an analysis of mastery level per topic: (a) will have more positive expressions of perceived usefulness of the feedback to guide study time on individual topics, direct the individual to relearn topics not mastered, and to judge which topics the student could answer correctly when compared to students who receive individualized performance feedback only as a progress graph, and (b) will have stronger correlations between perceptions of usefulness of the feedback to guide study time and reported time engaged in academic activities, as well as between perceptions of usefulness of the feedback to direct the student to relearn topics not mastered and measures of relearning, when compared to students who receive individualized performance feedback only as a progress graph.

Analyses are presented to compare perceptions on the usefulness on the feedback instrument between the two feedback conditions, followed by associations between perceptions of the usefulness of the form of feedback with measures of study time and instances of relearning.

Descriptive statistics. An exit questionnaire was distributed to all students in the feedback condition, which asked students to rate, on 5-point Likert scale, their perceived usefulness of the feedback instrument to guide study time on individual topics, to direct the individual to relearn topics not mastered, and to help the student judge which topics the student could answer correctly. Table 6 presents mean and median response values on the usefulness of the feedback instrument.

Table 7
Mean and median statistics for perceived usefulness of the feedback

	IFM Mastery Level Analysis		IFM Progress Graph		IF Progress Graph	
Variable	*M*	*Mdn*	*M*	*Mdn*	*M*	*Mdn*
Guide study time	3.30	3	3.43	3	2.85	3
Relearn topics not mastered	3.22	3	3.39	3	3.15	3
Self-efficacy to answer correctly on topics	3.09	3	3.30	3	3.21	3

Perceived usefulness of the individualized performance feedback. To determine whether frequencies of perceived usefulness of the feedback tool to guide study time, to direct the student to relearn topics not mastered, and to judge which topics the student could answer correctly differ by type of feedback received, a chi-square test for independence was performed for each variable. The responses of the IFM group on the usefulness of the mastery analysis and usefulness of the progress graph were averaged into a single measure. Responses were coded as positive expressions for students who responded with a score of 3(*somewhat*), 4 (*quite a bit*) or 5 (*very much*) and negative expression for students who responded with a score as 2 (*very little*), or 1 (*not at all*).

Guide study time on individual unit topics. There was a statistically significant association, χ^2 (1, n = 43) = 7.406, p = .006 between feedback group and perceived usefulness of the feedback instrument to guide study time on individual unit topics. The strength of the relationship was measured by Cramer's V = .415, which corresponds to a medium-size effect. The IFM group recorded positive perceptions on the usefulness of the feedback to guide study time 91.3% of the time (frequency = 21of 23) as compared to the 55% of positive perceptions reported by the IF group (frequency = 11 of 20).

Direct the learner to relearn material not mastered. A chi-square test of association indicated no significant association, χ^2 (1, n = 43) = 1.139, p = .286, between feedback group and perceived usefulness of the feedback to direct the learner to relearn material not mastered.

Judge what topics the student could answer correctly. A chi-square test for independence indicated no overall differences, χ^2 (1, n = 42) = .090, p = .764, between feedback group and perceived usefulness of the feedback to help the learner judge what topics he or she could answer correctly.

Hypothesis 4a, "the group that received individualized performance feedback as a progress graph supplemented with an analysis of mastery level per topic will have more positive expressions of perceived usefulness of the feedback tool to guide study time on individual topics, direct the individual to relearn topics not mastered, and to judge which topics the student could answer correctly when compared to students who receive individualized performance feedback only as a progress graph" is supported for only the perceived usefulness of the feedback to guide study time.

Associations between perceived usefulness of individualized performance feedback and measures of study time and relearning. To determine the extent to which the student's perceptions of perceived usefulness of the individualized performance feedback is related to measures of study time and instances of relearning, correlations using Spearman's *rho* were computed for each group (IFM and IF), each component of feedback (mastery level analysis and progress graph), and for each measure (study time and relearning).

Perceived usefulness of the feedback instrument to guide study time and measures of time engaged in academic activities. Perceptions of the usefulness of the feedback instrument to guide the participant's study time on individual unit topics were recoded into high ratings for students who responded with a score of 4 (*quite a bit*) or 5 (*very much*) and low ratings for students who responded with a score as 3(*somewhat*), 2 (*very little*), or 1 (*not at all*). For each group, time spent engaged in academic activities per unit, in minutes per week was computed into a single measure of average time of academic engagement per week.

Mastery level analysis. Average time spent engaged in academic activities per unit, in minutes per week, was recoded as '2' for time above the median, (*Mdn* = 85.43 minutes per week.*)* and recoded as '1' for time below the median for the IFM group that receive this supplementary feedback. A Spearman Rho correlation coefficient was calculated for the relationship between participants' perceived usefulness of the analysis of mastery level per topic to guide the participant's study time on individual unit topics and weekly time spent engaged in academic activities, as measured as minutes per week, over the research period. A strong negative correlation was found (*rho* = -.778, n =18, $p < .001$), indicating a significant inverse relationship between the two variables. Participants from the IFM group with high ratings of the usefulness of the analysis of mastery level per topic in guiding the participant's study time on individual unit topics reported less time engaged in academic activities per week. Students with high perceptions of the usefulness of the mastery analysis exhibited greater learning efficiency than those with low perceptions of usefulness of the mastery level feedback.

Progress graph. Average time of academic engagement was recoded as '2' for above the median time, (*Mdn* = 85.43 minutes per week for the IFM group, *Mdn* = 105.20 minutes per week for the IF group) and recoded as '1' for below median time. A Spearman *Rho* correlation coefficient for each group was calculated for the relationship between participants' perceived usefulness of the progress graph in guiding the participant's study time on individual unit topics and average weekly time spent engaged in academic activities, as measured as minutes per week, over the research period.

There was no correlation between perceptions of usefulness of the progress graph to guide study time on individual unit topics and reported study time for IFM participants, (*rho* = -.111, n =18, $p = .66$) or for IF participants, (*rho* = .169, n =17, $p = .52$). This result indicates no significant relationship between the two variables: perceptions of usefulness of the progress graph to guide study time and reported time engaged in academic activities.

Perceived usefulness of the feedback instrument to direct the student to relearn material not mastered and measures of relearning. Perceptions of the usefulness of the feedback instrument to direct the participant to relearn material not mastered were recoded into high ratings, as a value of 2, for students who responded with a score of 4 (*quite a bit*) or 5 (*very much*) and low ratings, as a value of 1, for students who responded with a score of 3(*somewhat*), 2 (*very little*), or 1 (*not at all*). Instances of relearning per unit of study were averaged into a single measure of instances of relearning per unit of study.

Mastery level analyses. The relearning measure was recoded as ‘2’ for above the median number of instances of relearning, (*Mdn* = 5.50 instances per unit) and recoded as ‘1’ for below the median number of instances of relearning. A Spearman Rho correlation coefficient was calculated for the relationship between IFM participants’ perceived usefulness of the analysis of mastery level per topic to direct the participant to relearn material not mastered and instances of relearning per unit over the research period. There is a strong direct correlation (*rho* =.533, n = 20, *p* = .02) between perceptions of the usefulness of the analysis of mastery level per topic to direct the student to relearn material not mastered and instances of relearning per unit. Students who rated the usefulness of analysis of mastery level per topic highly, as ‘very much’ or ‘quite a bit’, have a higher number of instances of relearning than students who rated the usefulness of the analysis of mastery level per topic as ‘somewhat’, ‘very little’, or ‘not at all’.

Progress graph. The relearning measure was recoded as ‘2’ for above median instances of relearning, (*Mdn* = 5.50 instances per unit for the IFM group, *Mdn* = 4.43

instances per unit for the IF group*)* and recoded as '1' for below median instances of relearning.

A Spearman Rho correlation coefficient was calculated to access the strength of the relationship between participant's perceptions of usefulness of the progress graph in directing the participant to relearn material not mastered and instances of relearning.

There was no significant correlation between perceptions of usefulness of the progress graph to direct the participant to relearn material not mastered and instances of relearning for the IFM group, *rho* = .203, n =17, p = .44, or for the IF group, *rho* = .101, n =20, p = .67. This result suggests no significant relationship between the two variables: perceptions of the usefulness of the progress graph to direct the student to relearn topics not mastered and instances of relearning for either group.

The perceived usefulness of the individualized feedback presented as an analysis of mastery level per topic and its strong association with measures of academic engagement and relearning supports Hypothesis 4b. The IFM group has stronger correlations between perceptions of usefulness of the feedback to guide study time and reported time engaged in academic activities, as well as, between perceptions of usefulness of the feedback to direct the student to relearn topics not mastered and measures of relearning, when compared to the IF group. These finding lend support for the inclusion of the additional information given to students with the analysis of mastery level per topic as a tool to direct study time and relearn material not mastered that is not evidenced with the progress graph as the feedback instrument.

Taken together both parts of this hypothesis lend some support for the usefulness of the mastery level analysis in conjunction with the progress graph to empower students to interpret and use the feedback information to guide time engaged in academic activities and to relearn material that has not been mastered.

Chapter 5

Discussion

This study was designed to examine the effectiveness of curriculum-based performance indicators as a tool to inform and direct the learning behaviors of first term engineering calculus students. This study addressed concerns that low levels of student achievement in math were due in part to ineffective studying, homework, and preparation for class. To investigate these concerns, I provided college students with key self-regulatory processes to become successful learners, such as advantageous goal setting, metacognitive monitoring, and adapting. More specifically, the present study used the CBM system to monitor students' achievement in a first course in engineering calculus, awarded partial credit to student responses, and allowed students to choose their educational course of action.

Hypotheses

Section I. This section discusses the impact of individualized performance feedback on measures of calibration accuracy, relearning, and academic performance when contrasted to corresponding measures from students who did not receive individualized performance feedback.

Hypothesis 1. This hypothesis posited that students who receive individualized performance feedback have more accurate measures of grade calibration when contrasted to students who do not receive individualized performance feedback on weekly Calculus-CBM quizzes. The results showed that students receiving individualized performance feedback had more accurate grade calibration than students not receiving performance feedback. Although this difference narrowly missed statistical significance, $p = .054$, it had an effect size of medium strength. Thus, when students were given individualized feedback on weekly Calculus-CBM quiz performance, they gained insight into their calculus capability as was demonstrated by their measures of calibration accuracy. This finding provides some support for this hypothesis.

Another notable finding was the greater monitoring accuracy of the IFM group compared to the no IF group. When students received the benefit of feedback as a progress graph supplemented with an analysis of mastery level per topic (IFM), they displayed significantly greater metacognitive awareness of the skills they bring to an exam than students who did not receive individualized performance feedback (no IF) on weekly quizzes. The magnitude of the effect size was medium. This result supports the learner-centered principle that successful learners have greater metacognitive awareness of how they think and learn than unsuccessful learners. With each additional CBM feedback component, the students increased the accuracy of their monitoring of calculus capability. The IFM group was 3.7 points more accurate in calibration than the IF group. The IFM group was 6 points more accurate in calibration than the no IF group, which is a significant difference.

The present study extended the research of Nietfeld et al. (2006) on calibration from the content area of educational psychology to the content area of calculus. The domain of mathematics is distinctive in that numerical rules and principles build closely upon each other, and high levels of precision are needed. More specifically, in calculus, the sequential integration of rules of differentiation and conceptual understanding is necessary for students to succeed in subsequent calculus courses. In the research of Nietfeld et al. (2006) and in the present study, students in treatment groups had more accurate measures of calibration than students in control groups. However, when examining the monitoring accuracy of the two feedback conditions in this study, there was no main effect for time. This result differs from that of Nietfeld et al. (2006) where groups improved calibration accuracy over time. The main effect for time of their research may be attributed to the explicit monitoring exercises that the students received. The monitoring exercises were designed to reinforce the connection between assessments of what one knows and monitoring accuracy. Students in the present study did not receive instruction on the association between feedback and monitoring accuracy.

Hypothesis 2. The second hypothesis concerned whether students who receive individualized performance feedback would have a higher rate of relearning topics when contrasted to students who do not receive individualized performance feedback on their weekly Calculus-CBM quiz. Differences in rate of relearning topics between the feedback groups and the no feedback group were statistically significant. The effect size of these differences was large. These findings lend support to the second hypothesis: namely, that the monitoring of performance can lead to adjustments in studying processes.

In "How People Learn," Bransford et al., (2000) noted that feedback is most valuable when students have an opportunity to correct their learning. Curriculum-based measurement practices afford students the opportunity to correct their learning by testing topics previously taught, currently taught, and topics yet to be encountered. In the present study, students who received individualized performance feedback corrected their learning by significantly relearning more topics than the no IF group.

Previous research by Stecker and Fuchs (2000) demonstrated that when skills analyses are part of CBM monitoring, teachers make more adjustments in instructional tactics than when given a monitoring graph or a static measure of student performance. Retesting topics does not guarantee that students will take the opportunity to relearn material. In the present research, the no IFM group relearned topics on average 37% of the time. When students were given either a visual representation of their progress or a visual representation supplemented with qualitative performance information, they relearned material on average 50% of the time. Monitoring ongoing performance appears to equip students with the necessary information to direct their learning behaviors.

Supplementary Analyses. A significant relationship between on-line monitoring of one's awareness of comprehension and task performance has been documented in both academic achievement and motoric skill. The supplementary analyses on differential performance on exams by groups over the research period revealed additional links between accurate monitoring judgments and performance.

The results of the present study reflect the positive relationship between monitoring accuracy and academic achievement. Nietfeld et al. (2006) noted that "monitoring ability in the form of calibration impacts not only performance on multiple-choice tests but also on more authentic, performance-based measures targeting the integration of course content" (pg. 174). In the present study, students who received individualized performance feedback scored significantly higher on exams than students who did not receive individualized performance feedback.

The ability to see small changes in performance over time is a key feature of curriculum-based measurement procedures. A greater awareness of one's learning behaviors by students in the feedback conditions was manifested as significantly higher rates of relearning and levels of performance compared to students in the no IF condition. This finding supports the research by Kitsantas and Zimmerman (2006). Their research revealed significant differences between students who graph their learning results and those who did not graph their ongoing progress. Students, who graphed ongoing performance, reported a greater awareness of their learning behaviors and performed at a higher skill level.

In summary, individualized performance feedback provided students with opportunities to self-assess, rate their self-efficacy in calculus, and make predictions regarding their performance on summative assessments. This section examined the effects of providing individualized performance feedback on weekly Calculus-CBM quizzes in terms of calibration accuracy, rate of relearning topics, and overall performance. When comparing the measures of students in the no IF group with those of the feedback groups, the powerful effect of performance feedback as a tool to inform and direct learning behaviors emerges. Students who received individualized performance feedback had more accurate calibration, relearned more material, and achieved at a higher level. Across the measures, the magnitude of the differences between the feedback groups and the no feedback group was an effect size of at least medium strength.

Section II. This section discusses the study habits and students' perceptions of the usefulness of feedback for the two groups that received individualized performance feedback. In particular, it discusses the impact of the supplementary mastery analysis (IFM) on students' time engaged in academic activities. This section also discusses the IFM group's perceptions of the usefulness of the feedback to inform and direct learning when compared to the corresponding measures of students who received feedback as a progress graph (IF).

Hypothesis 3. It was posited that students who receive individualized performance feedback on their weekly Calculus-CBM quizzes as a progress graph supplemented with an analysis of mastery level per topic would have greater learning efficiency in terms of engagement time in academic activities than students who received individualized performance feedback only as a progress graph. That is, IFM students would learn more in a shorter period of time than IF students.

The analyses showed a significant interaction between the unit of study and the feedback group in engagement time. The effect size of the interaction was medium. This result indicates that the impact of the type of feedback on academic engagement is influenced by the unit of study time. For the first unit of study, differences in academic engagement time between the IFM and IF groups approached statistical significance, $p = .057$. The IFM group was more efficient with their time engaged in academic activities than the IF group. There was also a significant main effect for time, with both feedback groups becoming more efficient in their engagement time in academic activities over the term. The effect size of this result is large. The form of the students' feedback explained an additional 27.6% of the variance in average time spent weekly engaged in academic activities over that explained by the interaction between the type of feedback and the unit of study. When students make informed decisions regarding what they know, they are better able to direct the time spent in academic activities on items that need their attention. This enables them to make gains in learning efficiency. The results suggest that both groups used the feedback to guide their study behaviors over time.

Although there is marginally significant support for differences between feedback conditions, IFM and IF, at the first unit of study, Hypothesis 3 was not fully confirmed.

There was evidence of differential self-monitoring activities reported by the students in the two feedback groups. The IFM group appears to be more effective in self-monitoring. Compared to the IF group, the IFM group reported studying topics not yet mastered significantly more often. The IFM group reported using all methods of study except reading the textbook to prepare for their weekly Calculus-CBM quiz. Students in the IFM group consistently reworked homework, reviewed notes, and used additional resources to prepare for the weekly Calculus-CBM quiz. There are no significant correlations between methods of study chosen by the IF group.

The self-monitoring protocol was based on the research of Lan (1996) who studied the effects of self-monitoring on performance and self-regulated learning strategies. The findings of the present study lend additional support to Lan's results. He found that when students were involved in a self-monitoring process, they were more alert to the effectiveness of their learning strategies. In the present study, IFM students report more instances of studying topics not mastered and have a consistent pattern of study behaviors.

To summarize, when students are cognizant of the details of their performance, they attend to academic activities more purposefully. Students in the IFM group appear to selectively address shortcomings in topic mastery. The initial differences between feedback groups in time engaged in academic activities suggest the study regimen of the IFM group was more effectively directed. The time engaged in academic activities by the IF group over the last two units of study is not statistically different from that of the

IFM group. This suggests that over time the IF group appears learns to use the progress graph to direct academic activities.

Hypothesis 4. The first part of this hypothesis concerns whether perceptions of the usefulness of the feedback differ by group (that is, IFM versus IF). It was hypothesized that students in the IFM condition would perceive the usefulness of the feedback more positively than students in the IF condition. There were three areas to rate the usefulness of the feedback: to guide study time, to direct the student to relearn topics not mastered, and to judge which topics the student could answer correctly. This hypothesis was supported only for the first area: usefulness of the feedback to guide study time on individual topics. There were significantly more positive expressions reported by the IFM group than the IF group. The magnitude of the differences has a medium size effect.

The second part of this hypothesis examines the strength of the associations between perceptions of usefulness and corresponding measures, for each type of feedback. Perceptions of the usefulness of the feedback to direct study time and measures of time engaged in academic activities was the first area of inquiry. This hypothesis is supported by a strong inverse association between the usefulness of the mastery level analysis to guide study time and the measures of study time. This suggests that those who rated the usefulness of the mastery analysis highly were able to learn efficiently using less time.

The hypothesis additionally examined the association between perceptions of the usefulness of the feedback to relearn material and measures of relearning. There was a direct relationship between perceived usefulness of the mastery level analysis to relearn material not mastered and instances of relearning. This finding implies the addition of the topic mastery analyses seems to keep students informed of their progress and motivates them to be in control of their learning behaviors. This result supports Schunk's (1991) conclusion that attainment of a goal is linked with a greater sense of self-efficacy; heightened self-efficacy sustains motivation to attain goals and improves skill development. The weekly changes in measures of topic mastery returned sufficient information to the IFM student to make learning choices.

To summarize, differences between feedback groups in perceptions of the usefulness of the feedback tool to guide study time emerged. Students in the IFM group, who reported more positive expressions of usefulness of the feedback to direct study time, displayed a consistent pattern of time spent academically engaged whereas students in the IF group changed their study time behavior during the research period. The results also showed that when students rate the usefulness of the mastery feedback highly, they appear to use the feedback to guide their study time and to relearn material. The mastery level analysis informed and directed the learning behaviors of the IFM group. This provides some support for Hypothesis 4.

Limitations

The specialized mission of this college within the maritime field may limit generalizations of these results beyond similar institutions. First year engineering students carry a full course load and are often challenged by the demands placed on them academically. The pressure to study among these students to meet course competencies to continue in the engineering program may be great. Academic benefits from higher quality feedback may be diminished for this group of participants as compared to those benefits derived from prior research in self-monitoring with psychology or education majors.

Furthermore, students in the present study were asked weekly to rate their self-confidence to use the feedback to improve academic performance and to reach a course grade goal. In both areas, the majority of students selected the four annotated values of a continuous 1 to 10 scale, rather than the full range of scale values. Subtle weekly changes in perception of one's ability to use the feedback were lost. Any future use should include fully annotated self-confidence scales.

The progress graph is a visual stimulus for students to quickly judge the impact of a single quiz on their overall performance. The student can immediately note the difference between the trendline in blue and the aimline in red. In order to transfer the visual benefits of the graph to the mastery analysis, future mastery level designations should be color coded as red for not mastered, orange for partially mastered, and green for mastered. Development of a dashboard for the coding would provide a concise visual stimulus.

The research was carried out in an authentic environment in that the feedback procedures were integrated into regular classroom practices. This can be a limitation in that some students may be unwilling to change their methods of studying if they have been effective for them. To derive meaning from the feedback, students need to feel a sense of ownership of it. This is an important consideration for future research.

Educational Implications

The findings regarding the effectiveness of curriculum-based measurement performance indicators as a tool to inform and direct the learning behaviors of first term engineering calculus students are encouraging. This study represents the first time CBM procedures were applied in an engineering calculus course. In the areas of calibration accuracy, rate of relearning, and achievement, the differences between the performance feedback and the no performance feedback groups were either statistically significant or closely approached statistical significance. The effect sizes for these differences were at least medium strength.

In the present study, no additional training on how to use the feedback was given to students that would have influenced the role of the performance feedback for one group over the other groups. Teachers, adopting curriculum-based measurement performance indictors, can maximize their effectiveness by reinforcing the connections between performance feedback, relearning, and achievement with their students. Students should take a more active role in utilizing their performance feedback. Students, who graph their ongoing progress, would benefit from a greater sense of control over their learning as changes in their trendline are revealed.

One of the strengths of students' self-monitoring of study habits and self-judgments of calculus capability was the regularity of these measures and their proximity to actual testing. These monitoring measures were given in real time immediately preceding a calculus assessment. Individualized performance feedback was returned to students at the following class meeting. This cycle can be motivating for students in that it links learning behaviors with changes in performance.

The analyses of study methods revealed that students did not use the textbook as a resource. This is a concern across many campuses. In order to determine the underlying factors involved with book ownership and use, my institution has included two additional questions on its most recent administration of the Student Opinion Survey regarding percent of books purchased and reasons for not purchasing books. The information gathered from these questions may guide decisions on using e-books or the frequency of changing editions of currently used textbooks.

Help-seeking should be included as a method of study in any future use of the self-monitoring form. This is an area where teachers can be proactive and provide students with guidance to become help-seekers. Tutors in learning assistance centers can be trained to help students ask questions that tap into what the student knows rather than providing answers on an assignment.

For educators, the use of CBM monitoring practices is a way to provide personalized feedback to students. For the teacher, the wealth of information that is returned from the skills analyses can refine instruction and close the feedback loop. For the student, self-assessments of progress can facilitate skill appraisal and self-directed learning. This study demonstrates the possibility of extending CBM practices to higher level courses. CBM would be a natural fit for licensure preparation, where items are drawn from the content area. More importantly, feedback can specifically address areas not mastered.

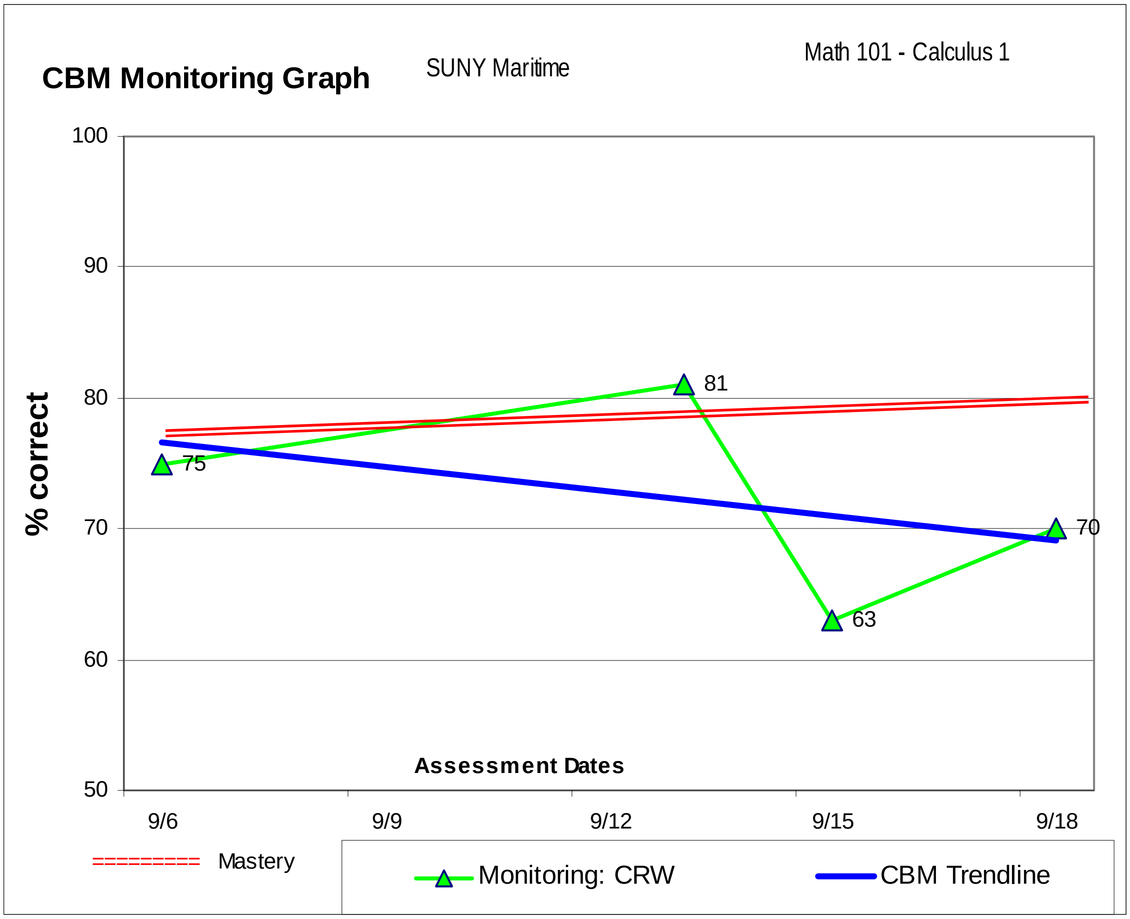
Conclusions

When academic feedback is presented visually as a progress graph, it can convey an overall trend in learning. When a student is not successful on a CBM assessment, his or her graph will reveal a drop in the trendline; but how should the student modify his or her approach to learning? The present research showed that shifts in students' trendlines can be analyzed in terms of academic topic. Such results enable learners to differentiate between topics that have not been mastered, those that have been partially mastered, or those that have been mastered. This high quality of individualized feedback can enhance students' ability to more effectively regulate their personal efforts to learn.

Appendix A

Performance Feedback Indicators

Progress Graph (IF and IFM Groups)



The red line is the grade goal. I've set this sample at a grade of B-, since that is an indication of topic mastery.

The green line is frequency graph of quiz scores.

The blue line is called a trend line. On this sample the trendline is sloping downward, because the grades are decreasing. **The trendline is not an average**. You can see this with some grades above and some grades below the trendline.

The important thing to notice is the gap between the grade goal and the trendline. The sample shows that if this student continues with this pattern of grades, the student will fall from the grade goal. The student can use this information to motivate him/herself to spend more time on certain topics.

Supplementary Qualitative Feedback for IFM Group:

TOPIC	Accuracy Quiz 1	Accuracy Quiz 2	Accuracy Quiz 3	Mastery Level		
3.3 Applications; higher order derivatives	83	64		Mastered	Partial Mastery	
3.4 Chain Rule for algebraic functions	78	60		Partial Mastery	Partial Mastery	
3.4 Chain Rule for non-algebraic functions		74	81		Partial Mastery	Mastered
3.5 Implicit Differentiation		59	68		Not Mastered	Partial Mastery
3.7 Related Rates			42			Not Mastered
3.6 Derivatives of Inverse Trig Functions			85			Mastered

Topic mastery is explained in terms of points earned on topic.

Level	Points Earned	Analysis
• Mastery:	At least 80% correct	Minor error in calculations, a misstatatement, or an almost complete response with a correct answer
• Partial Mastery:	At least 60% to 79% correct	The solution contains some correct aspects; but also exhibits conceptual flaw(s).
• Non-mastery:	Less than 60% correct.	Incomplete work, major error in concept

Although the trendline demonstrates that the student performance is moving away from the grade goal of mastery, the analysis of mastery level by topic provided information on which topics the student needs to relearn to achieve a level of mastery.

Appendix B

Letter of Consent

Ph.D. Program in Educational Psychology

The Graduate School and University Center
The City University of New York
365 Fifth Avenue
New York, NY 10016-4309
TEL 212.817.8285 FAX 212.817.1516

Consent Form

My name is Linda Sturges and I am a doctoral student in the Educational Psychology Ph.D. Program at The Graduate School and University Center of the City University of New York (CUNY) and Principal Investigator of a project entitled "Curriculum-Based Measurement Performance Indicators." This is a research study on an assessment approach, called Curriculum-Based Measurement and is expected to examine the effects of performance feedback on student learning and studying. I would like your permission to analyze your class information, as described below, after the course has been completed and grades have been submitted.

Participation in this study involves your consent to release the following class information after grades have been submitted:

1. **self-monitoring of study time**
2. **calibration scores**
3. **weekly Curriculum-Based Measurement scores**
4. **exit questionnaire**
5. **course grade information.**

This class information will be analyzed under a randomly assigned code number without any links to your name.

This consent form will be collected in a sealed envelope and filed under a student ID number. All information gathered will be kept strictly confidential and will be stored in a secured area, to which only I, and my advisor, will have access.

At any time you may withdraw from the study without any penalty. Your decision, to participate or not to participate in this research will not affect your grade in any manner. There is no obligation for you to participate in this research and you may withdraw or change your decision to participate at any time during the semester.

Should you wish to withdraw or change your decision, you simply give your revised consent form to the Science department secretary, who will give it to me after all grades have been submitted.

After all grades have been submitted, I will open the 'consent form' envelopes and recode, for data analysis purposes, the ID number of any student who has elected to participate in this research. Neither your name nor any identifying characteristics will be used in the data analysis.

The risk of participating in this research should not exceed that encountered in daily life. The information you release will be written up, without any direct reference to you as an individual, in a paper on curriculum-based measurement. The benefits of your participation will extend the generalized knowledge of curriculum-based measurement to a college setting. There will be approximately 40 participants taking part in this study.

I may publish the results of the study, but neither your name nor any identifying characteristics will be used in the publication. If you would like a copy of the study, please provide me with your address and I will send you a copy in the future.

If you have any questions about this research, you can contact me at 718 409-7382 or Lsturges@sunymaritime.edu or my advisor Professor Zimmerman at (212) 817-8291 or BZimmerman@gc.cuny.edu . If you have questions about your rights as a participant in this study, you can contact Kay Powell, IRB Administrator, The Graduate Center/City University of New York, (212) 817-7525 or kpowell@gc.cuny.edu

Thank you for your participation in this research. Please retain the second copy of this form for your records.

I agree to participate in this research. (please circle one): Yes No

____________________ ____________________

Student's Name (Print) **Student's signature** **Date**

____________________ ____________________

Student ID# **Investigator's signature** **Date**

Graduate School & University Center / CUNY
IRB
Approved
07-12-1414
Approved: Expired:
1/24/08 To 1/23/09

Appendix C

Self Monitoring Protocol Unit 2

Name ______________________ ID# ___________

Math 101 Unit 2 Basic Differentiation Techniques

Directions: Please indicate the topics you worked on in preparation for the next quiz. Second, please specify how you prepared for the quiz. Lastly, record the time you spent preparing for this course.

Chapter Section	Concept	Check Topics Studied	Review notes	Redo HW	Read text	End of chapter or online problems; DVDs
3.1	The Tangent Line Problem: • Derivative as a slope					
3.1	The Tangent Line Problem: • Differentiability at a point.					
3.1	The derivative of a function; • Relationship between a function and its derivative.					
3.2	Rules: • Power rule for polynomials					
3.2	Rules: • Trig functions & exponential functions					
3.2	Application: • rates of change					
3.3	Rules: • Product • Quotient					

This week, how much time would you estimate that you spent engaged in academic activities such as, doing homework, reviewing class notes, and studying in preparation for calculus?

Mon ____ Tues. _____ Wed. ____ Thurs. ______ weekend ______

Total time: ________________________

How confident are you in your ability to use the performance feedback you have received to improve your academic performance? ________________

[not very confident = 1 // somewhat confident = 4 // confident = 7 // very confident = 10]

How confident are you in your ability to use the performance feedback you have received to reach your course grade goal? ___________________

[not very confident = 1 // somewhat confident = 4 // confident = 7 // very confident = 10]

Appendix D

Self Monitoring Protocol Unit3

Name ______________________________ ID# ___________

Math 101 Unit 3 Advanced Differentiation Techniques

Directions: Please indicate the topics you worked on in preparation for the next quiz. Second, please specify how you prepared for the quiz. Lastly, record the time you spent preparing for this course.

Chapter Section	Concept	Check Topics Studied	Review notes	Redo HW	Read text	End of chapter or online problems; DVDs.
3.3	Applications of the product and quotient rules					
3.3	Higher order derivatives					
3.4	Chain Rule: • Algebraic functions					
3.4	Chain Rule: • Non-algebraic functions					
3.5	Implicit differentiation					
3.7	Related Rates					
3.6	Derivatives of Inverse Trig Functions					

This week, how much time would you estimate that you spent engaged in academic activities such as, doing homework, reviewing class notes, and studying in preparation for calculus?

Mon ____ Tues. _____ Wed. ____ Thurs. ______ weekend ______

Total time: _________________________

How confident are you in your ability to use the performance feedback you have received to improve your academic performance? ________________
[not very confident = 1 // somewhat confident = 4 // confident = 7 // very confident = 10]

How confident are you in your ability to use the performance feedback you have received to reach your course grade goal? ___________________
[not very confident = 1 // somewhat confident = 4 // confident = 7 // very confident = 10]

Appendix E

Self Monitoring Protocol Unit 4

Name ________________________________ ID# ___________

Math 101 Unit 4: Applications of differentiation

Directions: Please indicate the topics you worked on in preparation for the next quiz. Second, please specify how you prepared for the quiz. Lastly, record the time you spent preparing for this course.

Chapter Section	Concept	Check Topics Studied	Review notes	Redo HW	Read text	End of chapter or online problems; DVDs.
4.1	Extrema: • On a closed interval					
4.1	Extrema: • On an open interval					
4.3	Increasing & Decreasing Functions					
4.3	1st Derivative Test for Extrema					
4.4	Concavity					
4.4	2nd Derivative Test for Extrema					
4.5	Limits at Infinity					
4.6	Curve sketching					
4.7	Optimization					

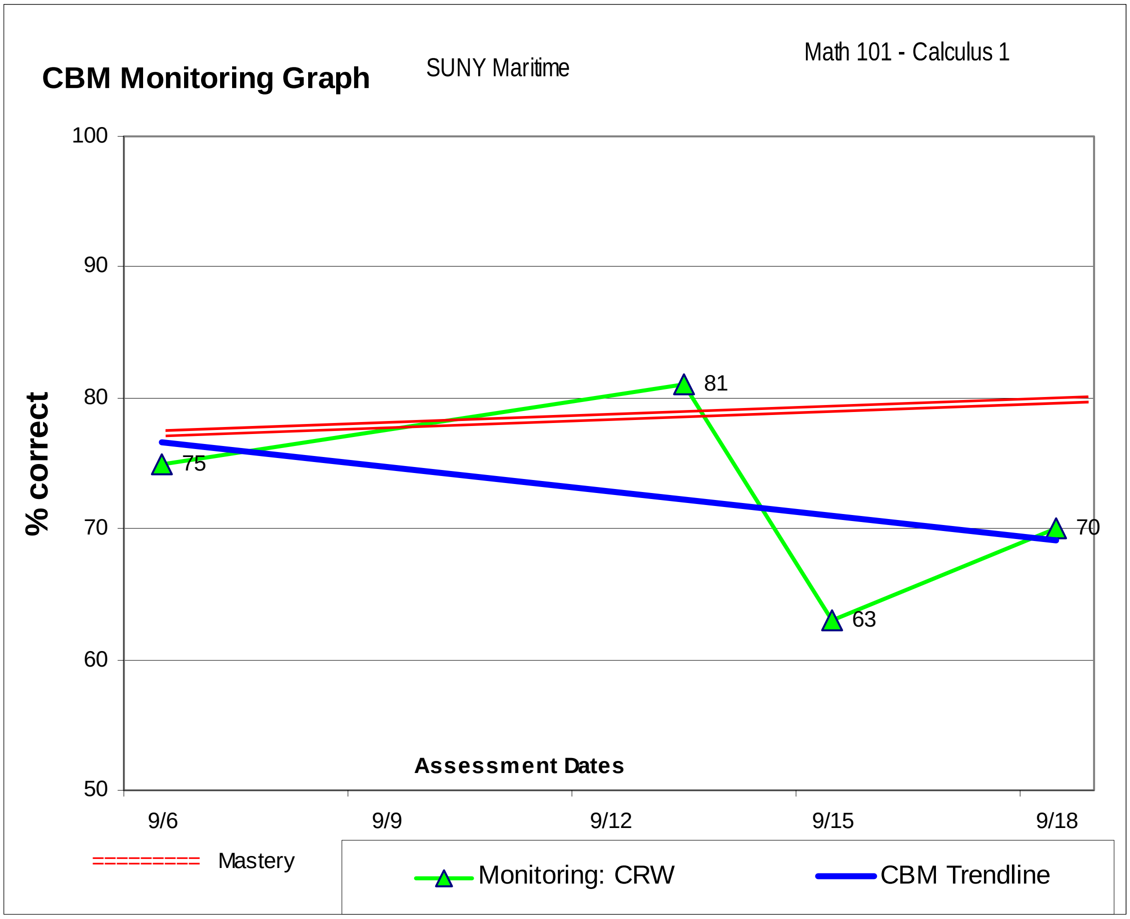
This week, how much time would you estimate that you spent engaged in academic activities such as, doing homework, reviewing class notes, and studying in preparation for calculus?
Mon ____ Tues. _____ Wed. ____ Thurs. ______ weekend ______

Total time: _________________________

How confident are you in your ability to use the performance feedback you have received to improve your academic performance? ________________
[not very confident = 1 // somewhat confident = 4 // confident = 7 // very confident = 10]

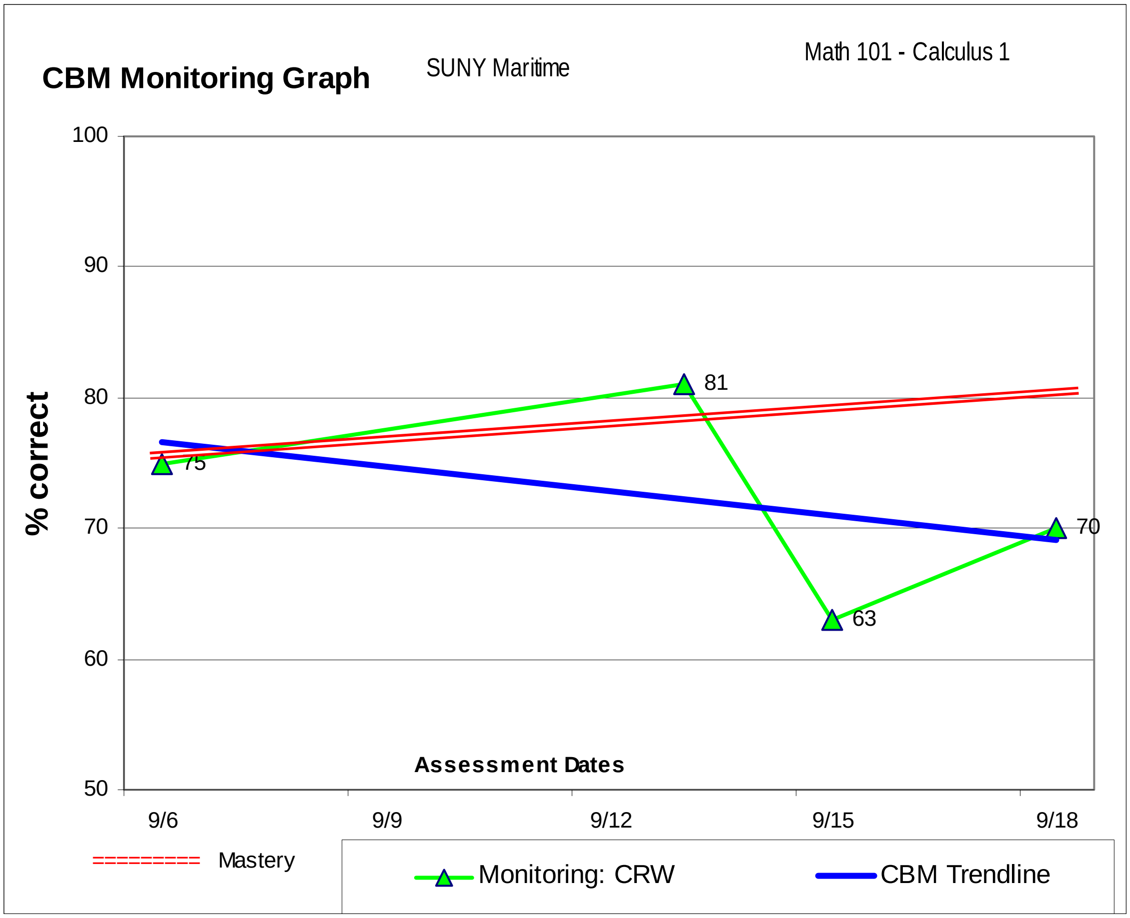
How confident are you in your ability to use the performance feedback you have received to reach your course grade goal? ___________________
[not very confident = 1 // somewhat confident = 4 // confident = 7 // very confident = 10]

Appendix F
Feedback for the IF Group



Appendix G

Feedback for IFM Group



TOPIC	Accuracy			Mastery Level		
	Quiz 1	Quiz 2	Quiz 3			
3.3 Applications; higher order derivatives	83	64		Mastered	Partial Mastery	
3.4 Chain Rule for algebraic functions	78	60		Partial Mastery	Partial Mastery	
3.4 Chain Rule for non-algebraic functions		74	81		Partial Mastery	Mastered
3.5 Implicit Differentiation		59	68		Not Mastered	Partial Mastery
3.7 Related Rates			42			Not Mastered
3.6 Derivatives of Inverse Trig Functions			85			Mastered

Appendix H

Self-Efficacy Form Unit 2

Name ________________________________ ID# ____________

Math 101 Unit 2: Basic Differentiation Techniques

Directions: Please record your self-efficacy or how confident you are in your ability to respond correctly to problems from each concept for this exam.

Use the continuous scale from 1 to 10 as described below:

1______________________4______________________7________________10

[not very confident // somewhat confident // confident // very confident]

Chapter.Section	Concept	Self-Efficacy
3.1	The Tangent Line Problem: • Derivative as a slope	
3.1	• Differentiability at a point.	
3.1	• Relationship between a function and its derivative.	
3.2	Rules: • Power rule for polynomials	
3.2	Rules: • Trig functions & exponential functions	
3.2	Application: • rates of change	
3.3	Rules: • Product • Quotient	

Use these rating to predict your grade on this exam and your level of confidence to earn the predicted score

Predict your score on exam 2 ______________________________________

How confident are you in your prediction *(Circle your choice)*:

Not very confident Confident Very Confident

What is your course grade goal, at this time? *(Circle your choice)*:

A (≥ 95%) A- (90 – 94) B+ (87-89) B (83-86) B- (80-82)

C+ (77-79) C (73-76) *C- (70-72) **required to continue into Calculus 2*

Appendix I

Self-Efficacy Form Unit 3

Name ________________________________ ID# ____________

Math 101 Unit 3 Advanced Differentiation Techniques

Directions: Please record your self-efficacy or how confident you are in your ability to successfully complete problems from each concept for this exam.

Use the continuous scale from 1 to 10 as described below:

1______________________4______________________7________________10

[not very confident // somewhat confident // confident // very confident]

Chapter.Section	Concept	Self-Efficacy
3.3	Applications of the derivative	
3.3	Higher order derivatives	
3.4	Chain Rule: • Algebraic functions	
3.4	Chain Rule: • Non-algebraic functions	
3.5	Implicit differentiation	
3.7	Related Rates	
3.6	Derivatives of Inverse Trig Functions	

Use these rating to predict your grade for this exam and your level of confidence to earn the predicted score.

Predict your score on exam 3 ______________________________________

How confident are you in your prediction *(Circle your choice)*:

Not very confident Confident Very Confident

What is your course grade goal, at this time? *(Circle your choice)*:

A (≥ 95%) A- (90 – 94) B+ (87-89) B (83-86) B- (80-82)

C+ (77-79) C (73-76) *C- (70-72) **required to continue into Calculus 2*

Appendix J

Self-Efficacy Form Unit 4

Name ______________________________ ID# ___________

Math 101 Exam 4: Applications of the derivative

Directions: Please record your self-efficacy (or how confident you are) in your ability to complete problems from each concept for this exam.

Use the continuous scale from 1 to 10 as described below:

1____________________4____________________7_______________10

[not very confident // somewhat confident // confident // very confident]

Chapter.Section	Concept	Self-Efficacy
4.1	Extrema: • On a closed interval	
4.1	Extrema: On an open interval	
4.3	Increasing & Decreasing Functions	
4.3	1st Derivative Test for Relative Extrema	
4.4	Concavity	
4.4	2nd Derivative Test for Extrema	
4.5	Limits at Infinity	
4.6	Curve sketching	
4.7	Optimization	

Use these rating to predict your grade on this exam and your level of confidence to earn the predicted score.

Predict your score on exam 4 ___________________________________

How confident are you in your prediction *(Circle your choice)*:

Not very confident Confident Very Confident

What is your course grade goal, at this time? *(Circle your choice)*:

A (≥ 95%) A- (90 – 94) B+ (87-89) B (83-86) B- (80-82)

C+ (77-79) C (73-76) *C- (70-72) **required to continue into Calculus 2*

Appendix K

Exit Questionnaire IF Group

Exit Questionnaire Form A ID #__________

Predict your course grade *(Circle)*.

A A- B+ B B- C+ C C- < C-

How certain are you of this grade:

☐ Not Sure ☐ Quite Sure ☐ Absolutely Sure

How useful was the feedback you received on your weekly quiz to help you reach your grade goal?

Very much Quite a bit Somewhat Very little Not at all

How useful was the feedback you received on your weekly quiz to help you improve your academic performance?

Very much Quite a bit Somewhat Very little Not at all

Please rate the following questions concerning the utility of the **graphed performance chart**. *(Circle your choice)*

How useful was the **graphed performance chart** in guiding your study time on individual unit topics?

Very much Quite a bit Somewhat Very little Not at all

How useful was the **graphed performance chart** in directing you to relearn material that you had not mastered?

Very much Quite a bit Somewhat Very little Not at all

How useful was the **graphed performance chart** in helping you judge what topic questions you could answer correctly?

Very much Quite a bit Somewhat Very little Not at all

Appendix L

Exit Questionnaire IFM Group

Exit Questionnaire Form B ID #__________

Predict your course grade *(Circle)*.

A A- B+ B B- C+ C C- < C-

How certain are you of this grade: ☐ Not Sure ☐ Quite Sure ☐ Absolutely Sure

---

How useful was the feedback you received on your weekly quiz to help you reach you grade goal?

Very much **Quite a bit** **Somewhat** **Very little** **Not at all**

How useful was the feedback you received on your weekly quiz to help you improve your academic performance?

Very much **Quite a bit** **Somewhat** **Very little** **Not at all**

Please rate the following questions concerning the **utility of the graphed performance chart.** *(Circle your choice)*

How useful was the **graphed performance chart** in guiding your study time on individual unit topics?

Very much **Quite a bit** **Somewhat** **Very little** **Not at all**

How useful was the **graphed performance chart** in directing you to relearn material that you had not mastered?

Very much **Quite a bit** **Somewhat** **Very little** **Not at all**

How useful was the **graphed performance chart** in helping you judge what topic questions you could answer correctly?

Very much **Quite a bit** **Somewhat** **Very little** **Not at all**

Please rate the following questions concerning the **utility of the concept mastery analysis.** *(Circle your choice)*

How useful was the **concept mastery analysis** in guiding your study time on individual unit topics?

Very much **Quite a bit** **Somewhat** **Very little** **Not at all**

How useful was the **concept mastery analysis** in directing you to relearn material that you had not mastered?

Very much **Quite a bit** **Somewhat** **Very little** **Not at all**

How useful was the **concept mastery analysis** in helping you judge what topic questions you could answer correctly?

Very much **Quite a bit** **Somewhat** **Very little** **Not at all**

References

Akst, G., Crook, D. Moreno, V., Littman, C., Grima, A. (2006) Performance in selected mathematics courses at the City University of New York: Implications for retention. http://www1.cuny.edu/academics/oaa/uei/cue/resources/OIRA.MATH.STUDY.FINAL.2006-03-06_PAGENUM.pdf

American Psychological Association (1997) Learner-Centered Psychological Principles: A Framework for School Reform and Redesign. http://www.apa.org/ed/cpse/LCPP.pdf

Batha, K., & Carroll, M. (2007). Metacognitive training aids decision making. *Australian Journal of Psychology, 59*(2), 64-69. doi:10.1080/00049530601148371.

Bransford, J., Brown, A., & Cocking, R. (Eds). (2000) How People Learn: Brain, Mind, Experience, and School. National Academy Press: Washington, D. C. Available: http://books.nap.edu/html/howpeople1/.

Cusumano, D. (2007). Is it working? An overview of curriculum-based measurement and its uses for assessing instructional, intervention, or program effectiveness. *The Behavior Analyst Today, 8*(1), 24-34. Retrieved from http://www.behavior-analyst-today.com/

de Ramirez, R. & Shapiro, E. (2006). Curriculum-based measurement and the evaluation of reading skills of Spanish-speaking English language learners in a bilingual education classrooms. *School Psychology Review, 35*(3), 356-369. Retrieved from http://www.nasponline.org/publications/spr/sprmain.aspx

Deno. S. (1992). The nature and development of curriculum-based measurement. *Preventing School Failure, 36*(2), 5. Retrieved from http://www.heldref.org/pubs/psf/about.html

Engaged Learning: Fostering Success for all Students. Annual Report 2008. http://nsse.iub.edu/NSSE_2008_Results/index.cfm

Espin, C & Tindal, G. (1998). Curriculum-based measurement for secondary students. In M. Shinn (ed.) *Advanced Applications of Curriculum-Based Measurement* (pp. 214-253). The Guilford Press: New York.

Espin, C. A., Shin, J., & Busch, T. W. (2005). Curriculum-Based Measurement in the Content Areas: Vocabulary Matching as an Indicator of Progress in Social Studies Learning. *Journal of Learning Disabilities, 38*(4), 353-363. Retrieved from http://ldx.sagepub.com/archive/

Espin, C., Scierka, B., & Skare, S. (1999). Criterion-related validity of curriculum-based measures in writing for secondary school students. *Reading and Writing Quarterly, 14*, 5-27. Retrieved from http://www.ingentaconnect.com/content/1057-3569;jsessionid=5qdu98kbj2895.alexandra

Everson, H. and Tobias, S. (2001). The ability to estimate knowledge and performance in college: A metacognitive Analysis. In H. Hartman, Ed. *Metacognition in Learning and Instruction: Theory, Research, and Practice* (pp.69-83). Kluwer Academic Publishers: Netherlands.

Foegen, A. Project AAIMS (Algebra Assessment and Instruction—Meeting Standards). http://www.ci.hs.iastate.edu/aaims/

Foegen, A., Jiban, C., & Deno, S. (2007). Progress monitoring measures in mathematics: A review of the literature. *Journal of Special Education, 41*(2), 121-139. Retrieved from http://sed.sagepub.com/

Foegen, A., Olson, J., & Impecoven-Lind, L. (2008). Developing progress monitoring measures for secondary mathematics: An illustration in algebra. *Assessment for Effective Intervention, 33*(4), 240 - 249. Retrieved from http://aei.sagepub.com.ezproxy.gc.cuny.edu/content/vol33/issue4/

Foegen, A., Olson, J., & Edwards, M. J. (April 2006). Monitoring student progress in algebra: An new approach to assessment. Session presented at the Annual Meeting of the National Council of Teachers of Mathematics, St. Louis, MO.

Fuchs, L. (2004). The Past, present, and future of curriculum-based measurement research. *School Psychology Review, 33*(2), 188-192. Retrieved from http://www.nasponline.org/publications/spr/sprmain.aspx

Fuchs, L. S. (1998). Computer applications to address implementation difficulties associated with curriculum-based measurement. In M. Shinn (ed.) *Advanced Applications of Curriculum-Based Measurement* (pp. 89-112). The Guilford Press: New York.

Fuchs, L., & Fuchs, D. (1990). The role of skills analysis in curriculum-based measurement in math. *School Psychology Review*, *19*(1), 6-22. Retrieved from http://www.nasponline.org/publications/spr/sprmain.aspx

Gall, M., Gall, J. Borg, W. (2003). Educational Research: An introduction. Allyn and Bacon: New York.

Good III, R. & Jefferson, G. (1998). Contemporary perspectives on curriculum-based measurement validity. In M. Shinn (Ed.) *Advanced Applications of Curriculum-Based Measurement* (pp. 61 – 88). The Guilford Press: New York.

Hacker, D., Bol, L., Horgan, D., and Rakow, E. (2000). *Test prediction and performance in classroom context.* Journal of Educational Psychology, *92*(1), 160-170. Retrieved from http://www.apa.org/journals/edu/

Helwig, R., Anderson, L., & Tindal, G. (2002, Summer). Using a Concept-Grounded Curriculum-Based Measure in Mathematics to Predict Statewide Test Scores for Middle School Students with LD. *Journal of Special Education*, *36*(2), 102. Retrieved from http://sed.sagepub.com/

Kitsantas, A., & Zimmerman, B. J. (2006). Enhancing self-regulation of practice: the influence of graphing and self-evaluative standards. *Metacognition and Learning, 1*(3), 201-212. doi:10.1007s/11409-006-9000-7.

Kuenzi, J., Matthews, C., and Mangan, B. (2006) CRS Reports: Science, Technology, Engineering, and Mathematics (STEM) Education Issues and Legislative Options. http://www.ncseonline.org/NLE/CRS/abstract.cfm?NLEid=1666

Lan, W. (1996). The effects of self-monitoring on students' course performance, use of learning strategies, attitude, self-judgment ability, and knowledge representation.

Journal of Experimental Education, 64(2), 101. Retrieved from http://www.heldref.org/pubs/jxe/about.html

Larson, W. C. & Ward, C. (2006). Developing a predictor of university students' course grades using curriculum-based measurement: An initial investigation. *Journal of College Reading and Learning, 37*(1), 45-59. Retrieved from http://findarticles.com/p/articles/mi_hb3247/is_1_37/?tag=content;col1

Lembke, E. S., & Espin, C. A. (2005). Curriculum-based measurement: using progress monitoring to influence instruction. In G. Sideridid, G. & Citro, T. (Eds.) *Research to Practice: Effective Interventions in Learning Disabilities* (pp. 144-167). Learning Disabilities Worldwide: MA.

Liu, A,.Sharkness, J & Pryor, J. H. (2008). *Findings from the 2007 Administration of Your First College Year (YFCY): National Aggregates* Los Angeles: Higher Education Research Institute, UCLA. http://www.gseis.ucla.edu/heri/publications-brp.php

Middle States Commission on Higher Education. (2007), Student Learning Assessment: Options and Resources. http://www.msche.org/publications/SLABook07070925104757.pdf

Nietfeld, J. L., Cao, L., & Osborne, J. W. (2006). The effect of distributed monitoring exercises and feedback on performance, monitoring accuracy, and self-efficacy. *Metacognition and Learning, 1*(2), 159-179. doi:10.1007/s10404-006-9595-6.

Nietfeld, J.& Cao, L., Osborne, J. (2005). Metacognitive monitoring accuracy and student performance in the postsecondary classroom. *The Journal of Experimental*

Education, 74(1), 7-28. Retrieved from http://www.heldref.org/pubs/jxe/about.html

Perkmen, S., Foegen, A., & Olsen, J. (2006b). *Reliability, criterion validity, and sensitivity to growth: Extending work on two algebra progress monitoring measures.* (Technical Report 12). Project AAIMS, Department of Curriculum and Instruction, Iowa State University, Ames, IA. Available: http://www.ci.hs.iastate.edu/aaima/Technical.html.

Perkmen, S., Foegen, A., & Olsen, J. (2006c). *A replication study of the reliability, criterion validity, and sensitivity to growth of two algebra progress monitoring measures.* (Technical Report 13). Project AAIMS, Department of Curriculum and Instruction, Iowa State University, Ames, IA. Available: http://www.ci.hs.iastate.edu/aaima/Technical.html.

Perkmen, S., Foegen, A., & Olsen, J. (2006d). *Technical characteristics of two algebra progress monitoring measures: Reliability, criterion validity, and sensitivity to growth.* (Technical Report 14). Project AAIMS, Department of Curriculum and Instruction, Iowa State University, Ames, IA. Available: http://www.ci.hs.iastate.edu/aaima/Technical.html.

Schraw, G. (2008, March). *Measuring Metacognitive Judgments*. Paper presented at the meeting of the AERA Annual Conference, New York, NY.

Schunk, D. (1991). *Self-efficacy and academic motivation.* Educational Psychologist, *26* (3 & 4), 207-231. Retrieved from http://www.apa.org/about/division/div15.html

Shinn, M. & Bamonto, S. (1998). Advanced applications of curriculum-based measurement: Big-ideas and avoiding confusion. In M. Shinn (Ed.) *Advanced Applications of Curriculum-Based Measurement.* The Guilford Press: New York.

Stecker, P. & Fuchs, L.S. (2000). Effecting superior achievement using curriculum-based measurement: The importance of individual progress monitoring. *Learning Disabilities Research & Practice, 15*(3), 128-138. Retrieved from http://www.wiley.com/bw/journal.asp?ref=0938-8982

Stecker, P., Fuchs, L. and Fuchs, D. (2005). Using curriculum-based measurement to improve student achievement: Review of the research. *Psychology in the Schools*, *42*(8), 795– 819. doi:10 1002/pits.20113.

State University of New York. (2005). Standards and Rubrics for Assessing General Education in Mathematics. http://www.suny.edu/sunypp/docs/628.pdf

The Nation's Report Card: 12 Grade Reading and Mathematics. http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2007468

Tindal, G.& Nolet, V. (1995). Curriculum-based measurement in middle and high schools: Critical thinking skills in content areas. *Focus on Exceptional Children, 27*(7), 1. Retrieved from http://www.lovepublishing.com/catalog/focus_on_exceptional_children_31.html

U. S. Department of Education. (2006). *A Test of Leadership: Charting the Future of U.S. Higher Education. Washington D. C.* http://www.ed.gov/about/bdscomm/list/hiedfuture/reports/pre-pub-report.pdf

Wright, J. CBM Manual for Teachers. www.interventioncentral.org/pdfdocs/cbaManual.pdf

Zimmerman, B. & Paulsen, A. (1995). Self-monitoring during collegiate study: An invaluable tool for academic self-regulation. *New Directions for Teaching and Learning, 63*, 13-27. Retrieved from http://www.josseybass.com/WileyCDA/WileyTitle/productCd-TL.html

Zimmerman, B. J. (2000). Attainment of self-regulation: A social cognitive perspective. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Self-regulation: Theory, research, and applications* (pp.13-39). Orlando, FL: Academic Press.