A comparison of former SIMMS and non-SIMMS students on three college-related measures
by Michael Allen Lundin

A dissertation submitted in partial fulfillment of Doctor of Education in Education
Montana State University

Abstract:
This study compared three college-related measures of former high school students from a reform-based mathematics curriculum, produced by the Systemic Initiative for Montana Mathematics and Science (SIMMS) Project, with those of students from more traditional mathematics curricula. The measures of interest were ACT and SAT scores, freshman mathematics grades, and initial choices of majors. The subjects entered one of two state universities in Montana in Fall 1998. An Instructor Survey of classroom practices measured the intensity of reform in college freshman mathematics classrooms, and Supervisor Interviews produced qualitative data as background.

Results from the study included evidence of tracking of former SIMMS and non-SIMMS subjects, born out by the higher nonmathematical ACT and SAT scores, mathematical ACT and SAT scores, and freshman mathematics grades of the non-SIMMS group. Conversely, SIMMS students chose science, mathematics, engineering, and technology majors more frequently than the non-SIMMS group but not significantly more frequently. Results also showed that, among three college mathematics Course Types of Developmental, Before Calculus, and Calculus, the Before Calculus courses were more reformed. However, there was no evidence to show that more reformed college mathematics courses favored either SIMMS or non-SIMMS group in terms of their grades. Interviews with 16 course supervisors affirmed the relatively traditional nature in most of the freshman mathematics courses at both Montana universities, which was quite different from what the NCTM Standards recommended. Additionally, freshman mean grades in the Before Calculus and Calculus groups were significantly higher than in the Developmental group.

The researcher concluded that 1) high schools need to monitor tracking of students; 2) the role of algebra in high school and college curricula needs to be redefined in terms of current technology and traditional value; 3) reconciliation between reformers and traditionalists must happen so that students can benefit from unified goals; 4) a new definition of academic intensity of the high school curriculum could provide focus and begin reconciliation between reformers and traditionalists; and 5) the gatekeeper nature of college developmental mathematics courses demands research into their effectiveness.

A COMPARISON OF FORMER SIMMS AND NON-SIMMS STUDENTS ON THREE COLLEGE-RELATED MEASURES

by

Michael Allen Lundin

A dissertation submitted in partial fulfillment

of

Doctor of Education

in

Education

MONTANA STATE UNIVERSITY—BOZEMAN
Bozeman, Montana

April 2001

APPROVAL

of a dissertation submitted by

Michael Allen Lundin

This dissertation has been read by each member of the dissertation committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the College of Graduate Studies.

Dr. Kenneth Borland _____ (Signature) 4/11/01 Date

Approved for the Department of Education

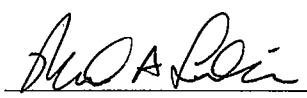
Dr. Gloria Gregg _____ (Signature) 4/20/01 Date

Approved for the Department of Graduate Studies

Dr. Bruce McLeod _____ (Signature) 4-20-01 Date

STATEMENT OF PERMISSION TO USE

In presenting this dissertation in partial fulfillment of the requirements for a doctoral degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library. I further agree that copying of this dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for extensive copying or reproduction of this dissertation should be referred to Bell & Howell Information and Learning, 300 North Zeeb Road, Ann Arbor, Michigan 48106, to whom I have granted "the exclusive right to reproduce and distribute my dissertation in and from microform along with the non-exclusive right to reproduce and distribute my abstract in any format in whole or in part."

Signature 

Date 4-20-01

This work is dedicated to my Mother and my Father without whose guidance I would have had neither the courage to begin nor the tenacity to finish. They are, and always shall be, my guides.

ACKNOWLEDGEMENTS

I wish to thank Dr. Ken Borland and Dr. Maurice Burke. Without their continued guidance throughout the process of planning and writing, this work would not have been possible. Through more than one storm, they held the tiller steadfast. This ship found port because of their foresight, hindsight, and insight.

I am also indebted to Dr. Rich Howard, who offered solid perspective to the methodology of this study, and to the other members of my committee, Dr. Chris Stryker, Dr. Elisabeth Swanson, Dr. Chris Yakymyshyn, and Dr. Lyle Andersen. They endured the requisite alterations inherent in the dissertation process with grace and dignity. Their suggestions led to a better product.

Finally, without the critical insights of Lynette Felix or the gracious concerns and encouragement of Diana Paterson in the SIMMS IM office, this work could not have been completed. I am grateful for their talents, their encouragement, and their friendship.

TABLE OF CONTENTS

TABLE OF CONTENTS - CONTINUED

TABLE OF CONTENTS - CONTINUED

TABLE OF CONTENTS - CONTINUED

LIST OF TABLES

LIST OF TABLES - CONTINUED

LIST OF FIGURES

Figure Page

ABSTRACT

This study compared three college-related measures of former high school students from a reform-based mathematics curriculum, produced by the Systemic Initiative for Montana Mathematics and Science (SIMMS) Project, with those of students from more traditional mathematics curricula. The measures of interest were ACT and SAT scores, freshman mathematics grades, and initial choices of majors. The subjects entered one of two state universities in Montana in Fall 1998. An Instructor Survey of classroom practices measured the intensity of reform in college freshman mathematics classrooms, and Supervisor Interviews produced qualitative data as background.

Results from the study included evidence of tracking of former SIMMS and non-SIMMS subjects, born out by the higher nonmathematical ACT and SAT scores, mathematical ACT and SAT scores, and freshman mathematics grades of the non-SIMMS group. Conversely, SIMMS students chose science, mathematics, engineering, and technology majors more frequently than the non-SIMMS group but not significantly more frequently. Results also showed that, among three college mathematics Course Types of Developmental, Before Calculus, and Calculus, the Before Calculus courses were more reformed. However, there was no evidence to show that more reformed college mathematics courses favored either SIMMS or non-SIMMS group in terms of their grades. Interviews with 16 course supervisors affirmed the relatively traditional nature in most of the freshman mathematics courses at both Montana universities, which was quite different from what the NCTM Standards recommended. Additionally, freshman mean grades in the Before Calculus and Calculus groups were significantly higher than in the Developmental group.

The researcher concluded that 1) high schools need to monitor tracking of students; 2) the role of algebra in high school and college curricula needs to be redefined in terms of current technology and traditional value; 3) reconciliation between reformers and traditionalists must happen so that students can benefit from unified goals; 4) a new definition of academic intensity of the high school curriculum could provide focus and begin reconciliation between reformers and traditionalists; and 5) the gatekeeper nature of college developmental mathematics courses demands research into their effectiveness.

INTRODUCTION TO THE STUDY

Chapter Introduction

Examined in this study was the question of whether two groups of Montana university freshmen differed with respect to college-related performance. Operationalizing this question meant comparing ACT and SAT test scores, freshman mathematics grades, and initial choices of majors for both groups. Comprising one group of subjects were students from a reform-based high school mathematics curriculum, produced by the Systemic Initiative for Montana Mathematics and Science (SIMMS) Project. In the alternate group were students from more traditional high school mathematics backgrounds. The students in this study entered one of the two state universities in Montana in 1998, and each site is treated separately. This study also included a survey of pedagogical practices in freshman mathematics classrooms at the same two Montana universities, as well as interviews with freshman mathematics course supervisors at those institutions. The survey and the interviews provided data for an informational framework, based on reform practices in freshman mathematics classes, instrumental in explaining potential differences in grades between the SIMMS and non-SIMMS students.

It was a premise of this study that comparing the performances of students from traditional and reform-based curricula, such as SIMMS, was not a trivial matter. Several facts, supported in the review of the literature, buttressed this premise. First, traditional

goals and prospective outcomes for learners differed from those of reform-based curricula, as did their respective theoretical foundations (Romberg & Wilson, 1995, p. 4). Therefore, assessment measures related to either type of curriculum alone, while useful for gauging performance for that particular curriculum, were not valid when comparing the two types.

Mathematics departments at both state universities in Montana, for example, used ACT and SAT scores or placement tests to sort students into their first university mathematics courses. However, a traditional test, such as the ACT, SAT, or a placement exam, does not and cannot measure ability outside its validity, and neither traditional mathematics content nor traditional test-taking processes align well with the content or processes valued by mathematics reformers (Romberg & Wilson, 1995, p. 4).

Second, university mathematics instructors practiced traditional teaching methods (LaBerge, Zollman, & Sons, 1997, p. 13) different from what many students may have experienced in high school. For SIMMS students in particular, more learner-focused interaction in high school (Dapples, 1995, p. 77; Thompson, 1992, p. 136) may have been replaced by lectures in college. Moreover, university faculty usually control their curricula (Janzow, Hinni, & Johnson, 1997, p. 499), including course assessments and evaluations. Consequently, former SIMMS students, caught between traditional and reform paradigms, found disturbing the traditional nature of the college mathematics classroom (Allinger, Lott, & Lundin, 1998, p. 29).

Third, reform-based standards, the products of mathematics professional organizations at secondary and post-secondary levels, were evolving during the period when subjects were in high school. While these various sets of guidelines seemed to

agree in their goals and practices for mathematics education, they were relatively novel. They were still changing, and all mathematics educators did not embrace them (LaBerge, Zollman, & Sons, 1997, p. 6).

Assessing the performance of SIMMS and non-SIMMS students in an equitable manner, then, was a matter of compromise. As such, ACT and SAT scores and freshman mathematics grades were of interest, since SAT and ACT scores correlate well with college freshman grade point averages (College Entrance Examination Board, 2000; ACT, Inc., 2000). Also, freshman course grades are "the single most revealing indicator of [students'] successful adjustment to the intellectual demands of a particular college's course of study" (Pascarella & Terenzini, 1991, p. 388). Importantly, data from both measures were readily available, since a concordance between ACT and SAT tests is often used to make decisions about admissions and placement (Dorans, Felicia, Pommerich, & Houston, 1997, p. 24). The traditional nature of ACT and SAT tests was of concern, however, because constructs measured by them may have underplayed the abilities of SIMMS students, whose experience had not been traditional. Hence, to better explain potential performance differences between former SIMMS and non-SIMMS students, an Instructor Survey produced data on practices in freshman mathematics classes. An analysis of this data provided some information about reform practices useful in explaining potential differences in freshman grades.

In addition to ACT and SAT scores and freshman grades, a third measure in this study, an initial choice of majors, was important. This measure represented student interest in a particular academic field, influenced, primarily, by pre-college experiences. This measure was chosen, then, to uncover differences in aspirations between former

SIMMS and non-SIMMS students. Especially, students' choices of majors in science, mathematics, engineering, and technology (SMET) were of interest here.

Since the reform movement is an embodiment of change, the theory of change entered into the subsequent exposition as a natural framework, without which, results would have been less meaningful. Additionally, various sets of professional standards stipulated the values, beliefs, and practices that might have guided mathematics instruction in a reform sense, and as such, are espoused by secondary and post-secondary professional organizations. That those standards have evolved and are still evolving with the reform movement is an idea that merited examination, if any recommendations were to stand firm as the bedrock of best practice. That evolution was examined here.

In conclusion, ACT and SAT scores, freshman grades, and initial choices of majors were measures that were important to Montana students (and others) in their university experiences. These measures represented available data that were used to compare the performance of former SIMMS and non-SIMMS students. In an academic arena challenged by the process of reform, the instructor survey of classroom practices and the interviews with freshman mathematics course supervisors provided background about the reform practices in Montana university mathematics classrooms. Differences in performance became more meaningful within the context of change theory, driven by various reform-based tenets of professional mathematics and collegiate organizations.

This chapter has three remaining parts. In the first part the research problem, purpose of the study, basis for the study, and research question are specifically addressed, beginning with brief history of the problem. In the second part of this chapter is a comprehensive introduction to the study, including its significance, operational

definitions, and assumptions and limitations. A chapter summary follows these two main sections.

Problem, Purpose, and Question

Background.

SIMMS vs. Traditional: A Comparison. The SIMMS curriculum materials embody the National Council of Teachers of Mathematics (NCTM) Standards (NCTM, 1989) for grades nine through twelve in a "learner-focused" (Thompson, 1992, p. 136) model of pedagogy. Each module begins with an Exploration or Activity that is built around a real-world problem. Embedded in the problem-solving process is the use of calculators or computers, and each lesson emphasizes cooperative learning by way of a discussion component. The use of physical materials or models is an integral part of many lessons, as are assessments that involve writing about mathematics. Students are encouraged to give presentations of work in progress and presentations of completed projects (MCTM, 1996-1998). These materials, in contrast to those that are more traditional, encourage a method of learning and teaching of mathematics that is anything but "linear subject, mainly concerned with mechanistically teaching facts and skills predominately related to number and generally characterized by paper-and-pencil activity..." (Nickson, 1992, p. 103).

On the other hand, Thompson (1992) would call that traditionalist approach "content focused" (p. 136), driven by the teacher-centered attempts to present content and explain it to students in order to foster conceptual understanding. A content-focused

approach may have an emphasis on understanding or an emphasis on performance, the first, driven by the structure of mathematics itself, and the second, driven by an instrumental or pedagogical approach. In either case, however, the mathematical content, more than the learners' abilities or interests, guides practice.

Consider, for example, the concept of linear equations, often a content item in high school freshman mathematics courses. A SIMMS approach to the subject, "Are You Just a Small Giant" (Carspecken, Eichenburger, Johnson, & Souhrada, 1997) begins with a brief reading about the world's tallest human, which is followed by formal definition of the geometric concept of similarity. This, in turn, is followed by a discussion of proportionality of various body parts. Students are guided through a lesson in which they measure each classmate's foot and compare this to his or her height. Gradually, they develop a natural constant that is the slope of a linear model of height as a function of foot size. Written activities and discussions that stem from the concepts of proportionality, growth patterns, and related subjects follow this lesson. The lesson, thereby, integrates activities, cooperative learning, technology, mathematical notions, and scientific topics in a manner consistent with the NCTM Standards. Notably, The Classroom Practices Inventory (APPENDIX A), used in this study to survey freshman mathematics instructors, captured the essence of those practices embedded in the SIMMS materials. Of course, the separate questions on that instrument were designed to measure practices recommended National Council of Teachers of Mathematics.

In contrast, a typical content-focused approach formally depends on definitions of proportions, slopes, linear functions, and similarity. Students might be given data but would not participate in its collection or the development of any model. Instead, a

mathematical model might be provided for them, and much of their subsequent work is analytic rather than synthetic. Any applications to science are secondary to the formal mathematical notion of linear function, and technology might not be a part of the process.

The literature review uncovered similarities between the NCTM Standards and other documents that purported to guide mathematics instruction. The Standards for Introductory College Mathematics Before Calculus (Cohen, 1995) and in the Guidelines for Programs and Departments in Undergraduate Mathematics (MAA, 1998a), for example, prescribed how college and university instructors should teach. More generally, The Principles of Good Practice for Undergraduate Education (Chickering & Gamson, 1987) encouraged active learning, cooperative learning, and diverse ways of learning. Therefore, much of what was prescribed in all of these documents was found in the SIMMS modules, and those practices were reflected in the Classroom Practice Inventory in this study.

In summary, substantial differences existed between traditional and SIMMS curricula. Those differences naturally led to questions about the comparative performance of former SIMMS and non-SIMMS students on college-related measures, and those questions have guided this study.

The SIMMS Project. In 1991 the National Science Foundation (NSF) funded the Montana Council of Teachers of Mathematics with a ten million-dollar grant, enabling the Systemic Initiative for Montana Mathematics and Science (SIMMS) Project to begin five years of operation (The SIMMS Project, 1996, p. 1). A major goal of the SIMMS Project was to write, publish, and put into practice a mathematics curriculum for grades

nine through twelve that would reflect the (then) recently published Curriculum and Evaluation Standards for School Mathematics (NCTM), 1989).

As planned in the grant proposal, the SIMMS Project implemented its curriculum, based on the NCTM Standards, in many Montana high schools (MCTM, 1997, p. 6). Four years later, in the fall of 1996, the first cohort of SIMMS-educated mathematics students began matriculating at Montana's institutions of higher education. Although the SIMMS Project had evaluated its pilot curriculum at the high school level (MCTM, 1998), it had not yet studied the effects of its reform-based curriculum on college students, since that cadre of subjects had not graduated from Montana high schools. Realizing that "Preparing students for post-secondary education involves more than simply guiding [students] through SAT or ACT tests and the admissions process" (Allinger, Lott, & Lundin, 1998, p. 16), the SIMMS Project staff conducted pilot research to better understand the attitudes and performance of college freshman who had been SIMMS students. That research was timely, having been prompted by public and professional interest in the college performance of SIMMS veterans, as communicated by Glenn Allinger, SIMMS Professional Development Co-Chairman (personal communication, March 28, 2000).

How did SIMMS educated students perform in college mathematics when compared with non-SIMMS students? Analyses of qualitative and quantitative data gathered in the 1997 pilot study began to address that question (Allinger, Lott, & Lundin, 1998), and the findings were noted in the literature review. Those findings, though meager, suggested possible differences between SIMMS and non-SIMMS groups.

However, the general question of the relationship between the SIMMS curriculum and college students' performance and initial choices of majors remained open.

Research Problem

It was not known whether former SIMMS and non-SIMMS students differed in terms of college-related measures. In particular, it was not known whether they differed with respect to ACT scores, SAT scores, college freshman mathematics grades, or initial choices of major.

Purpose of the Study

The purpose of this study was to determine if variations in college-related measures existed between two groups of former Montana high school students, those from a more traditional mathematics programs and those from the SIMMS integrated mathematics program. ACT scores, SAT scores, and freshman mathematics grades were the more traditional measures used to compare the SIMMS and non-SIMMS groups, and initial choice of majors, while not a performance indicator, was a measure of interest in mathematics or a related field. The literature review lent support to this choice of measures as meaningful to the study at hand. That review also supported a survey of classroom practices of college mathematics instructors and interviews with freshman mathematics course supervisors to provide a backdrop for the discussion of performance comparisons between former SIMMS and non-SIMMS students.

Basis for the study

Questions of National Interest. Although this study was specific to Montana, it was one that sprang from much national interest in mathematics reform, about a decade old, if dated from the first publication of the NCTM Standards. However, relatively little data existed on the effects of reform-based high school mathematics on college students. The dearth of literature existed despite much funding by NSF in an effort to promote reform in mathematics and science (NSF, 1999). Furthermore, the evolution of the reform movement had spawned many general questions, begging for extensive research related to the parameters of this study. For example, the following questions were of public concern and related to the questions that guided this study.

How do students educated with reform-based curricula in general compare with those who have studied in more traditional settings? How do high school mathematics departments best prepare their students for college mathematics courses? What are effective ways to formulate curricular goals and align them with practice in both secondary and post-secondary mathematics departments? What are effective strategies to promote dialog between secondary and post-secondary mathematics departments? To what extent do mathematics departments in high schools and colleges now conform to professional standards for educating students? (Green, 1999; Mervis, 1998)

The Perspective of Change Theory. Change theory has much to do with many aspects of reform, which serves as a rich example of the dynamics in the process of change. For this study, it was necessary to enter the realm of program comparison. Viewing mathematics programs from a perspective of change theory brought forth issues,

structures, and processes associated with these programs, allowing them to be better scrutinized and compared with one another. From these comparisons came suggestions for more effective programs as well as answers to questions about differences between SIMMS and more traditional curricula.

Without this backdrop of change theory, any findings of this study would have been less meaningful, because reform goals were evolving, and those goals tended to drive mathematics assessment. Due to the dynamic nature of mathematics reform, the reform goals represented, metaphorically, moving targets. That is, for many levels of assessment, the goals changed, exemplifying what Hall and Hord (1987) called "mutual adaptation" (p. 116) of beliefs and practices of various players. For this study, mutual adaptation explained, for example, why few assessments of reform projects included formal evaluations at the post-secondary levels. Also exemplifying mutual adaptation, an initial choice of majors by college students was once recommended indicator of systemic reform by NSF (MCTM, 1997, p. 2), although it was not specifically listed later (National Science Foundation, 2000). Mutual adaptation, then, is itself evidence that mathematics reform goals are evolving.

Principles, Drivers, and Professional Standards. In this study, four main sets of documents had prescriptions for mathematics educators, fortified by general principles for college and high school teaching. Comprising those sets of documents were the three-volume NCTM Standards (NCTM, 1989; NCTM, 1991; NCTM, 1993), the American Mathematical Association of Two-Year Colleges (AMATYC) Crossroads in Mathematics: Standards for Introductory College Mathematics Before Calculus (Cohen,

1995), the Mathematical Association of America's (MAA) Guidelines for Programs and Departments in Undergraduate Mathematical Sciences (MAA, 1998), and the Principles for Good Practice in Undergraduate Education (Chickering & Gamson, 1987).

The three-volume NCTM document included the Curriculum and Evaluation Standards for School Mathematics, the Professional Standards for Teaching Mathematics, and the Assessment Standards for School Mathematics. Together, these volumes strongly suggested a framework for the curricula, pedagogy, and assessment of high school mathematics. A revision of the Standards, released in Spring 2000 (NCTM, 2000), delivered even more input from the mathematics community following a draft presentation and a year of dialogue about that draft (NCTM, 1998).

At the post-secondary level, the AMATYC and MAA documents served as guidelines for mathematics educators. It should be noted that the MAA Guidelines were less prescriptive than were the AMATYC Standards, which were similar to the NCTM Standards in structure and wording. More generally, similar prescriptions for mathematics educators existed among the NCTM, AMATYC, and MAA documents, and those were emphasized in the review of the literature.

Noteworthy also were the best-practice principles developed for general college teaching by Chickering and Gamson (1987), because they shared common ground with those practices recommended by the NCTM, AMATYC, and MAA. According to Chickering and Gamson in the "Seven Principles for Good Practice in Undergraduate Education," good practice encourages contact between students and faculty, collaboration among students, active learning, prompt feedback, time on task, high expectations for students, and diverse but equitable ways of knowing (Chickering and Gamson, 1987, p.

7). The deployment of cooperative learning, active forms of learning, and equitable ways of learning were three of the recommendations also found in the NCTM, AMATYC, and MAA documents. (In the MAA Guidelines see section "3. Mathematics and General Education" on p. 8.) The Instructor Survey in this study was in part concerned with two of these principles, cooperative and active learning, and with standards from the other documents as well.

Additionally and importantly, at the state and regional levels, NSF's "Drivers of Reform" were designed to guide grant awardees in the direction of systemic change. (See APPENDIX B for the NSF Drivers of Reform.) Particularly, driver five was important to this study, since in its original form, it recommended counting college majors as an indicator of systemic change (Montana Council of Teachers of Mathematics, 1997, p. 1; National Science Foundation, 2000). More about the drivers and their apparent changes was stated in the review of the literature.

In conclusion, the various sets of educational guidelines shared goals for high school and college classrooms and were important in grounding this study. Those guidelines came from professional organizations and served to steer mathematics teaching and learning down the road of reform. Together with the NSF Drivers, those sets of guidelines provided a rationale for examining an initial choice of majors among university freshmen, and for measuring the reform-based practices in freshman mathematics classes. In general they buttressed the theory and practice of mathematics reform in secondary and post-secondary classrooms, and in particular they supported that theory and practice associated with the SIMMS curriculum. For the research at hand,

those documents provided reasons to monitor initial choices of majors among college freshmen and to profile the pedagogical practices in their mathematics classes.

Research Questions, Objectives, and Variables

Primary Question. How did former SIMMS and non-SIMMS students at either Montana state university in 1998 compare with respect to three college-related measures: ACT and SAT scores, freshman mathematics course grades, and initial choices of major?

Subsidiary Question. Did the reform-type classroom practices of freshman mathematics instructors differ by Course Type: Developmental, mathematics Before Calculus, or Calculus, and if so, were these differences evident in the performance of former SIMMS and non-SIMMS students, according to their freshman mathematics grades disaggregated into the same three categories?

Objectives. There were three research objectives. The first objective involved gathering data from those college freshmen entering the two Montana state universities in Fall 1998. This study compared former SIMMS students and non-SIMMS students with respect to ACT and SAT scores, performance in freshman mathematics courses, and initial declarations of majors. An analysis demanded a sufficiently large sample of former SIMMS students, which only became possible with the entering freshman class of 1998. The comparison of SIMMS and non-SIMMS students focussed on the traditional-aged freshmen cohorts at both universities, and the data included mathematics grades from both semesters of the students' freshman year. In what follows, data from objective one was referred to as "Freshman Data."

Objective two involved surveying college mathematics instructors at the freshman level as a basis for comparing the mathematics performance of former SIMMS and non-SIMMS students. Objective two was referred to as the "Instructor Survey," and data came from the instructors' self-reports about the intensity of their classroom practices as they aligned with those recommendations of the NCTM, MAA, and AMATYC.

Objective three involved interviewing freshman mathematics course supervisors to ascertain course goals, expectations, standards, and outcomes. Data from these interviews added depth and dimension to the instructor survey data. Objective three was referred to as the "Supervisor Interviews."

Together, the results from Freshman Data, the Instructor Survey, and Supervisor Interviews provided a rational means of accounting for differences in freshman mathematics grades between SIMMS and non-SIMMS students. In addition, the results from the Instructors Survey and Supervisor Interviews described the state of reform among freshman instructors at Montana's two state universities.

Variables: Freshman Data. The independent variable for this research piece was Curriculum Type, either non-SIMMS or SIMMS, used by the respective groups of students. There were three categories of dependent variables. Entering freshman ACT and SAT scores constituted the first category. Separately, those variables were ACT Mathematics, Science Reasoning, English, Reading, and SAT Mathematics and Verbal scores. Freshman mathematics course grades, disaggregated by course type, constituted the second category of dependent variable. Lastly, the levels of SMET, non-SMET, or Undeclared, determined the third category of dependent variable, students' choices of

majors. That variable represented one of three curricular choices: science, mathematics, engineering, or technology (SMET), a major that is none of the previous (non-SMET), or the non-declaration of a major.

Variables: Instructor Survey. The independent variable for objective two of this research study, the Instructor Survey, consisted of the Course Type, having three levels: Developmental, mathematics Before Calculus, or Calculus. The dependent variables were measures consistent with professional standards for high school and college mathematics: problem-solving, cooperative learning, connections with other disciplines and other areas of mathematics, use of technology, equity, and mathematical reasoning. There were twelve variables in all, and three of these represented measures that are more traditional: working problems from a text, lecturing in class, and using worksheet-based exercises in class. See the Classroom Practices Inventory in APPENDIX A.

Introduction to the Study

Significance of the Study

A National Need to Know: The Math Wars. Politically, the mathematics education community has been involved in what has been dubbed "math wars" (Green, 1999, p. 1), an ideological battle between traditionalists and reformers. Skirmishes occurred in secondary and post-secondary settings, in school board meetings, and at state and national levels. Respected mathematics journals, such as The American Mathematical Monthly, devoted editorial pages to both camps in an attempt toward reconciliation, as indicated in recent articles by Tucker (1999) and Krantz (1999). Green

(1999) and the January edition of Focus (2000) noted that, among those taking sides in the war, were several Nobel laureates and the Secretary of Education, Richard W. Riley, joined by of educators, parents, students, and community members.

Spurred on by the findings of a Department of Education Expert Panel, formed by Secretary Riley, traditionalists wrote a letter attacking the panel's choice of ten recommended "exemplary" and "promising" curricula, all, more or less, based on the tenets of reform found in the NCTM Standards. The letter to Secretary Riley, published November 18, 1999 as a paid advertisement in the Washington Post (NCTM, 1999), was found at the Mathematically Correct web site (http://www.mathematicallycorrect.com), while the NCTM's formal support of Secretary Riley's endorsement (Thorpe, 1999) was on-line at the NCTM web site (http://www.NCTM.org/rileystatement.htm).

In spite of the polemic fueled by ideological differences between reformers and traditionalists, any real clarity of issues must occur in the trenches of research, as in this study, which compared traditional curricula to one that is more reformed. Results from this study led to suggestions for more productive secondary and post-secondary mathematics programs and to smoother transitions for incoming college freshmen.

A Local Need to Know: High Interest, Little Data. Only eight years have passed since the National Science Foundation awarded the original twenty-six State Systemic Initiative grants, which included the SIMMS Project, and a little more than ten years, since the NCTM published its Curriculum Standards. The first SIMMS cohort graduated from college with the class of 2000. The mathematics reform movement weathered a hearty gale in a sea of politics (Mervis, 1998; Schontzler, 2000), yet there was little

information about the relationship between the SIMMS curriculum and college performance. More surprisingly, despite national involvement and interest in reform, there was little data on the effects of any high school mathematics reform curriculum on college students, although the results of mathematics reform at the college level are beginning to surface (Hurley, Koehn, & Gantner, 1999). This was a time for research, since interest was high, and the subjects were accessible.

The SIMMS materials comprised all or a substantial part of the curriculum in 17 high schools across Montana and supplemented the curricula in many more schools nationwide, according to SIMMS staff member, Lynette Felix (personal communication, April 11, 1999). The exact number of additional schools that used SIMMS materials as supplements was not known. Only about 100 former SIMMS students at each university in this study met the criteria of three years or more of SIMMS preparation to be included in the SIMMS group. Nevertheless, more than one seventh of Montana's 114 public high schools (OPI, 1998, p. 1) used SIMMS materials at some level at the time of this study, according to Felix, and Montana tax dollars supported that choice. Consequently, Montana citizens deserved to know as much as possible about that curriculum. This study answered several questions about the SIMMS curriculum and its relationship to college performance within the contexts of mathematics reform, change, and professional standards.

Definition of Terms

1. ACT Score: score on the American College Testing Service exam, purportedly measuring curriculum-based educational development, higher-order thinking skills,

and knowledge in college preparatory courses (Jeff Schiel, Senior Research Associate to ACT, personal communication, January 28, 1999).

2. Assessment: the process of gathering evidence about performance. (Adapted from NCTM, 1995, p. 87).
3. Concordance: statistical relationship between the SAT and ACT scores allowing conversion from one type of score to another (Jeff Schiel, Senior Research Associate to ACT, personal communication, January 28, 1999). Note that no psychometric relationship between the two tests is implied, because the tests measure different constructs.
4. Curriculum: academic plan consisting of purpose, content, sequence, learners, instructional resources, instructional processes, evaluations, and adjustments (Stark and Latucca, 1997, p. 15).
5. GPA: grade point average, based on a four-point scale.
6. Mathematics Program (high school): a full course of study of at least three years of high school mathematics (three Carnegie units) acceptable for entrance into Montana universities.
7. Mathematics Program (university): mathematics programs intended for the first two years of university learning, including developmental courses.
8. Program Evaluation: The process of determining the effectiveness of an educational program in achieving its goals and, therefore, its value in comparison to its required resources (NCTM, 1995. P. 90).
9. SAT Score: score on the Scholastic Aptitude Test I (SAT I) written by the Educational Testing Service, purportedly measuring the general ability to reason

quantitatively (Jeff Schiel, Senior Research Associate to ACT, personal communication, January 28, 1999).

10. SMET: science, mathematics, engineering, and technology.
11. SSI: State Systemic Initiative awards by the National Science Foundation.

Assumptions and Limitations of the Study

1. The data sets received from the two Montana state universities were both complete and valid. All data received were scanned and cross-checked for accuracy with smaller data samples used in previous studies. However, it is possible that errors in data retrieval could have compromised the validity of the data.
2. Each SIMMS and non-SIMMS subject received at least three years of high school mathematics, as required for admission to both Montana state universities and no more than four years of high school mathematics. It is a possibility that subjects in the SIMMS or the non-SIMMS groups may have received more or less mathematics instruction than was required for entrance into these institutions. In this study no documentation, experimental control, or statistical control was employed to explain variation due to this phenomenon.
3. Historical and qualitative evidence was presented in the literature review to argue that substantial differences exist between the SIMMS curriculum, as it was implemented in the years 1992-1998, and the more traditional high school mathematics curricula of that time period. However, gathering extensive data to compare and contrast the many mathematics curricula of that period was not in the scope of this study. It is possible that other curriculum materials in use at that time shared characteristics

similar to those of SIMMS and that these similarities could detract from the statistical analysis of the data and conclusions of this study. It is also possible that a comparative study of those curricula could enhance the understanding of factors affecting the dependent variables in this study.

4. In this study, for the purpose of statistical analysis of freshman mathematics grades, data were disaggregated into three categories: Developmental, mathematics Before Calculus, and Calculus. It is assumed that courses in each category shared enough characteristics to warrant inclusion into that category. No statistical analysis was presented to argue for this division, which constituted a first step in course placement rubric at both universities in this study. It is possible that a finer partition of courses could have yielded more telling results, but limited sample sizes in each of the resulting categories might also have prevented valid statistical analysis.
5. Attending college is done by self-selection rather than by random assignment (Pascarella and Terenzini, 1991, p. 566). By choosing to study those freshmen students at two universities, the researcher could not account for those who went elsewhere or those who did not go to college. It is possible that not including those who choose out-of-state schools affected the mean SAT-ACT concordance or otherwise biased the values of the dependent variables measured here. It was also possible that limiting the populations to two Montana universities biased grade point averages and SMET tallies. There was no evidence, however, that the SIMMS and non-SIMMS students who had chosen to go to colleges outside Montana were disproportionate.

Format for the Study

This study consists of four more chapters in addition to this introductory chapter. The review of the literature in Chapter Two supported the context and the current understanding of freshman college-related performance, initial choices of majors, and college mathematics classroom practices as measures. The supporting literature was summarized and examined for weaknesses and strengths, and avenues of further inquiry were noted. In Chapter Three, the methodology section included information about the populations studied, the instruments and materials used for the study, and the research design. Assumptions and limitations of the methodology were included there also, as was a timeframe for the study. In Chapter Four the results of the data analysis and a discussion of the meaning of those results were included as well as a Supplemental analysis. Finally, In Chapter Five a discussion of the broader theoretical and practical implications of this study as well as its limitations can be found. Recommendations for further research, in which the research question regarding differences between SIMMS and non-SIMMS college-performance and their curricular choices, were summarily included in Chapter Five.

Chapter Summary

Challenged by the new NCTM Standards and funded by NSF, the SIMMS Project created a high school mathematics curriculum and implemented that curriculum in Montana, as the mathematics reform movement evolved. Ideological conflicts, predicted by change theory, raised awareness of the issues surrounding reform and generated

demands for research on reform. Disparities existed between traditionalists and reformers about what constitutes curriculum, practice, and assessment of mathematics in high school and college, partially because the relationships between reform curricula and performance had not been well-documented. Also, disparities existed between what was recommended by professional standards for secondary and post-secondary mathematics programs and what was actually practiced. Finally, the evolution of mathematics reform created assessment goals that continued to evolve, confounding valid performance comparisons between former SIMMS and non-SIMMS students.

The main objective of this study was to compare the performance and initial choice of majors of incoming university freshman from two different groups: former SIMMS and non-SIMMS students. Each of the two Montana state universities was treated separately in the study. ACT and SAT scores, university freshman mathematics grades, and initial choices of majors of former SIMMS and non-SIMMS students comprised the measures here. Additionally, the Instructor Survey of classroom practices and course Supervisor Interviews provided information about university mathematics reform with which freshman grades were compared to explain differences between groups. Also, the instructor survey provided a measure of the disparity between what was recommended by professional standards and what was actually practiced in the university mathematics classroom at Montana's state universities.

The combination of traditional performance measures, ACT and SAT scores and freshman mathematics grades, and two reform-based measures, initial choice of major and the amount of reform in mathematics classes, represented the union of traditional and reform constructs. Such a union was needed to equitably compare both SIMMS and non-

SIMMS students. The measures and their associated instruments represented a compromise between accessibility of the data, validity of measured constructs, and value to both traditionalists and reformers. The results of this study began to clarify the relationship between traditional and reform mathematics in terms of high school and college performance, a relationship that the existing literature did not clarify.

LITERATURE REVIEW

Chapter Introduction

This literature review consists of four main parts: Criteria for Literature Selection, Theme Development, Conclusions, and Chapter Summary. The Themes section includes Problem Context, Understanding of the Problem, Previous Research and Findings, and a Review of Methodologies. The reader will note that problem context is complicated, and the section that describes it is necessarily detailed. The concluding section summarizes the review, points out strengths and weaknesses of the research, points to gaps and saturation levels in the knowledge base, and alludes to avenues of further inquiry.

Criteria for Literature Selection

The literature pertinent to this study was drawn from six main sources: 1) change theory; 2) guidelines for mathematics education, developed by professional organizations; 3) sources on learning, teaching, and the college curricula; 4) documents produced by the SIMMS Project; 5) research on reform in mathematics education; and 6) studies about the meaningfulness of high school and college-related performance measures.

In the literature on change theory and practice, books by Fullan and Stiegelbauer (1991) and Hall and Hord (1987) complement each other as accepted texts on planned

change. While the first text focused on the effects of change in elementary and secondary school settings, the second more formally outlined the structure and process of planned change from the perspective of the Concerns Based Adoption. The Concerns Based Adoption Model (CBAM), "provide [ed] ways to label change process phenomena, to take positive action in facilitating change, and to predict effects" (Hall & Hord, 1987, p. viii).

At the elementary, secondary, and post-secondary educational levels, sets of professional guidelines for mathematics education were written to inform practice. The pedagogical aspirations of various professional organizations were embodied in these guidelines, and they shared common prescriptions for educators. These sets of guidelines were of interest in this study, since the SIMMS curriculum (and other curricula) had roots in them, although the extent to which college mathematics instructors follow them at Montana universities was not known. Indeed, this was a question considered in the work at hand.

The three volumes of professional standards, published by the NCTM, constituted the primary framework for the mathematics reform movement. These volumes included prescriptions for curriculum and evaluation, pedagogy, and the assessment of elementary and secondary mathematics. A new volume of revised NCTM guidelines, Principles and Standards for School Mathematics (NCTM, 2000), which combined the previous three volumes with input from various professional organizations, was also of interest. After two years in draft form (NCTM, 1998), the published volume appeared in April 2000.

Documents analogous to the NCTM Standards, meant to guide post-secondary mathematics educators, were cited also. The American Mathematical Association of

Two-Year Colleges (AMATYC) published its Crossroads in Mathematics: Standards for Introductory College Mathematics Before Calculus in 1995. The Mathematical Association of America (MAA) developed its Guidelines for Programs and Departments in Undergraduate Mathematical Sciences (MAA, 1998a), and its Quantitative Reasoning for College Graduates: A Complement to the Standards (MAA, 1998b) to advise mathematics educators in American colleges and universities.

Providing a more general prescription for best practice were the Seven Principles for Good Practice in Undergraduate Education, developed for college teaching by Chickering and Gamson (1987). According to those researchers, good practice encourages contact between students and faculty, collaboration among students, active learning, prompt feedback, time on task, high expectations for learners, and diverse but equitable ways of knowing. (Chickering and Gamson, 1987, p. 7) Similarities in prescriptions for educators between the Principles and the NCTM, AMATYC, and MAA documents were clearly evident, and the principles themselves were grounded in Chickering's psychosocial research and theories (Evans, Forney, & Guido-DiBrito, 1998, p. 16; Pascarella & Terenzini, 1991, p. 19).

More specifically, cooperative learning, active forms of learning, and equitable ways of learning were three of Chickering's and Gamson's principles also advocated in the NCTM, AMATYC, and MAA documents. (In the MAA Guidelines, see section "3. Mathematics and General Education" on p. 8.) Additionally, Chickering's psychosocial theories affirmed the importance of measuring cooperative and active learning in the study at hand. Pertinent to this research study, the Instructor Survey contains questions regarding two of these principles, cooperative and active learning.

Texts by Pascarella and Terenzini, (1991), Stark and Latucca (1997), and Evans, Forney, and Guido-DiBrito (1998), documenting the effects of college on students and knowledge about college curricula, were cited in this study because these texts are accepted references on the college experience. Citations from various chapters in the anthology, Handbook of the Undergraduate Curriculum (Gaff and Ratcliff, 1997), appeared here for the same reason. Likewise, the work of Romberg and Wilson (1995) was taken to heart in this study, because the authors exhibited extensive knowledge about assessment and mathematics education.

Much particular information for the work at hand came from the many publications of the SIMMS project. The SIMMS curriculum, the source of the phenomenon under investigation, included six levels of materials with either three or four books of modules at each curricular level (MCTM, 1996-1998). In addition to the curricular materials, the SIMMS Project staff produced a number of monographs that document the project's goals, research and evaluation studies, and a final report. Those monographs were an indispensable source of data, while the structure of the curriculum materials themselves constituted strong evidence that former SIMMS students underwent a unique educational experience.

Also of related interest was research that profiled the beliefs and practices of high school teachers (Weiss, Upton, and Nelson,1992; Weiss, Matti, and Smith,1994) and university faculty (LaBerge, Zollman, and Sons, 1997) with respect to recommendations of the NCTM Standards. The Classroom Practice Inventory (APPENDIX A), the Instructor Survey instrument used in this study to measure alignment of classroom practices with the NCTM Standards, came directly from the work of the latter three

researchers, having been adapted (p. 6) from a study of Weiss, Upton, and Nelson (1992). Recent dissertations by Kull (1996) and Hawkins (1998), respectively, brought to light the degree of awareness of the NCTM Standards in preparatory programs in colleges and universities and the effect of NCTM standards-based education on the self-concept and achievement of university freshmen.

Two conflicting evaluations of the Core-Plus Mathematics Program (CPMP), a reform-based high school mathematics project and a contemporary of SIMMS, were of special interest because of similarities between Core-Plus and SIMMS. Core-Plus' own evaluation, Contemporary Mathematics in Context: Evaluation Results, included an examination of college student performance and is written and printed for public consumption. (Core-Plus Mathematics Project, 2000b). The glossy report contrasted in substance, appearance, and results with an independent study of college student performance prepared by two mathematicians, Bachelis (1998, 1999) and Milgram (1999). Those researchers worked in tandem to investigate the achievement and attitudes of college freshmen and former CPMP students. Like the Core-Plus evaluators, Bachelis and Milgram compared SAT and ACT scores, freshman mathematics grades, course placement, and attitudes of former CPMP students and students of more traditional curricula. Unlike Core-Plus' findings, however, Bachelis' and Milgram's analysis cast CPMP in unfavorable light. They self-published their findings, and because both researchers are mathematicians, their work, because of its implications and political ramifications, had an impact that could not be dismissed. Notably, that work was attacked for its lack of rigor by Core-Plus, but still merited attention, because the findings contradicted those of Core-Plus' own evaluation.

Unfortunately, the research of Bachelis and Milgram stood out, perhaps, more than it should have, because it stood alone. The reform movement had not matured enough to support similar studies. Such work has only recently become possible, as the first students involved in reform-based mathematics projects graduated from high school and entered the doors of colleges and universities.

Finally, much literature existed on the importance of college-related performance measures. This research supported the need for and use of ACT and SAT scores, freshman mathematics grades, and initial choices of majors in the study at hand. Especially, the work of Adelman (1998, 1999), based on data collected by the National Center for Educational Statistics, indicated a relationship between performance measures and degree attainment. McCormick (1999) correlated degree attainment to first year GPA, and Murtaugh, Burns, and Schuster (1999) established the same relationship for sophomore and senior years. Benefield (1996) established a strong correlation between ACT scores and successful completion of a pre-engineering program, and Levin and Wycoff (1990) determined those courses whose grades were predictors for persistence in engineering. Lewis (1996) and Kull (1996) conducted national descriptive studies, investigating remedial courses in colleges. Their work demonstrated the tendency toward remediation at the college level and some of the consequences of that phenomenon. This type of study was of interest here because remedial mathematics courses existed at Montana universities.

Why students choose their majors was important in this study, because initial choice of a SMET major is related to high school curricula and performance (Drew, 1996, p. 88). Dawson-Threat and Huba (1996, Strenta (1994), Drew (1996), Maple and

Stage (1991), and Pearson and Fechter (1994) examined the phenomena associated with choices of majors, and their work was considered herein.

Theme Development

Context of the Problem

Problem Statement in Context. It was not known if there were differences between former SIMMS mathematics students and those taught from more traditional high school mathematics curricula (non-SIMMS) in terms of ACT and SAT scores, freshman college mathematics grades, or initial declarations of college majors. Furthermore, a fundamental premise of this study was that the theoretical frameworks, goals, and expectations of the two program types, SIMMS vs. traditional, were different, and that difference confounded the process of comparing the achievements of students from both programs. Finally, the criteria against which achievement should be measured had evolved with the mathematics reform movement, resulting in the ambiguity of goals and outcomes for students. Despite the fact that traditional measures such as ACT and SAT scores and freshman course grades were important to this study, more modern measures were needed, based on professional standards, to round out a profile of mathematics performance for freshman students. Students' initial choice of majors and the survey of classroom practices served this purpose by adding reform-based criteria to the assessment.

The National Council of Teachers of Mathematics (NCTM) and the National Science Foundation (NSF) have been powerful forces behind the reform movement, each

serving the mathematics education community in a different manner. During the last decade NSF has awarded millions of dollars to promote substantial change in the way mathematics is taught, and its six Drivers of Reform (NSF, 1999) were meant to keep awardees at the project and state levels on the reform trail. At the district, school, and classroom levels, the NCTM Standards (NCTM 1989, 1991, 1995), three volumes of them, were written to guide educators along that same path.

The evolutionary nature of the mathematics reform movement cannot be overlooked, and the NCTM Standards and the NSF Drivers of Reform (National Science Foundation, 2000) are more a product of that movement than a pre-movement handbook, as this review will demonstrate. (See APPENDIX B for a concise listing of the Drivers of Reform and the NCTM Principles and Standards.) Reform planners considered both the Standards and the Drivers important, even though, ironically, neither was fully developed until after large reform projects began. Consequently, those who planned and implemented reform-based projects early in the movement lacked the guidance that the refined documents could have provided.

Change From the perspective of educational change theory in the sense of Hall and Hord (1987) or Fullan (1991), the mathematics reform movement was expected to have evolved in phases. Those nine phases were described by Hall and Hord (1987), as research, development, diffusion, dissemination, adoption, implementation, institutionalization, refinement, and abandonment (p. 331). Evidence in the literature affirmed that reform had evolved moderately in both secondary and post-secondary levels

(Weiss, Matti, & and Smith, 1994; LaBerge, Zollman, & Sons, 1997; Hurley, Koehn, & Gantner, 1999).

Mervis (1998) and Green (1999) affirmed criticism of the mathematics reform projects, the rough spots along the reform trail, predicted by theorists, but there also were planned interventions to deal with pot holes (Hall & Hord, 1987, p. 142). It was not always possible to predict when or what interventions were needed, however, or to keep "a lot of 'plates spinning simultaneously in the air'" (MCTM, 1997, p. 33) while in the process of implementing systemic change. Keeping all those plates spinning in the arena of statewide performance standards was a difficult trick in Montana, particularly, and that difficulty indicated a gap between dissemination and implementation phases of planned change.

More specifically, although Montana mathematics standards had been disseminated, (http://www.state.mt.us/usys/edu.htm), the state balked at implementing statewide assessment based on the tenets of reform, unlike some other states. Compare, for example, Vermont's progressive assessment system at the web site, http://www.FairTest.org. About Montana's current assessment practice, the assessment organization, FairTest, states " Montana has a bare-bones state assessment program that needs many major improvements. The state system relies entirely on multiple-choice, norm-referenced tests in three grades. This program should be replaced" (FairTest: The National Center for Fair and Open Testing, 2000, Montana data.). It is plausible, then, that Montana shared with some other states a more moderate phase of reform, as described in the change hierarchy of Hall and Hord (p. 331). More pointedly, Montana's educational institutions were not driven by a state-mandated reform agenda. Hence, there

were no statewide tests, based on the tenets of mathematics reform, with which to measure students' achievement.

Exemplifying a statement by Fullan, "Change is a Process, Not an Event" (p. 8), the guidelines for reform evolved with the movement itself. This evolutionary approach to reform was not without consequences, even if planned change was the reform strategy of the NCTM. One consequence was that the three volumes of "boiler-plate" Standards appeared and, indeed, evolved over a period of seven years. Likewise, NSF's Drivers for Reform, the guidelines meant to spur grant awardees down a narrower path of change, appeared mid-program for those awardees (MCTM, 1997, p. 1). The appearance of the NCTM Standards and NSF Drivers in this manner constituted rule-making with a play in progress. In any case, the latent appearance of the rules of the game undermined the belief that SIMMS, or any other reform-based curriculum project at that time, had access to all of the crystallized goals that would later appear in the NCTM Principles and Standards or the NSF Drivers.

Summarizing, because of the evolutionary nature of the reform movement, any equitable and valid comparison between the performance of SIMMS and non-SIMMS students was considered in light of the on-going changes. Appropriate measures were chosen as a compromise between reform and traditional philosophies in vogue at the time that students in this study took their mathematics in high school and college.

The Disparity Between Assessment Types. For entering freshman just out of high school, admission to Montana's two state universities required their ACT or SAT scores, high school grade point averages (GPA), and class ranks. (Montana State University

[MSU], 1999; The University of Montana [UM], 2000) These measures, together with high school course grades and scores from several other standardized tests, traditionally became the trails of achievement left by graduating high school seniors. Although grade point averages and class ranks may have been program specific, standardized tests were not, since those tests were not designed to measure program specific knowledge, skills, or aptitudes. As Stake (1995) implied, even when a mathematics program's content closely matches that of some standardized test, that test fails to measure alternative constructs, specific knowledge, or useful skills (p. 173). Instead, standardized tests do purportedly measure those constructs agreeable to test writers, even when the test writers have little in common with those tested. Thus, a trail of standardized test scores may fail to indicate strong points of a mathematics program, instead measuring some selected concerns of disconnected "experts."

In contrast to traditional assessment practices, Romberg and Wilson (1995) argued for an authentic assessment framework based on the NCTM Standards to meet the challenge of a modern paradigm for mathematics education, founded, in part, on constructivism. Doing mathematics is neither a static nor a linear process. That process is too rich to be adequately assessed by standardized tests alone, despite pervasive traditional testing methods that assess static and linear knowledge. (Romberg & Wilson, 1995, p. 4) On the contrary, doing mathematics in a manner recommended by the NCTM Standards demands richer assessments:

> If one considers mathematics to be a static, linearly ordered set of discrete facts, then the logical choice for a valid assessment system is the traditional standardized achievement test. On the other hand, if one views mathematics as a dynamic set of interconnected, humanly constructed ideas, then the assessment system must allow students to engage in rich

> activities that include problem solving, reasoning, communications, and making connections. (p. 4)

The importance of the latter statement to this study could be denied: traditional and reform-based assessments clearly differed on theoretical grounds.

The Disparities in Program Assessment and Evaluation. "The primary question to be answered in any program evaluation is, How well is the program working in relation to goals and expectations for the students?" (NCTM, 1995, p. 66) Implicit in this question was the notion that program assessments should align with program goals and expectations for students. Assessment instruments that gather information about student achievement serve as focal points for a specific program, because they (ideally) measure those constructs that are valued within that program.

A central premise in this section was that values and practices differed among mathematics programs, confounding any attempt to measure the achievement of SIMMS and non-SIMMS students by traditional means alone. More specifically, at least four types of value-practice conflicts, described subsequently, arose in serving both groups, because the two types of curricula were so different. A second premise responded to the first by solving the problem implicit in its statement. That is, values and practices must align with assessment goals. Hence, for the purpose of this study, comparing the performance of former SIMMS and non-SIMMS students depended on the recommendations of secondary and post-secondary mathematics professional organizations and best-practice recommendations as well. Pertinent professional organizations included the National Council of Teachers of Mathematics (NCTM), the Mathematical Association of America (MAA), and the American Mathematical