THE EFFECTS OF A FRAMEWORK FOR PROCEDURAL UNDERSTANDING ON COLLEGE ALGEBRA STUDENTS' PROCEDURAL SKILL AND UNDERSTANDING

by

Jon Fredrick Hasenbank

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Mathematics

MONTANA STATE UNIVERSITY
Bozeman, Montana

April 2006

APPROVAL

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 Division of Graduate Education.

Dr. Ted Hodgson (Chair)

Approved for the Department of Mathematical Sciences

Dr. Ken Bowers

Approved for the Division of Graduate Education

Dr. Joseph J. Fedock

STATEMENT OF PERMISSION TO USE

ACKNOWLEDGEMENTS

I would like to express my gratitude to the members of my graduate committee for their accessibility and insightful comments. Special thanks goes to my chair, Dr. Ted Hodgson, for the years of support, advice, and guidance he has offered me. Thanks also to Dr. Maurice Burke for putting pen to paper several years ago to lay out his distinctive vision of what procedural understanding should look like; without those early thoughts, this dissertation would not have been written. My thanks also goes to Dr. John Borkowski for his patient guidance while I analyzed my data; to Dr. Warren Esty for his always thoughtful advice; to Dr. Linda Simonsen for her support and guidance from my very first days at Montana State University; and to Dr. Janet Sharp for our lively debates about the nature of mathematical knowledge.

I would also like to thank my dear friends and colleagues in the department where I have spent some of the most memorable years of my life. If I begin to list their names, I fear I will not find a place to stop, so I will just broadly say, "thanks for the memories." And to my parents ("who taught me to fly"), I reserve my most heartfelt thank you. Lastly, words cannot express the deep thanks I owe to my beautiful wife Kate, as much for her love and support as for her never letting me lose sight of the things that really matter in life.

You cannot help but learn more as you take the world into your hands. Take it up reverently, for it is an old piece of clay, with millions of thumbprints on it. ~ John Updike

TABLE OF CONTENTS

TABLE OF CONTENTS – CONTINUED

TABLE OF CONTENTS – CONTINUED

TABLE OF CONTENTS – CONTINUED

LIST OF TABLES

LIST OF FIGURES

GLOSSARY

Conceptual: Referring to the facts, concepts, and ideas of mathematics.

Conceptual understanding: Refers to the degree to which one understands a specific concept. See Understanding, defined below.

Deep knowledge: As contrasted with shallow knowledge, deep knowledge refers to a level of knowledge characterized by recognition of the underlying structure or guiding principles of a mathematical concept or procedure.

Experienced instructor: An experienced instructor is an instructor who has taught at any level for at least one semester.

The Framework: Used as an abbreviation for the Framework for Procedural Understanding (NCTM, 2001), which serves as the basis of the instructional treatment for this study.

Framework-oriented: Used as an adjective to signify an alignment with the objectives of the Framework for Procedural Understanding, which strives to help students develop deep knowledge of mathematical procedures.

Instructor: Refers to graduate teaching assistants (GTAs) and adjunct instructors.

Procedural: Referring to the algorithms, procedures, and processes of mathematics.

Procedural understanding: Refers to the degree to which one understands a mathematical procedure. It is operationally defined to mean the degree to which one can address the eight objectives of the Framework for Procedural Understanding (in the context of a specific procedure). See also Understanding, defined below.

Shallow knowledge: As contrasted with deep knowledge, shallow knowledge is characterized by rote memorization without significant knowledge of underlying structure or guiding principles of a concept or procedure.

Understanding: In general, a mathematical procedure or concept is understood to the extent that it is part of a network of knowledge that is: "internally coherent, consistent, logical, and rational, representative of the true procedure or concept (i.e., it is essentially correct), and connected to other things the person knows" (Greeno, 1978). See also Procedural understanding and Conceptual understanding, defined above.

ABSTRACT

This dissertation examined the effectiveness of an instructional treatment consisting of lecture content, homework tasks, and quiz assessments built around a common Framework for Procedural Understanding. The study addressed concerns about increasing numbers of students enrolling in remedial mathematics courses because they did not develop sufficient understanding in previous courses. The Framework-oriented instruction was designed to help students develop deep and well-connected knowledge of procedures, which has been shown to facilitate recall and promote future learning.

Data collection spanned the Fall 2005 semester at a western land-grant university. In the quasi-experimental design, instructors from six intact sections of college algebra were matched into pairs based on prior teaching experience, and the treatment condition was assigned to one member of each pair. Equivalence of treatment and control groups was established by comparing ACT / SAT scores for the 85% of students for whom those scores were available. Data collection consisted of classroom observations, homework samples, common hour exams scores, procedural understanding assessments, supplemental course evaluations, and a final interview with treatment instructors. Analysis of covariance was the primary statistical tool used to compare treatment and control group performances while controlling for attendance rates and pre-requisite mathematical knowledge.

Treatment group students scored significantly higher than control group students on the procedural understanding assessments. Moreover, although treatment students were assigned 18% fewer drill questions than controls and 8% fewer problems overall, the gains in procedural understanding were realized without declines in procedural skill. The relationship between understanding and skill was also examined, and students with greater procedural understanding tended to score higher on the skills-oriented final exam regardless of which treatment condition was assigned to them. Finally, the interview with the treatment instructors provided insight into the implementation issues surrounding the treatment. They expressed concerns about time constraints and reported initial discomfort with, but eventual appreciation for, using the Framework for Procedural Understanding to guide instruction.

The Framework-oriented treatment was found to be effective at helping students develop deeper procedural understanding without declines in procedural skill. Additional implications and recommendations for future research are also discussed.

CHAPTER 1

STATEMENT OF PROBLEM

Introduction

A wealth of recent evidence supports the position that American mathematics students are not learning mathematics at a deep level. Enrollment in college remedial math courses has been increasing substantially in recent years. For example, the percentage of freshmen needing remedial math at California state colleges increased from 23% in 1989 to 54% in 1998 (Schultz, 2000), and over half of Ohio high school graduates took remedial courses in college in 2002 (Kranz, 2004). Mathematics students are simply not retaining the math they have been taught. Rather than developing a sense of the logical structure of mathematics, students come to see math as a collection of rules for rewriting strings of numbers and letters that must be remembered for the next test (Kaput, 1995). The vision of mathematics education spelled out in the National Council of Teachers of Mathematics (NCTM) *Principles and Standards for School Mathematics* is quite different: "In the twenty-first century, all students should be expected to understand and be able to apply mathematics" (NCTM, 2000, p. 20).

Teaching students in a manner that promotes understanding is no easy task. The author's personal experience teaching high school algebra made it clear that high school students have come to expect algebra class to be about procedures and rules. Attempts to redirect attention to concepts and ideas are merely tolerated by students, for they know that, in the end, the test will consist of a series of procedural problems with right or

wrong answers. Changing students' beliefs about the nature of mathematics will take time. Assessment practices will need to change so that assessments are better aligned with the goals of teaching for understanding. Classroom norms that support an inquisitive, "why does this work" attitude towards mathematics must be developed and implemented. Teaching practices will need to change so that instruction models the sort of higher order thinking and reflection that lead to the development of deeper mathematical knowledge (Kaput, 1995).

As Carpenter and Lehrer (1999) wrote, "the selection of appropriate tasks and tools can facilitate the development of understanding, [but] the normative practices of a class determine whether they will be used for that purpose" (p. 26). This research project examines whether changing the normative classroom practice so that it is built around a common Framework for Procedural Understanding can lead to deeper understanding and improved performance on traditional tests of procedural skill.

Statement of the Problem

An ever-present goal of mathematics instruction (including algebra, precalculus, and calculus) is to help students to develop a well-organized collection of versatile mathematical procedures that students can call upon to solve problems in a variety of situations (Carpenter & Lehrer, 1999; Hiebert & Carpenter, 1992). Unfortunately, this goal is not often attained. Instead, students develop a cluttered collection of algorithms and procedures that they have difficulty sorting through when it comes time to select the appropriate tool for the job. Moreover, students' knowledge of procedures is often quite

shallow. While students may become quite proficient at using a procedure in the context in which they were drilled, they may have difficulty seeing how the procedure can be extended to other uses. For example, students may become quite good at factoring trinomials. However, if students solely use factoring to solve quadratic equations, they may be unaware that (a) factoring may be used for other purposes (e.g., simplifying rational functions), that (b) there are other techniques available for solving quadratic equations (e.g., the quadratic formula), and that (c) solving by factoring (over the integers) is not always possible. In other words, students do not often develop deep knowledge of procedures, even when they learn to compute with them efficiently.

A second goal of mathematics instruction has been to help students to develop knowledge of the concepts that underlie mathematics procedures. Here again, the outcome often falls short of the goal. First, many mathematics assessments focus on procedural tasks, so that conceptual knowledge is either not valued or is assumed to be present in students who score well procedurally. But procedures are often learned by rote, so inferring conceptual knowledge from the presence of procedural skills is not always appropriate (Erlwanger, 1973). In fact, it is possible to develop well-automatized procedural knowledge that is not strongly connected to any conceptual knowledge network (Hasenbank, 2005; Star, 2005). Second, it is common for classroom instruction to fall to a superficial level of mathematical activity (Henningsen & Stein, 1997), so that there is little or no emphasis placed on the conceptual underpinnings of mathematical knowledge.

In summary, students often develop only rote procedural knowledge without understanding. Students who lack understanding of memorized procedures may be able to perform quite well on tests of familiar procedural skills, but their knowledge is fragile, inflexible, and soon forgotten. After years of algorithmic mathematics instruction, a classroom culture has emerged in which students rarely pursue deep and meaningful mathematical knowledge.

Purpose of the Study

The purpose of this study was to determine the effectiveness of an instructional treatment that explicitly encourages college algebra students to develop understanding of algebra procedures through the use of six guidelines, collectively called the Framework for Procedural Understanding (or simply, the Framework), which will be described later in this chapter. This study examined the correctness of the hypothesis that as procedural knowledge becomes deeper and better connected, students will make fewer mistakes on tests of procedural skills, and their scores on such tests will improve. Specifically, the purpose of the study was to determine the extent to which the instructional treatment will promote understanding and lead to improved performance on procedural skills-oriented exams in college algebra.

If improvements in procedural performance are observed, the use of the Framework as an instructional tool will be validated. However, success on procedural tasks is not necessarily indicative of understanding, so it was also necessary to determine

whether the instruction had the intended effect of improving students' procedural understanding.

If improvements in procedural performance are not observed, then there are three possible explanations. The first reason consists of potential statistical and experimental design limitations; given that the experimental design is sound, two possible explanations for a lack of improvement in procedural performance remain. Either (a) the instructional treatment did not significantly improve procedural understanding, or (b) the treatment did improve procedural understanding but the improvement did not lead to significant improvements in performance on the procedural skills tests. The experimental design addressed each of those possibilities.

The third purpose of the study was to provide an account of instructors' perceptions of implementation issues surrounding the use of the Framework for Procedural Understanding in college algebra. These perceptions were assessed through an end-of-study interview with the three treatment instructors. The results of the interviews provided insights into the results of the quantitative analyses and provided direction for future research.

Research Questions

The instructional treatment used in this study was based on the Framework for Procedural Understanding, described later in this chapter. The Framework was designed to help students develop deep and well-connected knowledge of procedures through explicit instruction in class and through reflection outside of class.

The research questions for this study were:

1. Were there significant differences in students' performance on tests of procedural skill in college algebra between treatment and control students?
2. Were there significant differences in the depth of students' procedural knowledge of college algebra between treatment and control students?
3. What were the treatment instructors' perceptions of the overall effectiveness of the instructional treatment?

Background

Procedural and Conceptual Knowledge

The relationship between procedural knowledge and conceptual knowledge has a long history in mathematics education. In 1992 Hiebert and Carpenter noted that one of the oldest debates in the history of mathematics education is over the relative importance of conceptual knowledge and procedural knowledge. Yet, they argued, this is the wrong question to ask. "Both types of knowledge are crucial. A better question to ask is how conceptual and procedural knowledge are related" (p. 78). There is mounting evidence that suggests that conceptual and procedural knowledge develop iteratively, with advances in one type of knowledge supporting advances in the other in a hand-over-hand fashion (Rittle-Johnson, Siegler, & Alibali, 2001). An iterative model of knowledge construction fits well with the constructivist theory of learning, which holds that students construct their own knowledge based on their unique mix of pre-existing knowledge (Slavin, 1997). Models of mathematical knowledge that seek to inform mathematics

teaching and learning would therefore do well to allow for flexibility the order in which procedural and conceptual knowledge develop.

The Framework for Procedural Understanding

The Framework for Procedural Understanding was developed for use in *Navigating Through Algebra for Grades 9-12* (NCTM, 2001). The Framework consists of the following six objectives for teaching and learning mathematics:

1. The student understands the overall goal of the algebraic process and knows how to predict or estimate the outcome.
2. The student understands how to carry out an algebraic process and knows alternative methods and *representations* of the process.
3. The student understands and can *communicate* to others why the process is effective and leads to valid results.
4. The student understands how to evaluate the results of an algebraic process by invoking *connections* with a context or with other mathematics the student knows.
5. The student understands and uses mathematical *reasoning* to assess the relative efficiency and accuracy of an algebraic process compared with alternative methods that might have been used.
6. The student understands why an algebraic process empowers her or him as a mathematical *problem solver* (NCTM, 2001, p. 31, emphasis in original).

Before applying the Framework in a classroom setting, it is useful to re-express the guidelines in the Framework as a series of student-centered questions. The following eight questions are adaptations of the six objectives that make up the Framework. For the remainder of this document, it will be assumed that the Framework has eight objectives (not six), based upon the eight questions presented below.

1. (a) What is the goal of the procedure, and (b) what sort of answer should I expect?

2. (a) How do I execute the procedure, and (b) what are some other procedures I could use instead?
3. Why is the procedure effective and valid?
4. What connections or contextual features could I use to verify my answer?
5. When is this the "best" procedure to use?
6. What can I use this procedure to do?

Theoretical Framework

Constructivism

Constructivist theories of learning hold that "learners must individually discover and transform complex information, checking new information against old rules and revising rules when they no longer work" (Slavin, 1997, p. 269). Constructivist theories are rooted in work of Piaget and Vygotsky. Consequently, they advocate a social, top-down approach to learning, with the pedagogical implication that students should work in cooperative groups on challenging tasks from which they create (or are guided to discover) their own knowledge. Like most modern theories of learning, constructivists contend that instruction alone cannot deepen students' knowledge: students must be actively involved in the learning process.

It is unrealistic in a traditional, fast-paced college algebra setting to consistently give students time to work together to grapple with deep questions during class. Therefore, the treatment used in the present study was not intended to be an implementation of constructivist learning theories. However, the study's design drew

upon constructivist theories in an important way: by recognizing that students learn best when they are active learners. The study supplemented deeper classroom lectures with additional reflection-oriented homework exercises. In this way, students were passively exposed to deep questions through day-to-day classroom lectures, but they were also given an opportunity to reflect on these questions individually outside of class. That exposure and practice was then reinforced by weekly quiz questions. Overall, it was hoped that the students would develop deeper knowledge that would also help them improve their performance on tests of procedural skill.

A Model of Mathematical Knowledge

Hiebert and Carpenter (1992) reference N. L. Gage's observation in his 1963 *Handbook of Research on Teaching* that everyone works from a theory, and they differ only in the degree to which they are aware of their theory and make it available for public scrutiny. Therefore, it is important to briefly examine the theory of mathematical knowledge from which the author is working. Hasenbank's (2005) model of mathematical knowledge is a translation of the mental representation that he developed after reading and reflecting on the procedural and conceptual knowledge literature. The model owes much to the previous work by Hiebert and LeFevre (1986), Hiebert and Carpenter (1992), and especially to the more recent work by de Jong and Furguson-Hessler (1996) and Star (2000; 2005).

Star, in particular, has been a strong proponent of a movement to re-conceptualize mathematical knowledge (Star, 2000, 2002a, 2005). In 2005 (and previously in 2000), he suggested that researchers attempt to disentangle the terms procedural knowledge and

conceptual knowledge by considering the 2x2 matrix shown in Table 1. It nicely illustrates the entanglement of knowledge quality with the traditional definitions of the terms procedural and conceptual knowledge. These historic definitions stem primarily from an influential book edited by Hiebert (1986) on the subject of conceptual and procedural knowledge in mathematics. In the opening chapter of the Hiebert book, a model of mathematical knowledge is proposed in which (a) conceptual knowledge is knowledge that is understood, and (b) procedural knowledge is knowledge of the processes, algorithms, and syntactical structure of mathematics (Hiebert & LeFevre, 1986). The model assumes that conceptual knowledge is complex, well-connected, and deep by its very nature, while procedural knowledge is shallow, disconnected, and generally learned by rote. Yet contemporary research suggests that mathematics concepts ("slope is the ratio of rise to run," for example) can be memorized without being understood, and mathematical procedures can be known deeply in ways that are not conceptual in nature (see Star, 2000, 2005).

Table 1 – Types and Qualities of Knowledge

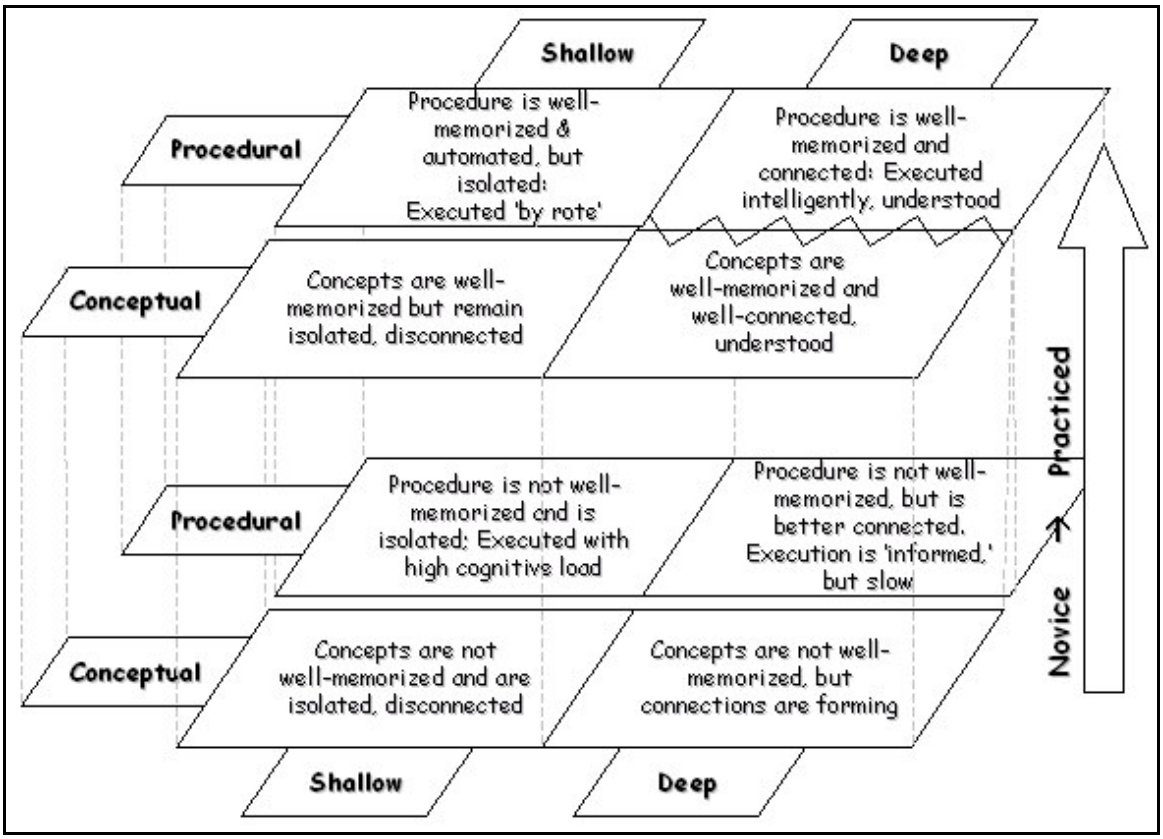
Knowledge Type	Knowledge Quality	
	Superficial	Deep
Procedural	Common usage of *procedural knowledge*	?
Conceptual	?	Common usage of *conceptual knowledge*

Note: This table was adapted from Star (2005, p. 408).

Owing heavily to Star's (2000; 2005) recommendations for filling in the unmarked cells in Table 1, Hasenbank's (2005) model also adds a third dimension accounting for students' knowledge development over time. The complete model of

mathematical knowledge, presented in Figure 1, includes three key dimensions: Knowledge Type (procedural vs. conceptual), Depth (shallow vs. deep), and Aptitude (novice vs. practiced). Note that the cells for deep, practiced knowledge have been visually stitched together to indicate that procedural and conceptual knowledge become increasingly difficult to separate at this level.

Figure 1 – Hasenbank's (2005) Model of Mathematical Knowledge



Significance of the Study

Most students who enroll in college algebra do so because they did not retain a sufficient understanding of algebra from their high school mathematics classes to meet

the prerequisites for enrollment in precalculus, calculus, or introductory statistics. James Kaput writes, “school algebra in the U.S. is institutionalized as two or more highly redundant courses, isolated from other subject matter, introduced abruptly to post-pubescent students, and often repeated at great cost as remedial mathematics at the post secondary level” (Kaput, 1995, p. 3). If instructional techniques at the college level are similar to the techniques these students experienced in high school, then similar problems with knowledge retention should be expected.

Yet when students learn mathematics with understanding, their knowledge lasts longer and it can be applied in a variety of situations (Carpenter & Lehrer, 1999, p. 21). When instruction explicitly models high-level thinking processes and strategies, students internalize them (Henningsen & Stein, 1997). Carpenter and Lehrer (1999) lists five ways in which classrooms can promote learning with understanding. They help students (a) develop appropriate relationships, (b) extend and apply their mathematical knowledge, (c) reflect about their own mathematical experiences, (d) articulate what they know, and (e) make mathematical knowledge their own. The Framework, when used to guide instruction and assessment in college algebra, has the potential to provide such opportunities. The explicit intent of the Framework for Procedural Understanding is to help algebra students understand the procedures they are learning. Consequently, the use of the Framework has the potential to help algebra students break the cycle of learning and forgetting that characterizes the current state of algebra instruction in the United States (Kaput, 1995).

CHAPTER 2

REVIEW OF LITERATURE

Introduction

This review of literature begins by laying out the theoretical framework of the study and describing the author's model of mathematical knowledge and its development from the literature base. Next, a review of research is presented which includes an examination of the historical distinction between procedural and conceptual knowledge, an analysis of classroom norms that promote the development of mathematical understanding, a discussion of how procedural or conceptual knowledge develop, and methodological considerations for assessing understanding. The chapter concludes with an overview of the research findings.

Theoretical Foundation

Constructivism

Slavin's (1997) text on educational psychology reports that constructivism grew out of the work of Piaget and Vygotsky, both of whom held that cognitive change grows out of a process of disequilibrium as new information is encountered that does not fit into existing mental structures. At its core, constructivism holds that students must "individually discover and transform complex information, checking new information against old rules and revising rules when they no longer work" (Slavin, 1997, p. 269). Constructivist theories reserve an extremely active role for the learner, and while the

treatment used in the present study was not an implementation of constructivist theories per se, constructivist theories are relevant to the extent that they recognize students learn best when they are actively involved in the learning process.

The study was based upon a three-fold instructional treatment. First, traditional college algebra lectures were supplemented with Framework-oriented ideas. In this way, lecture content was supposedly deepened. However, constructivist theories suggest that deepening the relatively passive lecture portion should have little impact by itself. In recognition of that belief, additional reflection-oriented homework exercises were included in the treatment. Later, some of these exercises also appeared on weekly quizzes. In this way, enhanced lecture content was reinforced by giving students the opportunity (and the incentive) to reflect on deep questions outside of class.

The Nature of Mathematical Knowledge

Mathematical knowledge has traditionally been classified along a single dimension: procedural knowledge vs. conceptual knowledge (Star, 2000, 2005; Wu, 1999). However, reading the literature pertaining to procedural and conceptual knowledge, the author developed a sense that the operational definitions of procedural and conceptual knowledge, as exemplified by the existing models, did not completely capture all relevant aspects of mathematical knowledge. For example, Hiebert and Carpenter's (1992) model holds that there can be no isolated bits of conceptual knowledge. Conceptual knowledge, they said, is knowledge that is understood. Yet mathematics concepts ("slope is the ratio of rise to run," for example) can be memorized without being understood. Similarly, it is possible to have a deep knowledge of

mathematical procedures that is not conceptual in nature (see Star, 2000). Finally, existing models do not include a developmental component so that educators and researchers can track a student's progress.

Star (2005) makes a strong case for rethinking the traditional meanings associated with the terms procedural and conceptual knowledge. Drawing upon an earlier, much more complex model set forth by de Jong and Ferguson-Hessler (1996), Star laid out a two-dimensional matrix illustrating the historical "entanglement" between two dimensions of knowledge. Knowledge type (procedural vs. conceptual), Star suggested, is fully independent of knowledge depth (shallow vs. deep). The author's model of mathematical knowledge extends Star's matrix to include a dimension for representing students' developing knowledge. Thus, as Figure 1 (p. 11) illustrates, the author contends that mathematical knowledge consists of three key dimensions: Knowledge Type (procedural vs. conceptual), Depth (shallow vs. deep), and Aptitude (novice vs. practiced). Together, these dimensions form a complete yet parsimonious way to classify students' mathematical knowledge.

As an illustration, consider how a student's knowledge about a new topic, say three-digit by two-digit multiplication, might unfold. Classroom instruction might begin by illustrating the correct procedure through a series of examples. The student's procedural knowledge begins to develop, but with little practice and little time for reflection, her knowledge would be at the novice, shallow level. Initially, she finds it difficult to recall or reproduce the steps of the procedure on demand, and her execution of the procedure is prone to errors and requires substantial cognitive load. With practice, she

finds that the steps of the procedure become more easily accessible, and ultimately they become a well-automated sequence that can be performed almost mindlessly. Once this happens, the student's procedural knowledge of the algorithm has moved up along the aptitude dimension, and she now has shallow, but practiced, procedural knowledge of that topic.

As her cognitive load decreases during execution of the procedure, it becomes possible for the student to think more deeply about the procedure while she is applying it. She can (perhaps with guidance) begin to estimate her answer in advance and check the reasonableness of her answer when she finishes the algorithm. She may also notice patterns in the results (such as more efficient ways to handle multiplication by multiples of 10 or 100) that might further increase her efficiency and help her detect or prevent errors. The depth of her knowledge of the algorithm grows as the procedure becomes more automated.

Similarly, as the procedure becomes more automated, she is able to notice similarities and differences between the current procedure and other algorithms she has practiced in the past. With reflection, such observations can be incorporated into existing conceptual knowledge networks (e.g., knowledge of general multiplication and place value concepts), so that advances in procedural knowledge along both the aptitude and depth dimensions can promote advances in conceptual knowledge, just as Rittle-Johnson et al. (2001) have suggested. Moreover, these advances in the student's conceptual knowledge may trigger additional observations that further deepen her procedural knowledge, and links may form between her procedural and conceptual networks. As

procedural knowledge and conceptual knowledge become deeper, it becomes more and more difficult to tease them apart.

From the preceding illustration it would be easy to get the impression that procedural aptitude (or automation) must come before deep procedural knowledge can develop. In fact, that is probably the most efficient way to proceed because cognitive resources are freed up as a procedure becomes better automated (Wu, 1999). However, constructivism advocates top-down instruction that attempts to help students construct their deep procedural knowledge from the very first days of practice (Hiebert & Carpenter, 1992). In this way, students develop a sort of rudimentary deep procedural knowledge even before they have developed the ability to carry out a procedure efficiently. Consider, for instance, a college student's growing knowledge of first-year differential calculus, or a master's student's emerging knowledge of modern algebra, or a doctoral student's developing knowledge of residue theory. It is possible for knowledge of such complex procedures to never become completely automated, even with frequent practice, despite the development of fairly deep procedural knowledge that allows answers to be predicted in advance and allows the procedure to be applied in a variety of contexts. Star (2000) has suggested that one of the reasons deep, non-conceptual knowledge of procedures has been overlooked for so long is that most of the existing research has focused on elementary school topics that are too simple to allow for such knowledge. As Star wrote, "It *is* difficult to conceive of having understanding of a procedure when one only considers the relatively simple and short procedures learned in elementary school" (Star, 2000, pp. 5-6, emphasis in original).

From the preceding discussion, it appears that procedural knowledge can grow deeper without necessarily being connected to a conceptual knowledge network and that procedural knowledge can draw upon and promote advances in conceptual knowledge. Next, consider a short example of how a unit of conceptual knowledge might develop. Suppose a lesson is designed to help students learn about the distributive property of multiplication over addition. The distributive property serves as yet another example of conceptual knowledge that may either be memorized at a shallow level or be understood at a deep level. Instruction may begin with a symbolic representation (such as $a(b + c) = ab + ac$), with a visual representation (such as using an appropriately divided rectangular area), or possibly with connections to number facts (such as $37\times4 = 30\times4 + 7\times4$). As instruction unfolds, a student's conceptual knowledge may grow along one of two dimensions. Perhaps her knowledge of the visual pattern $a(b + c) = ab + ac$ will become quite strong, so that her knowledge develops into shallow (disconnected), but well memorized knowledge. On the other hand, perhaps a rudimentary conceptual network begins to emerge, possibly containing a number of misconceptions such as the well known error of over-generalizing the distributive property to include situations such as $(a + b)^2 = a^2 + b^2$. In this way, the student's knowledge moves first along the depth dimension before it becomes well memorized. In time, misconceptions are corrected and the links in the conceptual network are strengthened, so that the student's knowledge moves up the developmental dimension and becomes not only deep, but also well memorized. The proposed model of mathematical knowledge allows for two independent

dimensions of mathematical growth in the early stages of learning about mathematical procedures and concepts (Hasenbank, 2005).

Review of Related Research

The following review of research is divided into three parts. The first part consists of an analysis of research literature on procedural and conceptual knowledge, including operational definitions of terms, theoretical and empirical evidence of the benefits of learning mathematics with understanding, a list of classroom norms that support the development of understanding, and a discussion of research that suggests that procedural and conceptual knowledge develop iteratively and that understanding should be emphasized at each stage of learning. The second part consists of research related to the use of writing and self-reflection in mathematics education, with emphasis on their use as tools for developing deeper mathematical knowledge. The third part considers methodological issues pertaining to the assessment of mathematical understanding.

Procedural and Conceptual Knowledge

For the purposes of the current research, procedural knowledge is operationally defined as knowledge of the algorithms, procedures, and processes of mathematics, while conceptual knowledge is operationally defined as knowledge of the facts, concepts, and ideas of mathematics. In light of the Hasenbank's (2005) three-dimensional model of mathematical knowledge and Star's (2005) two-dimensional model, the traditional connotations of procedural knowledge as shallow rote knowledge and of conceptual

knowledge as deeply understood knowledge are inappropriate. Each type of knowledge may be known at either a shallow or a deep level.

The depth of a student's knowledge is operationally defined as the extent to which the student is aware of the underlying structure and guiding principles of that knowledge. For example, a student with shallow procedural knowledge of derivatives may be quite adept at computing derivatives while being unaware of the fact that the product rule will suffice any time the quotient rule is applicable, or that the constant rule ($\mathrm{d}/\mathrm{d}x\,[a\,f(x)] = a\,\mathrm{d}/\mathrm{d}x[f(x)]$) is just a special case of the product rule. Similarly, a student with shallow conceptual knowledge might have a well developed mental image of the definite integral as "area under the curve" without being aware of its relationship to the anti-derivative of a function.

Each of the preceding examples could also be seen as an example of a partially understood bit of knowledge. Greeno (1978) defined understanding of a mathematical procedure or concept in terms of the extent to which it is part of a network of knowledge that is internally coherent, representative of the true procedure or concept (i.e., it is essentially correct), and connected to other things the person knows. Lately, the term understanding has been used in such a variety of situations and with such a variety of subtly differing operational definitions that it is difficult to use the term with any precision (compare, for example the definitions in: Carpenter & Lehrer, 1999; Glaser, 1984; Ohlsson & Rees, 1991; Rittle-Johnson et al., 2001; Wiggins & McTighe, 2001; Wu, 1999).

It is likely that the abundance of operational definitions in the literature is due to a general difficulty in assessing understanding (Wiggins & McTighe, 2001). Because the current study focuses on enhancing students' procedural understanding through the use of the Framework for Procedural Understanding (NCTM, 2001), it is appropriate to use the Framework to operationally define an "understanding of procedures." Unfortunately, by choosing to operationally define procedural understanding in this way, the present study adds yet another meaning to the term procedural understanding.

Despite the overuse of the term, the general meaning of understanding is apparently closely related to the author's definition of depth, which refers to one's knowledge of the underlying structure and guiding principles of a procedure or concept. The eight objectives of the Framework for Procedural Understanding scaffold these "underlying structures and guiding principles," so that the Framework encompasses what is typically referred to as understanding. In fact, most existing research has used the term understanding only in association with conceptual knowledge, and until recently, very few articles have made reference to deeply known procedural knowledge (de Jong & Ferguson-Hessler, 1996; Star, 2000, 2005). Therefore, in the review of literature which follows, the word understanding shall be used almost exclusively to reflect its common usage in the existing literature base.

Benefits of Understanding

The Learning Principle of the *Principles and Standards for School Mathematics* (NCTM, 2000) is "based on students' learning mathematics with understanding" (p. 20). When a unit of mathematical knowledge is understood, that knowledge is more easily

remembered (Carpenter & Lehrer, 1999; Hiebert & Carpenter, 1992; Van Hiele, 1986) and it can be applied more readily in a variety of situations (Hiebert & Carpenter, 1992; Kieran, 1992). This fits well with connectionism (Bereiter, 1991, 2000) and with results from cognitive psychology that show that knowledge best remembered is knowledge that is meaningful to the learner: "Things take meaning from the way they are connected with other things" (Carpenter & Lehrer, 1999, p. 20). When knowledge is part of a well-connected network, parts of the network can facilitate recall of other pieces of the network, thereby making it easier to recreate knowledge that may have been forgotten. Finally, when knowledge is understood it becomes easier to incorporate new knowledge into the existing structure, so that current understanding facilitates future learning (Hiebert & Carpenter, 1992).

Classroom Norms that Promote Understanding

Research points to the benefits of developing deep mathematical knowledge, but how can such depth be promoted in the classroom setting? It is easy for classroom instruction to fall to a low level of mathematical work, even in instructional programs that claim to engage students in high-level tasks (Henningsen & Stein, 1997). When this occurs, students often fail to develop sufficient understanding of the procedures they are learning and resort to memorizing patterns and rules in order to overcome (or hide) their lack of understanding (Kieran, 1992).

By establishing classroom norms that provide students with opportunities to (a) develop appropriate relationships, (b) extend and apply their existing knowledge, (c) reflect on their experiences, (d) articulate what they know, and (e) make mathematical

knowledge their own, mathematics classrooms can help students to develop deep mathematical knowledge (Carpenter & Lehrer, 1999). Such instruction aligns well with a constructivist approach toward learning, as well as with the vision of mathematics teaching set forth in the *Principles and Standards for School Mathematics* (NCTM, 2000).

It is important that classroom norms for promoting understanding be established and maintained for an extended period of time. Carpenter and Lehrer (1999) wrote that it takes time for students to develop knowledge structures.

> Understanding is not an all-or-none phenomenon. Virtually all complex ideas or processes can be understood at a number of levels and in quite different ways. Therefore, it is appropriate to think of understanding as emerging or developing rather than presuming that someone either does or does not understand a given topic, idea, or process (p. 20).

Like most reformed instructional techniques, a steady, consistent approach is the only way to see results.

Which Type of Knowledge Develops First?

A decade of research regarding the development of mathematical knowledge produced conflicting evidence regarding whether procedural or conceptual knowledge develops first (Rittle-Johnson et al., 2001). In fact, there is a growing consensus that procedural and conceptual knowledge develop iteratively, in a sort of hand-over-fist fashion, so that advances in one type of knowledge lead to advances in the other, and vice-versa (Rittle-Johnson & Koedinger, 2002; Rittle-Johnson et al., 2001). For example, limited conceptual knowledge in a domain might lead a student to focus more attention on certain aspects of a procedure, and after sufficient practice the subsequent advances in

procedural knowledge free up cognitive resources that allow the learner to notice broad conceptual similarities between the procedure and its underlying mathematical basis.

Regardless of whether one type of knowledge develops first or whether, as the evidence increasingly supports, the two types develop iteratively, Carpenter and Lehrer (1999) point to "a mounting body of evidence" in support of teaching for understanding from the very beginning. One of the reasons for teaching for understanding before teaching for rote skill is that through practice it is possible for students to develop misconceptions or systematic errors that can be very difficult to "unlearn." Cognitive psychologists use the terms *functional fixedness*, *cognitive narrowing*, or *tunnel vision* (Norman, 1993) to refer to a learner's inability to extend the use of a tool beyond the routine tasks that they have practiced. Even after students have mastered a procedure, they are often unable to transfer what they have learned to new tasks despite strong similarities between the tasks (Kieran, 1992).

Rittle-Johnson and Koedinger (2002) compared an iterative instructional model (in which procedural and conceptual instruction were interwoven) and a concepts-first instructional model (in which conceptual lessons preceded procedural lessons). The lessons were presented by a computer-based intelligent tutoring system to 72 sixth-grade students who were learning about decimal place value and regrouping concepts along with procedures for adding and subtracting decimals. Their findings showed a significant main effect for condition ($F(1, 72)$[1] $= 3.92$, $p < 0.06$), representing significantly greater

---

[1] The original paper reported that the degrees of freedom for the error term was 72, yet this appears to be inconsistent with the reported sample size of 72.

gains from pretest to posttest for students in the iterative condition than students in the concepts-first condition. Moreover, there was an interaction effect between condition and knowledge type ($F(1, 72)$[2] $= 4.91$, $p < .03$), so that the effect of condition was mostly associated with the procedural knowledge items. The authors interpret these findings to mean that iterative instruction highlighted the links between conceptual and procedural lessons, and that these connections helped students overcome common procedural errors.

Writing in Mathematics

The use of writing in mathematics provides students with opportunities to actively reflect on what they know. Such reflection is a valuable part of the constructivist's view of the learning process, which holds that the learner must discover and actively transform complex information if they are to make it their own (Slavin, 1997). Writing in mathematics helps students become less passive learners and to think more critically about the content they are learning. When students write about mathematics content, they must relate, connect, and organize their knowledge while discovering relationships between the ideas and drawing on connections with their own prior knowledge (Brozo & Simpson, 1999). Writing also provides opportunities for students to analyze arguments, compare and contrast ideas, and pull together knowledge into a more cohesive knowledge base, and all of this can be done in a safe, non-public forum, possibly as a prelude to a classroom discussion (Dougherty, 1996). These activities are all in alignment with classroom norms that promote mathematical understanding (Carpenter & Lehrer, 1999).

---

[2] Again, the degrees of freedom appear to have been reported incorrectly in the original paper.

Assessing Understanding

Understanding is notoriously difficult to assess (Hiebert & Carpenter, 1992). Wiggins and McTighe (2001) capture the nature of the difficulty:

> Understanding could be sophisticated in the absence of a good explanation. The quality of the insight may be less *or* more than the quality of the explanation or performance in general…. To judge understanding of how someone makes sense of something, we typically need the person to explain it to us (p. 80, emphasis in original).

Understanding cannot be assumed to be present simply because a student has performed well on a single task: it is possible to consistently perform a task correctly without understanding the underlying mathematics (Hiebert & Carpenter, 1992). Interviews, journals, and other writing prompts are therefore natural tools for assessing students' understanding.

Erlwanger (1973) used problem-based interviews to examine a sixth-grade student's (Benny's) understanding of decimals and fractions and his beliefs about the nature of rules, relationships, and answers in mathematics. Benny was among the top students in his mathematics class, which used the Individually Prescribed Instruction (IPI) Mathematics program. Benny had progressed through the levels of the IPI program at a faster rate than many of his peers, and so it was assumed his understanding of mathematics was progressing as well. In fact, Erlwanger's problem-based interviews with Benny revealed that Benny had developed serious misconceptions about not only decimals and fractions but also about the very nature of mathematical activity.

Problem-based interviews based on Framework-oriented questions were considered for the present study because they would provide the greatest level of detail

about students' procedural understanding. Unfortunately, using problem-based interviews as a data collection method is very time intensive, and it would be unrealistic to sample more than a handful of the approximately 150 students who took the final exam. Without data on understanding from a majority of students, it would have been difficult to reliably compare treatment and control students' understanding.

The process of designing curricula and assessments for understanding should begin with the question, "What would count as evidence of successful teaching?" (Wiggins & McTighe, 2001, p. 63). Because procedural understanding has been operationally defined in terms of the Framework for Procedural Understanding (NCTM, 2001), assessments of procedural understanding should be derived from the eight learning objectives of the Framework. Wiggins and McTighe also warn that "understanding could be sophisticated in the absence of a good explanation," and that to judge a person's understanding of a topic, "we typically need the person to explain it to us" (Wiggins & McTighe, 2001, p. 80). Therefore, it was determined that a series of Framework-oriented journal prompts would provide the broadest and the most useful evidence of students' procedural understanding.

Summary

Mathematical knowledge of procedures and concepts develops in a hand-over-hand fashion, with advances in one type of knowledge leading to advances in the other (Rittle-Johnson & Koedinger, 2002; Rittle-Johnson et al., 2001). As these types of knowledge develop, it is important to encourage students to develop deep knowledge

(understanding) of the mathematics so that benefits such as improved memory, increased transfer, facilitation of future learning, increased flexibility, and reduced rates of nonsensical errors will be realized. Writing tasks can serve as a basis for mathematical reflection, which in turn can lead to the development and assessment of deeper mathematical knowledge. Mathematical understanding cannot be assessed through any individual task; therefore, it was determined that a sequence of journal tasks designed around Framework-oriented objectives would be used to assess procedural understanding. The review of literature just presented forms the foundation upon which the methodological considerations addressed in the next chapter were derived.

CHAPTER 3

METHODOLOGY

Introduction

This chapter begins by restating the research questions that appear in Chapter 1, and then proceeds to describe the design of the study, including both the quasi-experimental nature of the quantitative component as well as the debriefing session held with treatment instructors at the end of the study period. The population and sampling procedures are then described, followed by a discussion of the pilot study during which the instructional materials were developed. The remainder of the chapter contains a detailed description of the data collection procedures, including an annotated timeline and a description of the data analysis methods. Finally, a list of assumptions and limitations pertaining to the study is presented.

Research Questions

The instructional treatment was based on the Framework for Procedural Understanding. The Framework was designed to help students develop deep and well-connected knowledge of procedures through explicit instruction in class and through students' reflections outside of class.

The research questions for this study were:

1. Were there significant differences in students' performance on tests of procedural skill in college algebra between treatment and control students?
2. Were there significant differences in the depth of students' procedural knowledge of college algebra between treatment and control students?
3. What were the treatment instructors' perceptions of the overall effectiveness of the instructional treatment?

Overview of Methodology Used

The study primarily employed quantitative methodology to examine the relationships between procedural skill and procedural understanding. In addition, instructors' perceptions of the instructional treatment were assessed using qualitative methods in the form of an end-of-study interview. The results of the interview were used to guide recommendations for future research and to provide insights into the results obtained from the quantitative analyses.

The quantitative component employed a semester-long quasi-experimental design using intact groups of college algebra students. The six experienced instructors teaching college algebra during the study period were recruited, divided into three matched pairs based on prior teaching experience, and assigned to either the treatment or control condition. While common in educational research, the use of intact groups of students nonetheless presents a limitation to the external validity of the study.

The qualitative component consisted of a single interview with all three treatment instructors. The researcher's advisor was also present. The researcher used a semi-

structured interview script to guide the interview, which was audio taped and later transcribed by the researcher. The transcription covered 36 minutes of discussion.

Population and Sample

Student Population

The population of interest for the study consisted of undergraduate students enrolled in college algebra at Montana State University, a research-intensive land grant university located in the northern Rocky Mountains. Because college algebra does not satisfy the university's core mathematics requirement, students who enroll in college algebra typically do so in order to prepare themselves for future mathematics or statistics coursework. These students are typically freshmen in college who have not satisfied the prerequisites necessary for precalculus, calculus, or introductory statistics.

Student Sample

The sample was drawn from the population of college algebra students at Montana State University. The instructors selected for the study were already assigned by the mathematics department to teach a specific section of college algebra, and these teaching assignments determined which college algebra students participated in the study. Therefore, intact groups of students were used for this study.

Instructor Population

College algebra instructors are typically either adjunct instructors or graduate teaching assistants from the mathematics department. Adjuncts often teach more than one

course per semester at the university, which may or may not include multiple sections of college algebra. Adjuncts typically have earned a master's degree and have substantial teaching experience.

Graduate teaching assistants (GTAs) teach one section of college algebra (or another course) in exchange for a stipend and a tuition fee waiver for their graduate coursework. These graduate students take between 6 and 9 credits of graduate coursework while preparing lecture notes, writing quizzes, and grading homework for their section.

Instructor Sample

At the start of the study period, the researcher identified six college algebra instructors who had prior teaching experience, meaning they had one or more semesters of teaching experience at any level. The six instructors were recruited via telephone during the week before the semester began. All agreed to participate, although one instructor expressed some concern about the extra work it might require.

In order to control for teaching experience, the six instructors were divided into three matched pairs. The Highly Experienced pair of instructors were both adjuncts. Of these, one was the course supervisor and the other had been a long-time high school math teacher at a local high school. The course supervisor taught two sections of college algebra; only one of them was selected for inclusion in the study. The Moderately Experienced pair of instructors were both GTAs who had each been teaching for more than two semesters. The Less Experienced pair of instructors were GTAs who had only one semester of teaching experience.

After matching instructors based on experience, the researcher intended to randomly assign the treatment condition to one member of each pair. He attended the first course supervisor meeting the week before school began and asked the six participating instructors to remain briefly after the meeting concluded. The researcher explained the anticipated responsibilities of treatment and control instructors. The one instructor (Moderately Experienced) who had initially expressed concerns about the extra work stated that he felt the weekly treatment instructor meetings would represent a greater time investment than he was willing to give, and he opted out of being a treatment instructor. His matched instructor agreed to participate as a treatment instructor, and so random assignment of treatment was not attempted for this pair. The other instructors agreed to random assignment of the treatment, and a subsequent coin toss accomplished the random assignment. It should be noted that the course supervisor was (randomly) assigned to the control condition.

Time of Day and Room Assignments

Table 2 displays the meeting times of the six college algebra classes selected for participation in the study.

Table 2 – Class Meeting Times for Participating Groups

	Instructor Experience		
	High	Moderate	Less
Treatment	8 am	10 am	12 pm
Control	11 am	10 am	9 am

All but two of the classes met in the same room on campus at different times throughout the day. The two classes that did not meet in the common room were the treatment

sections offered at 10am and 12pm. The meeting times and classroom assignments were scheduled by the university well before assignment of treatment condition was attempted.

Course Description

College algebra is a three credit hour class offered every semester at Montana State University. Twelve sections of college algebra, taught by 11 different instructors, were scheduled in the Fall 2005 semester. The description of college algebra as it appears in the Montana State University Bulletin (2004) is: "Algebra for College Students. Further development of algebraic skills through the study of linear, quadratic, polynomial, exponential, and logarithmic functions."

To enroll in college algebra, students are expected to have earned an SAT Math score of at least 530 or an ACT Math score of 23, or to have satisfactorily completed a previous course in intermediate algebra. Students who do not meet one of those requirements but wish to enroll in college algebra are required to pass the math placement exam that is administered by the mathematics department at the start of each semester.

Pilot Study

From May 16 to June 23, 2005, the researcher taught a six-week summer session of college algebra. During this period, the researcher developed curriculum materials including homework assignments, quiz questions, and lecture examples that reflected a Framework-oriented approach to learning college algebra. In addition, the pilot study period allowed the researcher to become familiar with the content of the curriculum and format of the Blitzer (2004) textbook, as well as the types of errors, quality of work,

academic backgrounds, and mathematical attitudes and beliefs that could be expected from college algebra students.

Design of the Study

The Instructional Treatment

Three sections of college algebra received specialized instructional treatment based upon the Framework for Procedural Understanding. The treatment included Framework-oriented lecture content, homework tasks, and quiz items. Each of these aspects of the treatment is discussed in turn below.

Lecture Content. The researcher met with the three treatment group instructors every Monday morning of the semester. Each meeting was centered around a packet of instructor notes developed by the researcher. Each packet was divided into separate sections representing the three lessons to be delivered in the coming week. For each of these three lessons, the packet included a) the section objectives from the required Blitzer (2004) algebra text, b) the assigned homework problems, c) the course supervisor's instructor notes, and d) a series of Framework-oriented examples developed by the researcher. At the end of each packet was a section containing several (typically four or five) sample quiz questions. As an example, the packet from Week 5 is included as Appendix A.

Instructors were expected to incorporate most of the provided examples into their daily lectures and to include a subset (typically two) of the provided quiz questions on their quiz for that week. Instructors were also expected to supplement the provided

examples and quiz questions with some of their own. In this way, the instruction was supplemented with the Framework while control of the day-to-day workings of the course ultimately remained with the instructors.

Quiz Questions. Quiz questions were included in the treatment in order to reinforce the importance of students completing the assigned Writing questions (see below). In some cases, the quiz questions provided to treatment instructors were direct copies of a recently assigned Writing question. In other cases, the quiz questions reflected topics that had been emphasized in the lecture notes from the previous week. Grading keys were not supplied by the researcher. Instructors were given discretion to select quiz questions from those provided that they felt best matched their instruction. To ensure that treatment instructors used the recommended quiz questions, all six participating instructors were asked to submit a blank copy of their quiz each week.

Homework Tasks. Students from all non-treatment sections of college algebra were given a complete list of homework questions during the first week of the course. It was divided into book sections, and the exercise set for any given section typically included between 20 and 25 problems, of which fewer than two (and quite often none) were taken from the Writing in Mathematics subsection of each homework set. These Writing questions asked students to reflect on or explain the mathematics they were learning, and many of them were aligned with the instructional Framework implemented in this study.

The Writing questions served as the basis for the homework component of the instructional treatment. Students in the treatment sections were given a modified homework list supplemented with (on average) 3.3 Writing questions from each section. To offset any potential extra time requirements of the Writing questions, the treatment homework list contained fewer problems overall. The original list distributed to treatment students contained 738 total homework questions, of which 96 were Writing questions. The original list given to control students had 802 total questions, of which 15 were Writing questions. Thus, treatment students were assigned 18% fewer drill questions than controls and 8% fewer problems overall. Occasionally, the course supervisor saw fit to make minor modifications to the control students' list during the semester. The treatment list was adjusted to compensate, so that the initial ratios were approximately maintained.

The Writing in Mathematics exercises were listed before all other exercises and were offset with a "W" (for "Writing"). Treatment instructors were told to remind students to complete the Writing problems before attempting the others, under the pretext that he or she believed they were important questions and that some of the questions would serve as the basis for quiz questions. The complete homework lists for both control and treatment students can be found in Appendix B.

Data Sources

The differing level of emphasis given to Framework-oriented instruction in treatment and control groups served as the primary independent variable for the study. The corresponding dependent variables were procedural skill and procedural understanding.

Four distinct groups of data were collected. The first group of data measured the primary outcome variables of the study. Procedural skill was measured in terms of students' scores on the three common hour exams and the final exam, and procedural understanding was assessed with a series of six Framework-oriented journal tasks given in pairs during the last three weeks of the semester. The second group of data measured fidelity of implementation, or the degree to which the three components of the instructional treatment (lecture, homework, and quizzes) were successfully carried out. Classroom observations were the main source of data in this area, but secondary sources included homework samples, quiz samples, and a supplemental course evaluation form. The third group of data consisted of background information about students, including attendance rates and ACT / SAT scores. The fourth group of data consisted of the transcribed end-of-semester interview conducted with the three treatment instructors. Each of these four groups of data is discussed below.

Procedural Skill – Exam Scores. The common hour and final exams were written by the Math105 course supervisor without direct knowledge of the specific nature of the instructional treatment. It should be noted that the course supervisor was one of the instructors assigned to the control condition. The course supervisor agreed to write the exams with the intention of assessing computational skill. All Math105 instructors shared in the exam grading process, including those inexperienced instructors who did not otherwise participate in the study. Each instructor was responsible for grading a small subset of the exam questions across all sections to ensure consistency in grading. After each exam, the six participating instructors submitted a list of their students' individual

scores to the researcher, who then recorded them in a master gradebook for later analysis. These scores were identified with each student by first and last name.

Procedural Understanding – Writing Tasks. The researcher developed six journal tasks for assessing students' procedural understanding. Because procedural understanding has been operationally defined in terms of the Framework for Procedural Understanding, each of the six tasks was designed to reflect one of the Framework objectives. In this way, every objective was sampled once[3]. The procedures used to provide the context for the journal tasks varied from week to week and were selected from recent coursework whenever possible. These tasks collectively provided a measure of students' procedural understanding.

The six tasks were divided into three two-question quizzes given during the first 10 minutes of class on Monday of Week 13, Wednesday of Week 14, and Monday of Week 15. Students from all six participating sections were informed in advance that these six journal tasks would be combined into one quiz grade, and that they would receive full credit on each task as long as they made an honest effort to complete them. The six tasks appear in their original form in Appendix C. Task-specific rubrics used to score each task are provided in Appendix D.

Fidelity of Implementation – Classroom Observations. Classroom observations were conducted to assess the instructional emphasis on procedural understanding and to

---

[3] Of the eight objectives, items 1a and 1b were collapsed into a single task, and item 2a (performing the procedure) was omitted.

measure the extent to which the treatment and control instruction differed. Prior to the first observation in Week 2, the researcher developed an observation form, trained four peers, and established inter-rater reliability through pilot observations of a precalculus class taught by an experienced college professor. The training session consisted of a joint meeting between the researcher and three reviewers during which the researcher read through a fictional classroom scenario as reviewers attempted to code individual events. The fourth reviewer, who did not attend the original meeting, was trained separately prior to the pilot observation of the precalculus class. The classroom scenario was developed by the researcher specifically for the training session, and it is available in Appendix F.

Observers worked in pairs to produce independent assessments of the Wednesday classes in each of Weeks 2, 6, and 10. The two-page observation form used to guide observations is available in Appendix G. The first page of the form was used to record observations with tally marks to record and classify each Framework-oriented event they observed. At the end of the observation period, each observer reviewed his or her tally marks and recorded a holistic score for each of the eight Framework dimensions on the second page of the observation form. Each pair of ratings was averaged to form a single assessment for the observation along each of the eight dimensions of the Framework.

In order to cultivate a non-threatening atmosphere of observation, the observations were announced in advance either orally or via email to all participating instructors. During classroom observations, observers were directed to remain inconspicuous (i.e., to sit near the back and act as if they were regular students in the class) and to sit apart from one another so that they could produce independent ratings. In

at least two instances, some observers were forced to interact with students when group tasks were assigned or when students noticed they were not regular members of the class. A cover story was given in those instances to preserve students' blindness to the experiment. In two other instances, the researcher assumed the role of a helpful student and assisted a struggling instructor (one treatment, one control) by asking a leading question (e.g., "what if we try to … instead") in finding a mistake that had been made. It should be noted that the course supervisor had a policy of observing each instructor on at least one occasion during the semester. In no instance was it felt that the observers' true role was exposed to the class, nor did the observers' presence unduly influence class instruction.

Fidelity of Implementation – Homework Tasks. In addition to classroom observations, three secondary measures of fidelity of implementation were part of the design of the study. The secondary measures were compared with the results of the classroom observations in an effort to triangulate the results.

First, the researcher examined homework samples from the treatment sections during Weeks 3, 7, and 11. Because the course supervisor required all instructors to collect all assigned homework every week, the students were not made aware that some homework samples were being saved for analysis. The Writing questions from those homework tasks were coded by the researcher using the following general scoring rubric.

0 – No response (problem was omitted or not attempted).

1 – Terse and Incorrect / Incomplete

2 – Terse but Essentially Correct

3 – Verbose but Incorrect / Incomplete

4 – Verbose and Essentially Correct

Terse responses were often brief or vague, used incomplete sentences or poor grammatical structure, and were apparently written hastily and without much reflection. Verbose responses, on the other hand, used clearer language, were more specific, were typically written in complete sentences, and often used key words or phrases from the original question. Note that while a level 4 response should be considered better than a level 3, and a level 2 better than a level 1, one should not presume that a level 3 response is better than a level 2. Terse and verbose responses are qualitatively different, so that a lengthy, well-written (but ultimately incorrect) response may be coded level 3 and yet be quite inferior to most level 2 responses. The coding scheme was designed primarily for the purpose of assessing the degree of attention and effort students put forth on the Writing in Mathematics homework tasks. The correctness of the responses was considered to be of secondary importance.

The following examples illustrate how the scoring rubric was used in the study. One of the journal tasks asked students to "explain how to determine which numbers must be excluded from the domain of a rational expression." One student's response read: "Because rational expressions indicate division and division by zero is undefined, we must exclude numbers from a rational expression's domain that make the denominator zero." A second student wrote: "The numbers that need to be excluded are the ones that would make the denominator zero and the answer Ø [the empty set]." Both responses

were clear and made reference to the original question, the sentence structures were sound, and the answers were essentially correct. Both were coded level 4.

A third student's response read: "The numbers that will be excluded from the domain are the numbers that are [in] the denominator." The response was written using proper sentence structure and made reference to the original question, but was incomplete. A fourth student's response was: "They are the numbers, [*sic*] that when substituted in the domain for x equal 0." While it did not correctly answer the question, the response used key words from the original question, was written as a complete sentence, and used (nearly) correct grammatical structure. Both responses were coded level 3.

A fifth student's response read: "Whatever make [*sic*] the equation in the denominator can't [*sic*] equal zero." The illogical sentence structure suggests a lack of care and attention, and there is no direct reference to the original question. It was apparent, however, that the student had the right idea. The response was coded level 2.

A sixth student's response read: "#'s that make $x = 0$ must be excluded." The response is brief, makes little mathematical sense, and is ultimately incorrect. It was coded level 1.

<u>Fidelity of Implementation – Quiz Tasks.</u> Quiz samples were collected from treatment sections during Weeks 2, 6, and 10. However, it was decided that the evidence for fidelity of implementation gleaned from these tasks would not add significantly to the information already obtained by analyzing homework samples. Therefore, those data

were not analyzed; they are mentioned here only to acknowledge their role in promoting the fidelity of implementation by the treatment instructors.

Fidelity of Implementation – Student Course Evaluations. Students from treatment and control sections were given an opportunity at the end of the semester to rate their own instructor's emphasis on each of the eight Framework objectives. These ratings took the form of a supplemental course evaluation form completed at the same time the traditional course evaluations were given. The treatment students' supplemental form appears in Appendix H. The control students' form differed slightly in that it did not contain the last two questions, which asked about specific elements of the treatment condition that the control students had not experienced.

Student Information. Pre-existing mathematical ability was assessed by examining the ACT and SAT math scores. These scores are routinely made available to instructors for the purposes of determining which students may not have the pre-requisites necessary for succeeding the course. Although a separate pre-test for mathematical ability was scheduled for Week 2, this was not implemented due to personal matters beyond the researcher's control. It was later determined through consultation with the doctoral committee chair that max(ACT, SAT) scores would suffice as a measure of pre-requisite knowledge.

Student attendance rates were also assessed. During Week 12, each instructor was given a class roster and asked to rate how frequently each student attended class.

Instructors relied on their memory and their class records to estimate each student's attendance rate using the following rubric:

3 = always attends - only one or two absences all semester

2 = usually attends - absent once every couple weeks

1 = often absent - absent about once a week

0 = usually absent - absent twice a week or more

Reliability of these instructor ratings of student attendance was not assessed. This potential limitation is discussed further on page 60.

Interview Data. At the end of the study period, the three treatment instructors met with the researcher in a private conference room on campus for a debriefing session. The researcher developed an interview script, available in Appendix I, which he used to guide the semi-structured interview. The interview was audio taped, and the researcher's advisor was asked to attend in order to serve as an objective party who had not been present during the weekly meetings. All three instructors were present in the room at once, so they had an opportunity to share their comments, add to others' responses, and clarify their own positions. The audiotapes were transcribed soon after the interview was completed, and the transcripts and other interview notes were used to help interpret the results and to support the conclusions derived from other data sources.

Timeline

The study period was divided into four sections: the period before the first exam, the period between the first and second exams, the period between the second and third

exams, and the period after the third exam. Altogether, the study period spanned the full Fall 2005 semester of college algebra. The week-by-week timeline was as follows.

Study Period One: Prior to First Exam

- Week 0 (8/24 – 8/26). The first meeting between the course supervisor for college algebra and the instructors took place the week before classes began. The researcher attended the meeting, selected the three treatment instructors, and scheduled the weekly meeting time. The first meeting with the treatment instructors took place before classes began. At that time, instructors were given a copy of the Framework as well as the first set of instructor notes.
- Week 1 (8/29 – 9/2). The instructional treatment began on the first day of class. Treatment group instructors distributed the modified homework list and emphasized the importance of working the Writing in Mathematics questions.
- Week 2 (9/5 – 9/9). There was no class on Monday, 9/5, in observance of the Labor Day holiday. The first set of classroom observations of all six instructors took place on Wednesday. The treatment students' completed quizzes for this week were photocopied and archived by the researcher. A pre-test for mathematical ability was scheduled for this week, but due to personal matters beyond the control of the researcher, the assessment was not given. Weekly meetings with treatment instructors continued this week.
- Week 3 (9/12 – 9/16). Each treatment group instructor began collecting homework from Section P.6. The deadline for homework collection was the night of the first exam. As homework was collected, the researcher photocopied it and

returned the original copies to the instructors. The instructors marked them and returned them to the students. Weekly meetings with treatment instructors continued.

- Week 4 (9/19 – 9/23). The first common hour exam took place on Tuesday, 9/20, from 6 – 7pm. Each instructor submitted his or her exam scores to the researcher after all grades were tallied. Weekly meetings with treatment instructors continued.

Study Period Two: Between the First and Second Exams

- Week 5 (9/26 – 9/30). Weekly meetings with treatment instructors continued.
- Week 6 (10/3 – 10/7). The second round of classroom observations all six instructors took place on Wednesday. The treatment students' completed quizzes for the week were photocopied and archived by the researcher. Weekly meetings with treatment instructors continued.
- Week 7 (10/10 – 10/14). Each treatment group instructor began collecting homework from Section 2.3. As before, the last day to turn in the homework was the night of the next common hour exam. As homework came in, the researcher photocopied it and returned the originals to the instructors, who marked them and returned them to the students. Weekly meetings with treatment instructors continued.
- Week 8 (10/17 – 10/21). The second common hour exam took place on Tuesday, 10/18, from 6 – 7pm. Each instructor submitted his or her exam scores to the

researcher after grades were tallied. Weekly meetings with treatment instructors continued.

Study Period Three: Between the Second and Third Exams

- Week 9 (10/24 – 10/28). Weekly meetings with treatment instructors continued.
- Week 10 (10/31 – 11/4). The third round of classroom observations took place for all six instructors. The treatment students' completed quizzes for the week were photocopied and archived by the researcher. Weekly meetings with treatment instructors continued.
- Week 11 (11/7 – 11/11). Each treatment group instructor collected homework from Section 3.2. As before, the final date homework could be turned in was at the time of the next common hour exam. As homework came in, the researcher photocopied it and returned the originals to the instructors, who marked them and returned them to the students. No instructor meeting was held this week because the researcher was traveling for a job interview. Despite the researcher's absence, a typical set of instructor notes was provided to the treatment instructors.
- Week 12 (11/14 – 11/18). The third common hour exam took place on Tuesday, 11/15, from 6 – 7pm. Each instructor submitted his or her exam scores to the researcher after grades were tallied. Instructors were also given a list of all students on their original class rosters and asked to estimate their attendance rate for the entire semester. Finally, all instructors were asked to announce the three upcoming journal tasks that would be spread over the remaining three weeks of the semester. Weekly meetings with treatment instructors continued.

Study Period Four: After the Third Exam

- Week 13 (11/21 – 11/25). This was a short week due to the Thanksgiving holiday. The first of three journal tasks was given on Monday of this week. Weekly meetings with treatment instructors continued.
- Week 14 (11/28 – 12/2). The second of three journal tasks was given on Wednesday of this week. The final weekly meeting with treatment instructors, spanning the last two weeks of the semester, was held on Monday.
- Week 15 (12/5 – 12/9). The third of three journal tasks was given on Monday of this week.
- Finals Week (12/12 – 12/16). The final exam was administered on Wednesday, 12/14, from 2 – 3:50pm. Each instructor submitted his or her exam scores to the researcher after grades were tallied. The debriefing interview with the three treatment instructors was held on Monday of this week.

Data Analysis

The data analysis for this study consisted of a number of different phases. The quantitative analyses provide the most substantial portion of the results, and *SPSS 13.0 for Windows* was used for most calculations. First, equivalence of treatment and control groups was established by examining ACT and SAT scores. Second, fidelity of implementation was examined using the results from the classroom observations and triangulated with information gathered from the homework samples and the course evaluation supplements. Third, the effect of treatment condition on procedural skill was

explored using the common hour exam scores. Fourth, the effect of the treatment condition on procedural understanding was tested using the six journal tasks given during the last three weeks of the semester. Lastly, the relationship between procedural understanding and procedural skill was explored by correlating students' exam and journal task scores. Qualitative analyses of the interview transcripts were also conducted. The methods used for each of these components are described separately below.

Effect Size Calculations

Effect size is defined as "the strength of the relationship between the independent variable and the dependent variable" (Gliner & Morgan, 2000, p. 177). Cohen's *d* is a common measure of effect size; it is calculated by dividing the difference in group means by the pooled standard deviation. While Cohen's *d* is the most commonly discussed effect size measure (Gliner & Morgan, 2000), it is not available as output from *SPSS 13.0 for Windows*. However, if the correlation coefficient *r* between the levels of the independent variable and the values of the dependent variable is known, the equation

$$d = (2r) / (\text{sqrt}(1 - r^2))$$

can be used to calculate Cohen's *d* (see Becker, n.d.; Thompson, n.d.). Cohen, as cited in Gliner and Morgan (2000), provided the following rules of thumb for interpreting the practical significance of *d*: $d = 0.2$ (small effect), $d = 0.5$ (medium effect), $d = 0.8$ (large effect). Cohen's *d* will be used to report the size of the effect throughout the remainder of this paper.

Analysis of ACT / SAT Scores

Equivalence of control and treatment sections was established by converting students' ACT and SAT math scores into standardized *z*-scores. These were computed using the means and standard deviations obtained from the publishers' annual reports (ACT, 2005; College Board, 2005). For ACT, the mean math score was taken to be 20.6 with standard deviation 5.1. For SAT, the mean math score was taken to be 520 with standard deviation 115. Approximately 85% of the students who took the final exam had either SAT or ACT math scores available. If a student had both scores available, the maximum ACT / SAT math *z*-score was used to be consistent with the college algebra prerequisite policy described above (p. 34). Initial equivalence of the treatment and control groups was then established using independent sample *t*-tests.

Fidelity of Implementation

Fidelity of implementation was assessed primarily through classroom observations, but also through examination of homework samples and student evaluations of their instructor. Analysis of classroom observation data was primarily descriptive in nature, as small sample sizes precluded the possibility of rigorous statistical analyses.

Reliability of the classroom observations was established through a pilot observation episode in which all observers participated. Pearson's correlation coefficients between all pair-wise combinations of observers were calculated from independent observations along the eight dimensions of the Framework for Procedural Understanding (NCTM, 2001). In addition, a frequency table was created to explore how frequently

observers agreed, or disagreed, on their assessments of the pilot observation. For more detail on these analyses, please refer to the results in Chapter 4 (p. 65).

Homework tasks were analyzed by comparing raw completion rates as well as homework quality for the three different treatment sections. Analysis involved comparing relative frequencies of each homework task code (0 through 4, see p. 41) for each of the sampled Writing tasks across the three different instructors and across the three different sampling periods (Weeks 3, 7, and 11). Specific attention was given to searching for differences between instructors, including temporal differences. Such differences could help explain different performances in procedural skill and understanding between the different treatment sections.

The researcher assessed the reliability of the homework task scoring rubric by recoding the homework tasks of nine treatment students who completed at least one homework task. These nine students were those selected as part of the 28-student sample used to assessing reliability of the journal task scoring rubric (see p. 55). The tasks were recoded approximately five weeks after the initial batch was coded. Cohen's *kappa* was then calculated to record the extent to which the observed agreement rate surpassed the agreement rate that would be expected by chance (Landis & Koch, 1977). The results are presented in Chapter 4 (p. 75).

Analysis of student evaluations was limited to comparing line graphs for pairs of matched instructors. The evaluations were anonymous (labeled only by section number), and reliability of the evaluations could not be established. It is not clear how effectively students can rate their own instructor's effectiveness on the eight dimensions of

procedural understanding that are encompassed by the Framework, so these data are unfortunately of little practical significance (for additional information, see p. 75).

Effect of Treatment on Procedural Skill

The three common hour exams and the final exam score were averaged together to create an overall assessment of each student's procedural skill. Analysis of covariance was then used to determine whether instructional treatment explained a significant portion of the variability in exam scores after accounting for attendance rates and max(ACT, SAT) scores. Because max(ACT, SAT) scores were only available for approximately 85% of the students, it was also necessary to examine the remaining 15% of students using attendance rate as the only covariate. Individual analyses were also conducted for each exam using attendance rate and max(ACT, SAT) scores as covariates. The results are presented in Chapter 4 (p. 79).

Effect of Treatment on Procedural Understanding

The six journal tasks given during the last three weeks of school were used to assess students' procedural understanding (Appendix C). Analysis of these results consisted of several phases. First, approximately six weeks after the data collection was completed, the researcher developed a task-specific scoring rubric for each task (Appendix D). The rubrics were developed after the researcher had first scanned through the completed tasks to get a sense of the level of knowledge and types of errors that were present. Before the tasks could be coded, identifying information had to be removed. To that end, the researcher generated a random five-digit identification code for every

student who participated in the study and wrote the appropriate code at the bottom of every journal task. He then photocopied the original student responses and removed all identifying information (except the random code) from each photocopied response. Finally, the researcher printed out six lists of the identification codes, arranged in ascending order, for use in recording individual assessment task scores.

The tasks were then shuffled and scored in batches, so that all responses to Task 1 were scored at once, followed by responses to Task 2, and so on. Due to difficulty in reading photocopies of student work that was written in pencil, the researcher often rewrote some responses in his own hand just above the student's original writing. This was done to record the researcher's interpretation and to ensure consistent coding in the future. In rare cases, the researcher was unable to decipher small portions of the photocopied remarks. If it was felt that the missing portions might influence the proper coding of the response, the task was set aside for later cross-referencing with the original document. Otherwise, a code was assigned using only the available information.

It should be noted that by the time the procedural understanding assessments were conducted, the treatment students had had 12 weeks of practice putting their mathematical ideas into words through the Writing in Mathematics homework questions. Therefore, to avoid a sort of communication bias in favor of treatment students, the theoretical perspective for interpreting journal task responses was intended to reward the diversity of ideas present in each response, not the specific words used to convey them. Consequently, coders were forgiving of the poor, incorrect, or incomplete use of mathematical language that was often present in the responses as long as the underlying

ideas appeared to reflect one or more items from the task-specific scoring rubrics. The coding notes in Appendix E, which were used to train the independent reviewer who helped establish reliability, more thoroughly discuss the theoretical perspective from which the researcher worked while coding the journal tasks.

Stability and inter-rater reliability of the coding system were assessed approximately four weeks after the original coding. The scores originally assigned to the assessments were tested for both inter- and intra-rater reliability by randomly selecting 28 students (20% of the total sample) whose responses would be independently recoded by the researcher and by a trained peer. Training for the independent reviewer consisted of a pilot coding of the six tasks from each of three students, who were purposely selected from the remaining students because they represented a variety of scores. After the independent reviewer coded the pilot responses, the researcher compared the results with the originally coded responses and discussed any discrepancies with the reviewer. The researcher then proceeded to recode the 28-student sample before creating a page of coding notes, available as Appendix E, which explained the broad theoretical perspective the researcher used when coding the responses. The independent reviewer read these notes and proceeded to code the 28-student sample.

Agreement between reviewers was calculating for each journal task using both Pearson's *r* and Cohen's *kappa*. Pearson's *r* was used to capture the association between the reviewers' ratings, and Cohen's *kappa* was used to record the extent to which the observed agreement rate surpassed the agreement rate that would be expected by chance (Landis & Koch, 1977). Interpretation of measures of agreement such as Cohen's *kappa*

depend on the assumption that the raters "must understand and respond to the same question under the same instructions" (Bloch & Kraemer, 1989, p. 270). Pearson's correlation coefficient *r*, on the other hand, records the degree to which high scores by one reviewer are linearly associated with high scores by the other, and vice-versa, and makes no assumptions about the scale being interpreted the same way by all parties. It is possible, therefore, to have no agreement and yet have a high association (e.g., if one reviewer's code was always one unit higher than that of the second reviewer). Low agreement in the presence of high association might indicate bias, scaling differences, or choice of criteria; these "are not forgiven in the context of agreement, and are forgiven in the context of association" (Bloch & Kraemer, 1989, p. 271). Therefore, both statistics were examined to assess reliability. The results are presented in Chapter 4 (p. 77).

Because the six tasks were given as three separate pairs during three different weeks, not all students completed all six tasks. Attendance rates for the three testing dates were 79% in Week 13, 74% in Week 14, and 73% in Week 15. As a result, approximately 8% of 140 total students completed no journal tasks, 11% completed one pair, 28% completed two pairs, and 54% completed all three pairs of tasks. In order to include as many students as possible, the researcher also chose to analyze results using the average score calculated from all available tasks and to weight them according to the percentage of journal tasks completed. Therefore, a student who completed all six tasks was given a weight of 1.0, while a student who completed four or two tasks was given a weight of 0.7 or 0.3, respectively. Weighting of data points for *t*-tests was accomplished using the "Weight Cases…" command in *SPSS 13.0 for Windows*. Weighting of data

points for analysis of covariance was accomplished using the "WLS Weight" option in the Univariate General Linear Model command dialog box.

Independent samples *t*-tests were conducted for the weighted condition for the 92% of students with at least two tasks completed and for the unweighted condition for the 54% of respondents who completed all six tasks. Next, an analysis of covariance was conducted under both weighted and unweighted conditions to create a general linear model for average journal task score using attendance rate and max(ACT, SAT) as covariates and using treatment condition and instructor experience as fixed factors. The results for the weighted and unweighted conditions were similar.

Relationship Between Understanding and Skill

The relationship between understanding and skill was examined by comparing the students' scores on the journal tasks (administered during Weeks 13, 14, and 15) with their subsequent scores on the final exam (Week 16), and also by comparing the students' average scores on the first three exams (administered during Weeks 4, 8, and 12) with their subsequent scores on the journal tasks. Because the average journal task scores and average exam scores were quantitative variables (not categorical), analysis of covariance was not appropriate. Therefore, multiple regression using weighted least squares was used to examine the relationship between the exam scores and the journal task scores for the treatment and control groups. Weighting of data points was accomplished using the percentage of journal tasks completed by each student, reflecting the increased confidence afforded to scores derived from fuller participation in the six-part assessment.

Instructor Debriefing Session

The 36-minute interview was transcribed from the two independent audio recordings shortly after the interview was conducted (for the semi-structured interview script, see Appendix I). Later, the results were manually transferred into a spreadsheet program for coding and analysis. Each row in the spreadsheet represented a related snippet of text (typically, one or two sentences), and there was one column for each interview participant. Codes were then assigned to snippets of text based on emergent themes that arose from repeated readings of the spreadsheet transcript (see Glesne, 1999), and not all snippets were assigned a code. The set of codes (and sub codes) that emerged, as well as the frequency of coded responses in each category, is presented as Table 16 in Chapter 4 (p. 90).

Assumptions and Limitations

Assumptions

Writing in Mathematics Problems Evoke Reflection. The Writing in Mathematics problems in the Blitzer (2004) text were assumed to evoke students' reflection on meaningful mathematical topics which, in turn, helped students develop understanding of the procedures they were learning.

The Framework Appropriately Characterizes Procedural Understanding. The Framework for Procedural Understanding was used to guide the development of the six journal tasks that were used to assess procedural understanding. It was thus implicitly

assumed that these Framework-oriented questions would represent valid measures of understanding.

The Common Hour Exams Measure Procedural Skill. The common hour exams were used in this study to assess students' procedural skill. It was necessary, therefore, to assume that the course supervisor wrote the exams in a manner consistent with assessing procedural skill.

Independence Between Groups. It was assumed that the instructional treatment was restricted to students in the treatment sections, so that control students did not routinely attend treatment class meetings nor obtain lecture notes and homework sets from treatment students. This assumption may not be completely tenable given the nature of multi-section college courses. However, it was further assumed that any interaction between sections would benefit control students and therefore decrease, not inappropriately increase, the power of the study to detect differences between treatment and control sections.

Limitations

Quasi-experimental Design. The study employed a quasi-experimental design using a purposeful sample to select experienced instructors. Through the selection of instructors, intact groups of students were simultaneously selected. This imposed some limitations on the external validity of the study. However, steps were taken to ensure the equivalence of control and treatment groups. Specifically, instructors were matched on

teaching experience before treatment conditions were assigned. In addition, equivalence of prior mathematical ability was examined and determined to be satisfactory. Finally, variability due to any differences in prior math ability and attendance rates was accounted for prior to drawing conclusions from any observed group differences.

Reliability of Attendance Rate Covariate. Attendance rate was used as a covariate throughout the data analysis. Unfortunately, the stability of these instructor-ratings of students' attendance rates was not established (see p. 44). It is not clear how reliably the instructors were able to estimate each students' attendance rate for the semester. However, two pieces of information tend to mitigate concerns about this possible limitation. First, the average attendance rate calculated for the 8am section was lower than all other sections', and multiple independent observations by the researcher and by his advisor support the claim that attendance rates were especially poor for this section. Second, the attendance rate covariate was often a highly significant predictor in the analyses of covariance; this suggests that the attendance rate covariate measured something useful about the students, even if it was not attendance rates per se.

Treatment Students Had Practice Writing. Because the treatment students were assigned numerous Writing in Mathematics homework questions in the homework portion of the treatment, they had 12 weeks of practice putting their ideas into words before the procedural understanding assessments were conducted. This may have lead to a sort of communication bias in favor of treatment students on the journal tasks used to assess procedural understanding. To minimize this effect, the journal task responses were

coded with the intention of forgiving improper use of mathematical language as long as the underlying ideas apparently reflected one or more of the items listed on the task-specific scoring rubrics (see p. 54). Nonetheless, the possibility of a practice effect on the writing tasks should be acknowledge as a possible limitation to the study.

Possible Hawthorne Effect. A Hawthorne Effect refers to the psychological phenomenon by which participants are either more or less motivated simply because they know they are participating in a study (Gliner & Morgan, 2000). In the present study, the treatment group students received a modified homework list on which some drill exercises were replaced with "Writing in Mathematics" exercises that were to be completed before other problems are attempted. Because the original, unmodified list of recommended exercises was distributed among all other sections, the treatment students were told that several instructors had asked the course supervisor for permission to distribute a modified list because they felt the Writing questions were important. Therefore, a Hawthorne effect is possible if treatment group students realized they were participating in a research study based on their classroom performance. However, other than differing homework tasks, all aspects of the treatment were either transparent (e.g., the use of deeper examples during lecture) or equivalent across all sections (e.g., classroom observations and journal tasks given to all six sections). It was therefore deemed unlikely that the results were substantially influenced by students' knowledge of their participation in a study.

CHAPTER 4

RESULTS

Introduction

This chapter presents the results of the data analysis, and it is divided into four sections. The first section examines the covariates used in the study. The second section contains descriptive analyses establishing the degree to which the treatment was implemented. The third section focuses on quantitative analyses designed to explore the relationships between the instructional treatment, procedural skill, and procedural understanding. The last section consists of qualitative analysis of the interview transcripts and describes the treatment instructors' perceptions of implementation issues surrounding the instructional treatment.

Examination of Covariates

Pre-requisite Knowledge – Max(ACT, SAT)

Pre-requisite mathematics knowledge was available in the form of either ACT or SAT for 84% of control students and 86% of treatment students who took the final exam. An independent samples *t*-test was performed to examine the hypothesis that the mean maximum standardized ACT / SAT score was the same for treatment students as for control students. Across all sections, the mean standardized scores were 0.021 for controls and 0.183 for treatment. Cohen's *d* was calculated as a measure of effect size (see discussion on p. 50), and this was calculated to be $d = .301$, indicating a relatively

small effect. Levene's test for equality of variances was not significant (p = .266), so equal variances were assumed for the subsequent independent samples t-test. Using an alpha level of .05, the difference was not significant (t_{117} = 1.625, p = .11), so the treatment and control groups were not found to differ significantly in their pre-requisite mathematics knowledge. Results for individual matched pairs revealed that only the students in the Highly Experienced pair of classes came close to having a statistically significant difference in pre-requisite mathematics knowledge between treatment and control groups. See Table 3.

Table 3 – max(ACT, SAT) for Treatment and Control Groups

	Instructor Experience		
	High	Moderate	Less
Control	-0.02 (n = 26)	0.10 (n = 16)	0.01 (n = 23)
Treatment	0.28 (n = 18)	0.11 (n = 22)	0.18 (n = 14)
t-test	t_{42} = 1.859	t_{36} = 0.098	t_{35} = 0.774
p-value	p = .070	p = .923	p = .444

Note: an alpha level of .05 was used for these tests.

Attendance Rates

During Week 12, all instructors were given a class roster and asked to estimate each students' overall attendance rate for the semester. Overall, the treatment classes had lower average attendance rates than controls (t_{138} = -2.83, p = .005). As an estimate of effect size, Cohen's d was calculated to be d = .481, indicating a medium effect (see p. 50). The results for individual matched pairs, displayed in Table 4, reveal that the Highly Experienced treatment instructor, who taught at 8am, had significantly lower attendance rates than her matched control instructor. No other pair-wise differences were significant.

Table 4 – Attendance Rates for Treatment and Control Groups

	Instructor Experience		
	High	Moderate	Less
Control	2.61 (n = 33)	1.71 (n = 17)	2.33 (n = 27)
Treatment	1.57 (n = 23)	2.04 (n = 26)	2.14 (n = 14)
t-test	t_{54} = -4.863	t_{41} = 1.163	t_{39} = -0.712
p-value	p < .001*	p = .252	p = .481

* Result was significant using alpha = .05.

Drop Rates

The drop rates for the six participating sections, displayed in Table 5, revealed no systematic bias due to differences between treatment and control groups. The average drop rate across all groups was 31% (so that 69% of students were retained). On average, treatment sections retained 72% of their students, while control sections retained 66% of their students. It was noted that the drop rate for the Moderately experienced control group greatly exceeded that of other sections. Only students who completed the final exam were included in the data analyses conducted for this study.

Table 5 – Drop Rates for All Sections

	Highly Experienced		Moderately Experienced		Less Experienced		Overall	
	Treat.	Cont.	Treat.	Cont.	Treat.	Cont.	Treat.	Cont.
Started Course	30	37	36	39	21	40	87	116
Took Final Exam	23 (77%)	33 (89%)	26 (72%)	17 (44%)	14 (67%)	27 (68%)	63 (72%)	77 (66%)

Fidelity of Implementation

This section characterizes the differences in implementation among various sections of college algebra. Fidelity of implementation results are divided into three groups: classroom observations, students' course evaluations, and homework samples. Each group is discussed separately below.

Classroom Observations

Classroom observations were conducted for each of the six participating sections on Wednesday of Weeks 2, 6, and 10. The researcher chose not to examine the observation results in aggregate form because it would obscure the underlying patterns. Thus, the results of each of the three observations will be discussed separately in order to examine the relationship between the intended and observed lesson content. First, however, inter-rater reliability is discussed.

Reliability of the Observations. Inter-rater reliability was established after the initial training period through a small pilot study (as described in Chapter 3, p. 39). Pair-wise correlations were computed to compare the five observers' independent ratings of a single 50-minute class period. The codes assigned by the five observers along the eight Framework dimensions are shown in Table 6. Note that the researcher's codes are included in the last row. Raw agreement rates were not perfect, and while the Pearson correlations between pairs of observers (see Table 7) appear satisfactory, they are unduly

strengthened by the high leverage associated with item 2a. When item 2a was removed from the analysis, the correlations were substantially weakened.

Table 6 – Pilot Observation Codes Assigned by Each Observer

	Framework Objective							
	1a	1b	2a	2b	3	4	5	6
Observer 1	2	1	6	1	0	2	1	0
Observer 2	3	0	6	0	2	3	1	0
Observer 3	3	1	6	0	2	3	1	1
Observer 4	2	2	6	1	1	2	1	1
Researcher	2	1	6	1	0	1	0	2
Mean	2.4	1.0	6.0	0.6	1.0	2.2	0.8	0.8
Std. Err.	0.24	0.32	0.00	0.24	0.45	0.37	0.20	0.37

Table 7 – Pearson Correlations for Pilot Observation Results

	Observer 2	Observer 3	Observer 4	Researcher
Observer 1	0.87	0.88	0.97	0.88
Observer 2		0.98	0.85	0.73
Observer 3			0.90	0.80
Observer 4				0.92

Table 8 shows the pair-wise associations between observers' codes, computed by counting the number of times each pair of codes appeared together in a column of Table 6. For instance, a 0 was paired with another 0 three times in Table 6, and a 1 was paired with a 0 a total of 19 times. These frequencies make up the first two cells in

Table 8. With perfect association between ratings, all entries would occur along the main diagonal (Norušis, 2005).

For the pilot observation results, reviewers gave codes that matched in 41% of cases, codes that differed by one in 48% of cases, and codes that differed by two in 11% of cases. While this was not considered ideal, it was noted that variability in the

subsequent observations would be reduced because observers would be working in pairs and their independent codes would be averaged together. Therefore, it was concluded that the associations were high enough to proceed with data collection.

Table 8 – Frequency of Pair-wise Associations for Pilot Observation Results

		Code Assigned						
		0	1	2	3	4	5	6
Code Assigned	0	3	19	7	.	.	.	.
	1		13	9	2	.	.	.
	2			5	10	.	.	.
	3				2	.	.	.
	4					.	.	.
	5						.	.
	6							10

<u>Week 2 Instructor Notes.</u> The first set of observations was conducted on Wednesday of Week 2. The lesson covered Section P.5 (Factoring Polynomials) of the Blitzer (2004) text. The instructor notes provided by the researcher on this day included the following:

1. A sampling of future situations where students would use factoring, including solving quadratic equations, simplifying rational expressions, subtracting rational expressions, and calculating limits.
2. A series of factoring problems asking which method(s) of factoring would be appropriate for each problem.
3. A list of expressions with a question asking whether each was completely factored or not.

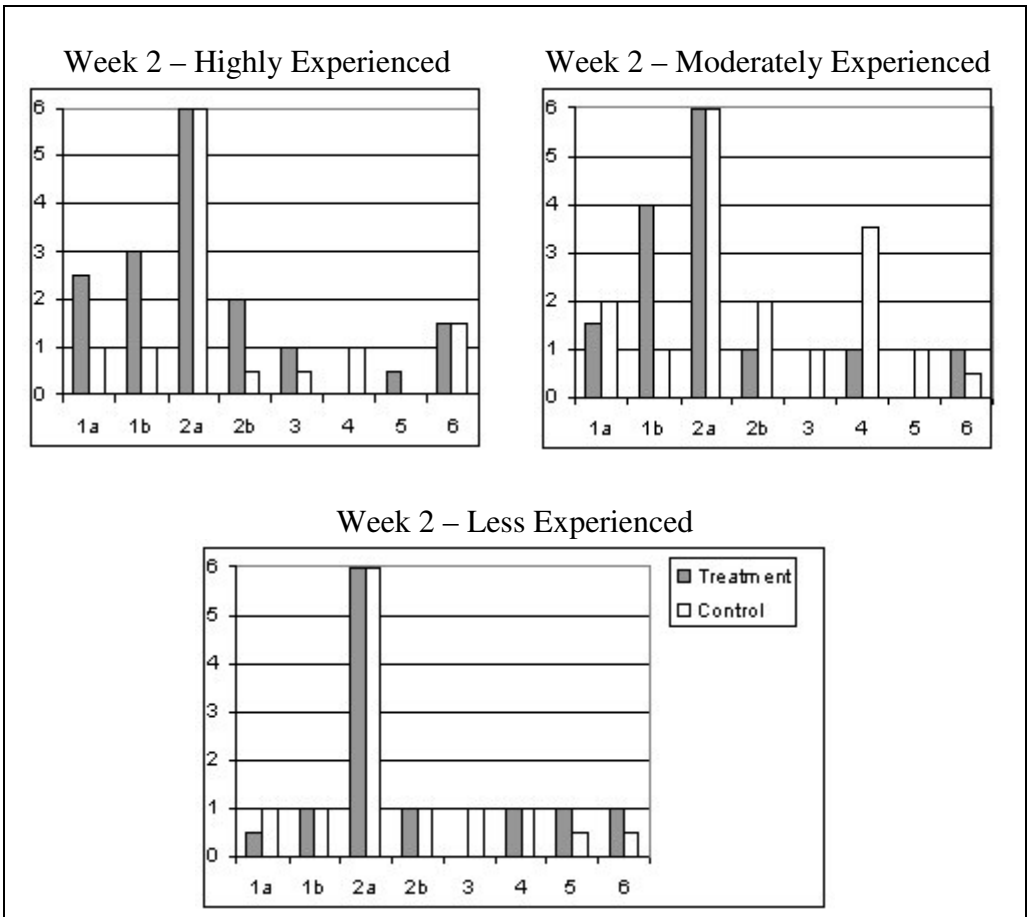
Week 2 Observation Results. The results from the Week 2 observations for each pair of instructors are shown graphically in Figure 2. Note that Objective 2a (Performing the procedure) dominated the lesson, but the Highly Experienced treatment instructor also incorporated comparatively greater levels of Objectives 1a (Overall goal), 1b (Predicting the answer), and 2b (Alternate methods) than the matched control instructor. These results reflect a successful implementation of the last two examples provided in the instructor notes.

In addition to thoroughly covering Objective 2a (Performing the procedure), the Moderately Experienced treatment instructor also incorporated comparatively greater levels of Objective 1b (Predicting the answer) than the matched control instructor. These results reflect a fairly successful implementation of the last two examples provided in the instructor notes. It is worth noting that the control instructor incorporated greater levels of Objective 4 (Verifying the answer) than any other instructor for the Week 2 observation period.

Neither of the two Less Experienced instructors incorporated significant levels of any of the Framework objectives except for Objective 2a (Performing the procedure). This may reflect a relative lack of confidence in the presence of classroom observers.

Week 6 Instructor Notes. The second set of observations was conducted on Wednesday of Week 6. The lesson covered Section 2.1 (Lines and Slope) of the Blitzer (2004) text.

Figure 2 – Results of Classroom Observations for Week 2



Note: The vertical axes represent the average score assigned by the pair of observers who observed each particular lesson. The scale ranges from 0 (Absent), 2 (Infrequent), 4 (Frequent), to 6 (Pervasive).

The instructor notes provided by the researcher on this day included the following:

1. Two examples discussing how linear equations could be used to model certain real-world phenomena, particularly in statistics.
2. Several examples for graphing horizontal and vertical lines that asked students to explain their reasoning. Some of the examples represented non-standard equations

of horizontal and vertical lines, such as $3(y - 2) = 6$, in order to stimulate discussion about how to predict the answer before working the problem.

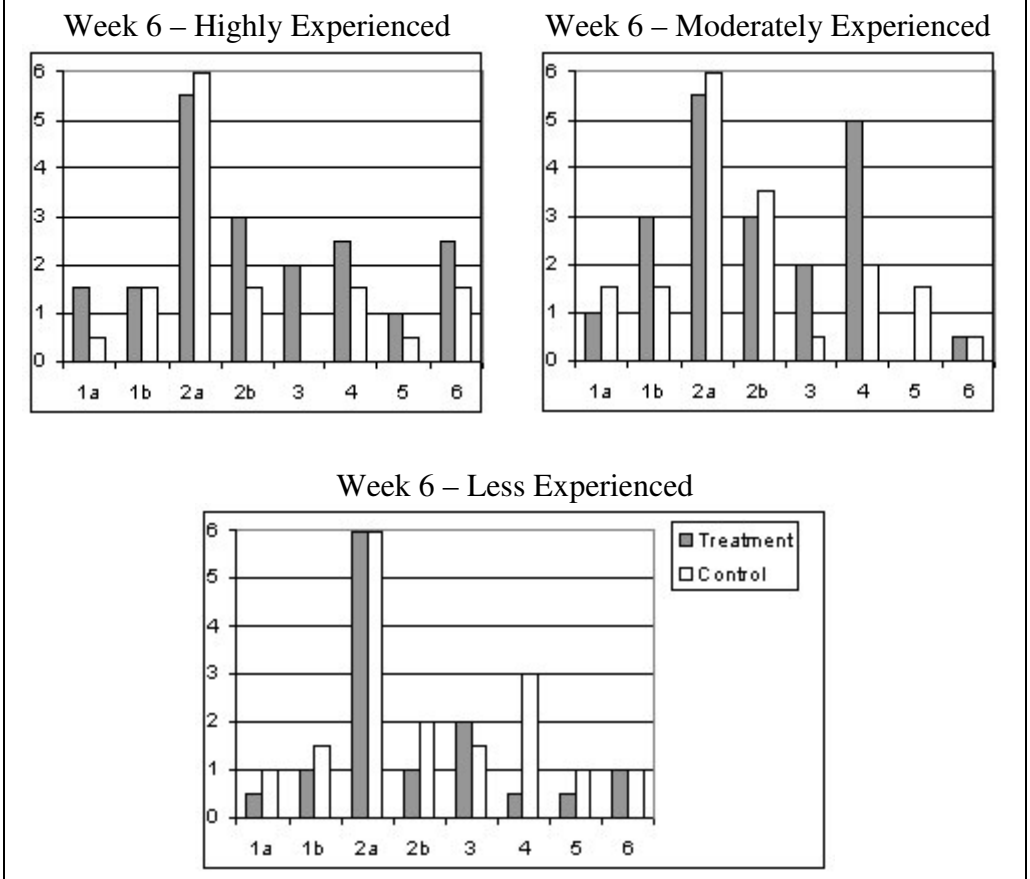
3. A general strategy (i.e., Find the slope, then find a point, and then use the point-slope formula) for finding the equation of a line that works for a variety of contexts (specifically, given two points, given a point and a slope, given a point and a parallel line, or given a point and a perpendicular line).

Week 6 Observation Results. The results from the Week 6 observations for each pair of instructors are shown graphically in Figure 3. Once again, Objective 2a (Performing the procedure) was consistently high, yet the Highly Experienced treatment instructor also incorporated comparatively greater levels of Objectives 1a (Overall goal), 2b (Alternate methods), 3 (Why is this effective?), 4 (Verifying the answer), and 6 (What can we use this to do?) than the matched control instructor. These results reflect a rich implementation of the examples provided in the instructor notes.

In addition to adequately covering Objective 2a (Performing the procedure), the Moderately Experienced treatment instructor incorporated comparatively greater levels of Objectives 1b (Predicting the answer), 3 (Why is this effective?), and 4 (Verifying the answer) than the matched control instructor. These results also reflect a successful implementation of the examples provided in the instructor notes.

The Less Experienced treatment instructor once again had difficulty successfully implementing the examples provided in the instructor notes. Again, this may reflect a relative lack of confidence in the presence of classroom observers.

Figure 3 – Results of Classroom Observations for Week 6



Note: The vertical axes represent the average score assigned by the pair of observers who observed each particular lesson. The scale ranges from 0 (Absent), 2 (Infrequent), 4 (Frequent), to 6 (Pervasive).

Week 10 Instructor Notes. The final set of observations was conducted on Wednesday of Week 10. The lesson covered Section 2.7 (Inverse Functions) of the Blitzer (2004) text. The instructor notes provided by the researcher on this day included the following:

1. Several functions and a question asking whether an inverse function should be expected based on visual cues.

2. Methods for verifying the answer, including the use of graphing technology.
3. Examples prompting students to consider the underlying relationship between the vertical line test (for testing whether a relation is a function) and the horizontal line test (for testing whether a function has an inverse that is a function).

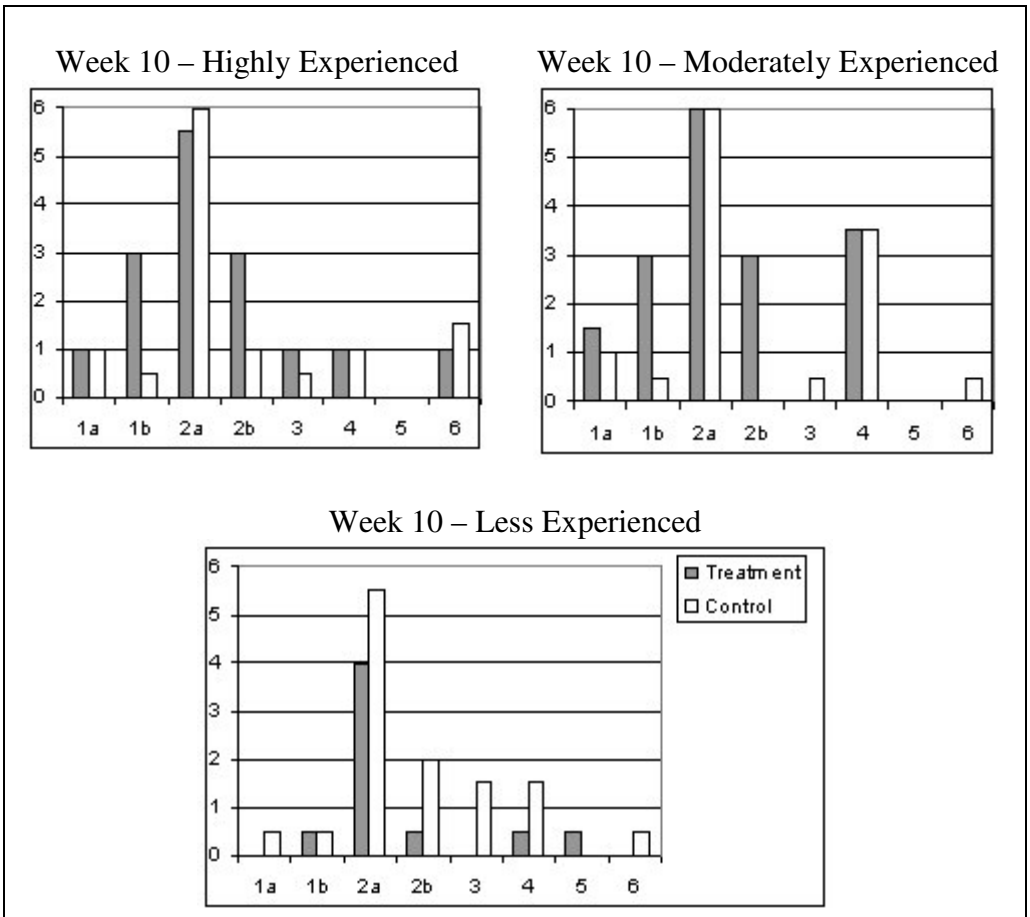
Week 10 Observation Results. The results from the Week 10 observations for each pair of instructors are shown graphically in Figure 4. As before, the Highly Experienced treatment instructor successfully covered Objective 2a (Performing the procedure) in addition to incorporating comparatively greater levels of Objectives 1b (Predicting the answer) and 2b (Alternate representations) than the matched control instructor. These results reflect a successful implementation of the examples provided in the instructor notes.

In addition to adequately covering Objective 2a (Performing the procedure), the Moderately Experienced treatment instructor also incorporated comparatively greater levels of Objectives 1b (Predicting the answer) and 2b (Alternate representations) than the matched control instructor. Objective 4 (Verifying the answer) was also quite high for both instructors. These results reflect a rich implementation of the examples provided in the instructor notes.

The Less Experienced treatment instructor once again had difficulty successfully implementing the examples provided in the instructor notes. The instructor notes requested that each treatment instructor incorporate the graphing calculator overhead display into the lesson. The additional pedagogical demands this requirement imposed, combined with the presence of classroom observers and the instructor's limited

classroom teaching experience, may explain the overall lack of focus on Framework-oriented tasks during this observation.

Figure 4 – Results of Classroom Observations for Week 10

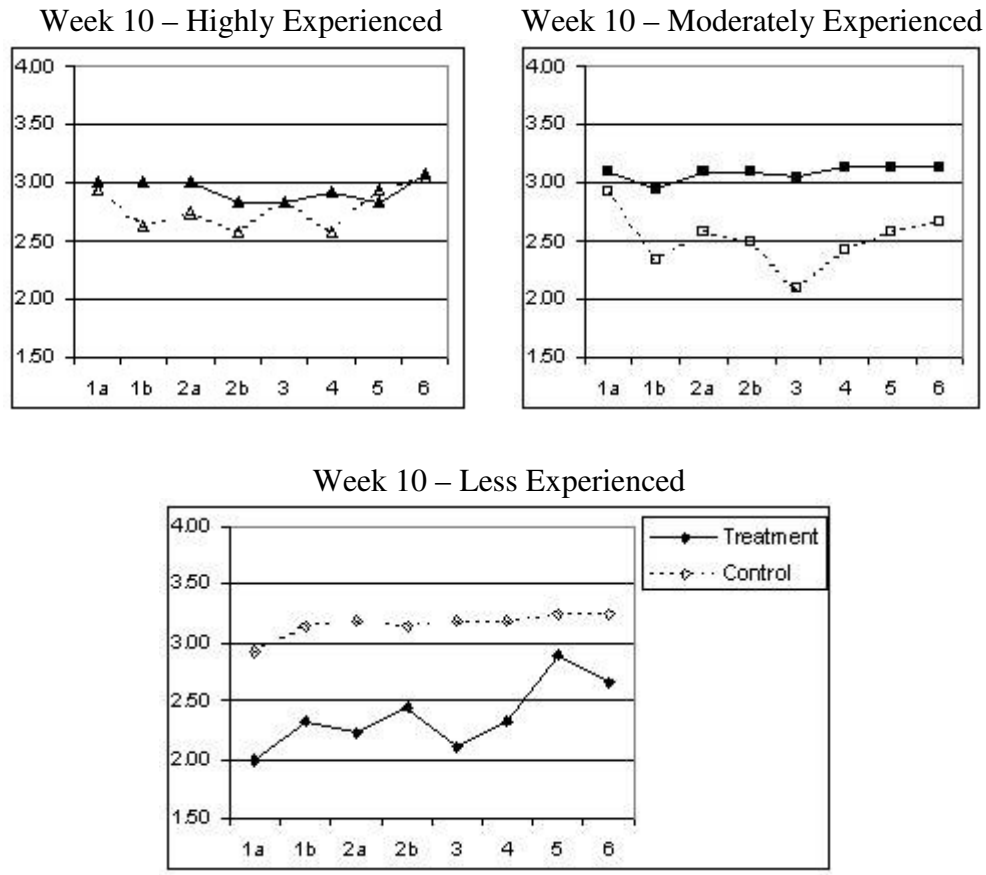


Note: The vertical axes represent the average score assigned by the pair of observers who observed each particular lesson. The scale ranges from 0 (Absent), 2 (Infrequent), 4 (Frequent), to 6 (Pervasive).

Course Evaluations

The results of the supplemental course evaluations are summarized shown in Figure 5.

Figure 5 – Results of Supplemental Course Evaluations



Note: The vertical scale represents the average rating students gave in response to the supplemental course evaluation prompts. Scores have been reverse coded and range from 0 (much worse) to 4 (much better).

Students responded to the prompt: "Compared with previous algebra courses you have taken, how effective would you say this course has been in developing…" their knowledge of each of the eight Framework objectives. The observation form used for treatment instructors is available in Appendix H.

Observe that the Highly Experienced treatment instructor's students rated their section's effectiveness slightly higher on the implementation of the eight Framework objectives than the matched control students rated their section. Also, the Moderately

Experienced treatment instructor's students rated their section's effectiveness substantially higher than the matched control students rated their section. However, the Less Experienced instructor's students rated their section's effectiveness substantially lower than her matched control students rated their section. Comparing the results with the classroom observations revealed no striking associations except that, in both cases, the Less Experienced Treatment instructor fared poorly.

Unfortunately, it was not clear how well students would be able to estimate their instructor's semester-long emphasis on each of the eight Framework objectives. In fact, the results of one question, given only to students in the three treatment sections, suggests that the results may in fact reflect students' overall satisfaction with their instructor instead of instructional emphasis on understanding. That question asked treatment students to "rank the effectiveness of the Writing in Mathematics homework tasks in terms of deepening [their] understanding of Math105 topics." The average rankings on this question were 3.17, 2.52, and 1.33, respectively, for the Highly, Moderately, and Less Experienced instructors. While it is possible that each instructor emphasized the Writing questions differently, it would be surprising to find differences that large given that the classes received identical homework sets. It was concluded that the student evaluations did not provide a reliable measure of the instructors' relative emphasis on procedural understanding along the eight Framework dimensions.

Homework Writing Tasks

Homework samples were collected from all three treatment sections during each of Weeks 3, 7, and 11. These tasks were coded using the rubric described in Chapter 3 (p.

41). Results from the analyses of the Writing problems will be discussed following a brief discussion of reliability. Note that the all analyses were restricted to those students who completed the final exam.

To assess the intra-rater reliability of the scoring rubric, nine students' homework tasks were recoded by the researcher as described in Chapter 3 (p. 52). Analysis of the 113 homework tasks (out of 135 total) completed by the nine students revealed an agreement rate of *kappa* = .767. According to the benchmarks described by Landis and Koch (1977) and summarized in this paper as part of Table 11, *kappa* = .767 represents a substantial rate of agreement. Therefore, it was concluded that the intra-rater reliability of the homework task scoring rubric was acceptable.

Analysis of the Writing problems revealed that completion rates declined as the semester progressed. In addition, Writing task completion rates were generally lowest for the Highly Experienced Instructor and highest for the Less Experienced instructor, particularly later in the semester. See Table 9.

Table 9 – Writing Task Completion Rates over Time

	Week 3	Week 7	Week 11
Highly Experienced	87%	74%	52%
Moderately Experienced	81%	77%	73%
Less Experienced	86%	86%	79%

The other notable result from analysis of the Writing problems is that verboseness, when averaged across all three samples, was highest for the Less Experienced instructor. These results are summarized in Table 10. See Chapter 3 for an explanation of the scoring rubric used to code the Writing tasks. No clear trends in

verboseness were observed over time, suggesting that while completion rates declined over time, the students who continued to complete the Writing problems maintained their level of effort.

Table 10 – Writing Task Verboseness by Instructor

	Verbose (3 or 4)	Terse (1 or 2)	None (0)
Highly Experienced	34%	51%	15%
Moderately Experienced	44%	45%	11%
Less Experienced	56%	32%	11%

Relationships between Treatment, Skill, and Understanding

This section addresses the primary research questions for the study. The reliability of the measures of skill and understanding will be discussed, and then the results pertaining to the relationships between the instructional treatment, students' procedural skill, and students' procedural understanding will be presented in turn.

Reliability and Validity Issues

The course supervisor wrote the exams with the intention of measuring procedural skill. Therefore, because the exams serve as the primary basis for assigning students' grades, they were assumed to be both reliable and valid measures of procedural skill.

The assessments for procedural understanding were considered valid because they were based on the Framework for Procedural Understanding, which was used to define the meaning of the construct called *procedural understanding*. The methodology used to assess inter-rater and intra-rater reliability have been described in Chapter 3 (p. 55). The recoded responses were compared with the original responses in Table 11, which also

includes the benchmarks for interpreting Cohen's *kappa* presented in Landis and Koch (1977).

It was determined that the stability of the researcher's codes over time (intra-rater reliability) was acceptable for all tasks, with the possible exception of Task 3 (which had a relatively low value of *kappa*). It was also observed that, for Tasks 3 and 5 only, the values of *kappa* and *r* between the researcher and the independent reviewer were higher than the corresponding values for the researcher's stability over time. Because differences over time were greater than differences between coders, this suggests some change in the researcher's interpretation of the scoring rubrics for these two tasks, which was then shared with the independent reviewer during the training which occurred at the time of recoding.

Table 11 – Reliability of Journal Task Instrument

Journal Task Number		Researcher vs. Original Codes		Researcher vs. Independent Reviewer	
	n	*r*	*Kappa*	*r*	*kappa*
1	23	.881	.687	.792	.455
2	23	.806	.562	.592	.324
3	21	.557	.360	.699	.645
4	21	1.00	1.00	.933	.857
5	22	.818	.600	.872	.732
6	22	.832	.716	.512	.263

Benchmarks for interpreting *kappa* (Landis & Koch, 1977):
< 0.0 (Poor); .00-.20 (Slight); .21-.40 (Fair); .41-.60 (Moderate);
.61-.80 (Substantial); .81-1.00 (Almost Perfect)

Although agreement rates and stability were not ideal for some individual tasks, the statistical analyses that used the journal task results would ultimately be based upon the mean of the available journal task scores for each student. Therefore, pair-wise

Pearson's correlations, weighted by the proportion of journal tasks completed, were calculated for the three sets of coded sample responses (n = 28). Cohen's *kappa* was not used because it is not appropriate for non-categorical data (Bloch & Kraemer, 1989). The results, shown in Table 12, revealed strong correlations in the average journal task scores. It was concluded that despite some instability in the codes assigned to some of the six tasks, the average journal task scores were stable enough to proceed with data analysis.

Table 12 – Associations between Average Journal Scores

	Researcher's Second Coding	Independent Reviewer's Coding
Researcher's Original Coding	r = .868	r = .828
Researcher's Second Coding		r = .805

Effects of Treatment on Procedural Skill

In order to answer the first research question, the differences in procedural skill between treatment and control groups were examined. Procedural skill was measured in terms of the three common hour exams and the final exam. Table 13 shows the mean differences between exam scores, scaled to 100 points, for all three pairs of instructors. It is apparent that treatment students generally scored higher on the exams, particularly later in the semester.

The researcher used analysis of covariance (ANCOVA) to control for attendance rates and max(ACT, SAT) scores while comparing students' exam scores. Table 13 shows the mean differences between exam scores after controlling for max(ACT, SAT) and attendance rate using ANCOVA. Note that the inclusion of the max(ACT, SAT) covariate restricted the data set to the 85% of students for whom that pre-requisite

information was available. Thus, direct comparisons between Table 13 and Table 14 are not appropriate. However, it is apparent from Table 14 that after adjusting for the covariates by fixing them at their mean values in the linear model, the treatment students generally outperformed the control students, particularly near the end of the semester.

Table 13 – Exam Performance Differences by Experience

	Instructor Experience			Average Across All Experienced Levels
	High	Moderate	Less	
Exam 1	-0.9	-4.2	0.3	-0.8
Exam 2	-1.7	3.5	4.3	1.6
Exam 3	3.0	8.1	5.5	4.3
Final	0.4	6.5	3.9	3.6
Across All Exams	0.2	3.5	3.5	2.2

Note: Positive values are in favor of treatment students.

Table 14 – Exam Performance Differences Adjusted for Covariates

	Instructor Experience			Average Across All Experienced Levels
	High	Moderate	Less	
Exam 1[a]	-5.0	-5.5	0.0	-3.5
Exam 2[b]	2.1	0.0	5.9	2.6
Exam 3[c]	7.2	2.7	5.9	5.3
Final[d]	5.1	1.4	4.9	3.8
Across All Exams[d]	2.7	-0.6	4.2	2.1

Note: Positive values are in favor of treatment students.
[a] Covariates evaluated at max(ACT, SAT) = .0864, attendance rate = 2.0932
[b] Covariates evaluated at max(ACT, SAT) = .0999, attendance rate = 2.1111
[c] Covariates evaluated at max(ACT, SAT) = .0848, attendance rate = 2.1379
[d] Covariates evaluated at max(ACT, SAT) = .0945, attendance rate = 2.1008

The ANCOVA models used to produce the adjusted group means for each exam row total in Table 14 included a term for the interaction between treatment condition and

instructor experience. However, the interaction terms were all non-significant (for the last row, for instance, $F(2, 111) = .463$ and $p = .631$), and including the interaction terms in the models only produced negligible increases in explanatory power (for the last row, for instance, r^2 increased from .348 to .354). Therefore, the interaction term between treatment condition and instructor experience level was not entered into subsequent ANCOVA models (Norušis, 2005).

The ANCOVA test for differences in average exam scores revealed highly significant main effects for both the attendance rate covariate ($F(1, 113) = 28.809$, $p < .001$) and the max(ACT, SAT) covariate ($F(1, 113) = 34.161$, $p < .001$), but it revealed no significant main effects for instructor experience ($F(2, 113) = .437$, $p = .647$) or treatment condition ($F(1, 113) = 1.018$, $p = .315$). As an estimate of effect size, Cohen's d was calculated to be to be $d = .183$, indicating a small effect (see p. 50). It was concluded that the instructional treatment had no significant overall effect on procedural skill among the 85% of students who had max(ACT, SAT) scores available.

Students without max(ACT, SAT) scores available ($n = 21$) were analyzed separately using attendance rate as a covariate and treatment condition and instructor experience as fixed factors. No significant main effects were found for any factor, so it was concluded that the instructional treatment also had no significant overall effect on procedural skill for these students.

Individual analyses of covariance were also conducted for each cell in Table 13. In most cases, the results revealed significant main effects for one or both of the covariates but not for the treatment condition. The notable exception was Exam 3, where

a significant main effect of treatment was found across all instructor experience levels, both with ($F(1, 110) = 5.357$, $p = .023$) and without ($F(1, 132) = 10.716$, p = .001) the max(ACT, SAT) covariate.

Overall, treatment students performed slightly better than control students on the common hour exams. However, there was little statistically reliable evidence to suggest that the treatment led to significant differences in procedural skill between treatment and control groups.

Effect of Treatment on Procedural Understanding

Procedural understanding was assessed using a series of six journal tasks, given in three pairs during Weeks 13, 14, and 15 of the study period. Of the 140 students who completed the course, approximately 92% completed at least one pair of journal tasks, while only 54% completed all three pairs.

Journal task responses were scored using the rubric described in Chapter 3. Table 15 shows the relative frequencies of codes assigned for treatment and control groups. Approximately 55% of treatment students' responses were coded at or above Level 2, meaning the response revealed either a *moderate* or a *high* degree of understanding. Only 37% of control students' responses were classified at or above Level 2, suggesting that the treatment was effective at increasing students' understanding of the procedures tested.

For just those students who completed all six tasks, the average of the mean journal scores was 1.138 for the control group and 1.5805 for the treatment group. An independent samples *t*-test revealed that the difference was highly significant ($t_{73} = 3.982$, $p < .001$), with a 95% confidence interval of (0.2198, 0.6647).

Table 15 – Frequencies of Coded Levels of Understanding on Journal Tasks

Coded Level of Understanding	Treatment	Control
3 – High	17.6%	5.7%
2 – Moderate	37.5%	31.2%
1 – Low	28.2%	35.3%
0 – None	16.7%	27.8%

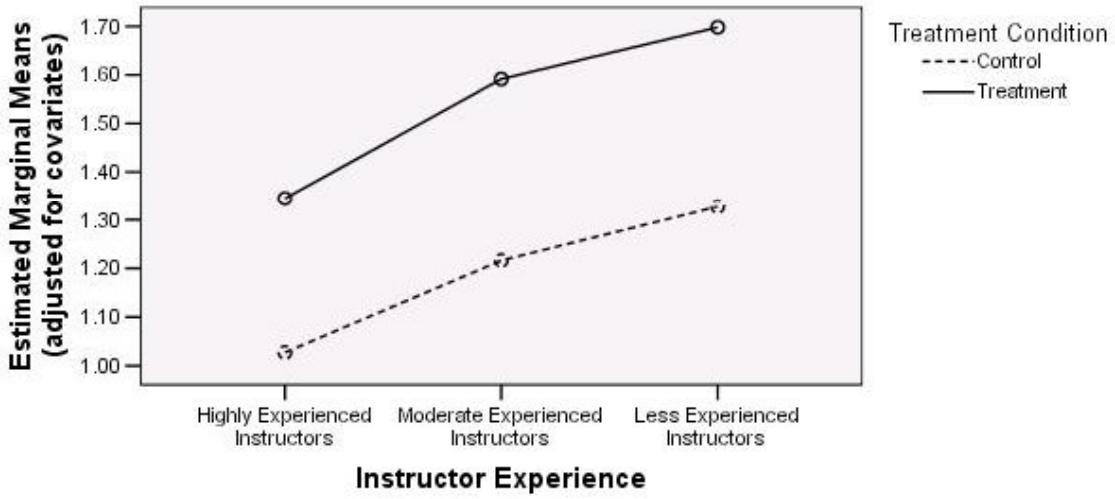
The results were similar when the data were weighted by completion rates for the 92% of students who completed at least two tasks: $t_{105} = 4.317$, $p < .001$, and 95% confidence interval for the difference of (0.2246, 0.6061). As a measure of effect size, Cohen's *d* was calculated to be $d = .928$ for those who completed all six journal tasks, signifying a large effect (see p. 50). When the data were weighted by completion rates, Cohen's *d* was .845, also signifying a large effect. It was concluded that treatment group students scored significantly higher than control group students on the tests for procedural understanding.

Analysis of covariance was used to control for variability due to attendance rates and max(ACT, SAT) scores. For the 54% of students with all six scores, the (unweighted) analysis revealed significant main effects for attendance rate ($F(1, 57) = 5.914$, $p = .018$), instructor experience ($F(2, 57) = 4.952$, $p = .010$), and treatment condition ($F(1, 57) = 10.015$, $p = .002$). The proportion of total variability explained by the model, calculated as r^2, was .363. Additional tests revealed that the main effect for instructor experience was due to a significant difference between the Higher Experienced and Less Experienced instructor groups ($p = .014$).

The results were similar when a weighted least squares analysis was used to compensate for the different precision afforded by average journal task scores calculated from different numbers of tasks. Using the 92% of data with at least two tasks completed,

the analysis revealed significant main effects for the max(ACT, SAT) covariate ($F(1, 103) = 6.114$, $p = .015$), instructor experience ($F(2, 103) = 4.615$, $p = .012$), and treatment condition ($F(1, 103) = 14.589$, $p < .001$), and the proportion of total variability explained by the model, calculated as r^2, was .266. Both ANCOVA analyses supported the results of the *t*-tests, suggesting that the treatment students scored significantly better than control students on the tests for procedural understanding. After adjusting for covariates by fixing them at their mean values in the weighted linear model, the predicted journal task scores for the 92% of students with at least two tasks are graphed in Figure 6.

Figure 6 – Estimated Marginal Means for Procedural Understanding



To interpret the main effect found for instructor experience, additional analyses were conducted to compare the mean differences between each of the three levels. The adjusted means for High, Moderate, and Less Experienced instructors were estimated to be 1.186, 1.404, and 1.513, respectively. Only the difference between the Highly and Less Experienced instructors was found to be statistically significantly ($p = .011$) using the Bonferroni adjustment for multiple comparisons. It was concluded that the students

with Less Experienced instructors scored significantly higher on the tests of procedural understanding than did those with Highly Experienced instructors. A possible explanation for this interesting result is discussed in Chapter 5 (p. 110).

It was further noted that the Less Experienced control instructor had a slightly higher average journal task score than the Highly Experienced treatment instructor. While individual pair-wise comparisons effectively reduce the sample size to $n = 3$ (for the three experience levels), it is interesting that the Less Experienced control instructor would outperform one of the treatment instructors on the journal tasks. One possible explanation for this result might be the relatively low attendance rate associated with the 8am section of the Highly Experienced treatment section (see Table 4, p. 64). Independent external observations of this low attendance rate (mean = 1.57) were noted by the researcher and by his advisor on numerous occasions. By comparison, the Less Experience control instructor taught at 9am and had a mean attendance rate of 2.33. Nonetheless, this observation of a control group matching the performance of a treatment group on the journal tasks does slightly temper the practical significance of the observed gains in procedural understanding.

In summary, the results revealed that on the tests for procedural understanding the treatment students scored higher than control students, and follow-up analyses revealed that the mean journal task scores for the Less Experienced pair of instructors were higher than the mean journal task scores for instructors who had more experience.

Relationship Between Understanding and Skill

Thus far, the results of this study have shown that treatment students performed significantly better on the procedural understanding journal tasks (p. 82), and they performed slightly better on the tests of procedural skill (though the latter difference was not statistically significant; p. 79). Two regression analyses were conducted to explore the relationship between understanding and skill. The first analysis explored whether understanding could predict skill, and the second explored whether skill could predict understanding. Each analysis is described below.

Multiple regression analysis using weighted least squares regression (weighted by percent of journal tasks completed) was carried out to examine how well a linear combination of max(ACT, SAT), attendance rate, average journal score, and treatment condition could predict final exam score. Predictors were entered in steps, and changes in r^2 were observed. When only max(ACT, SAT) and attendance rate were entered, r^2 was .315. Adding average journal score to the model significantly improved the predictive power of the model ($p < .001$), increasing r^2 to .471. However, adding treatment condition to the model did not significantly improve its predictive power ($p = .554$), increasing r^2 by only .002. Therefore, it was concluded that the most parsimonious model for predicting final exam score from available data included max(ACT, SAT), attendance rate, and average journal score.

The fact that treatment condition was not a significant predictor in the model is important. While previous analyses revealed that treatment students' average journal task scores were significantly higher than controls', those who did well on the journal tasks

tended to also do well on the final exam regardless of treatment condition. This suggests that, while the treatment was successful in improving students' procedural understanding, other approaches that similarly advance students' procedural understanding should be expected to produce similar advances in procedural skill.

A similar weighted least squares multiple regression analysis was conducted to determine how well a linear combination of max(ACT, SAT), attendance rate, average score on the first three exams, and treatment condition could predict the average journal score. As before, predictors were entered in steps, and changes in r^2 were observed. When only max(ACT, SAT) and attendance rate were entered, r^2 was .089 ($p = .006$). Adding the average score on the first three exams to the model significantly improved the predictive power of the model ($p < .001$), increasing r^2 to .278. Adding treatment condition to the model also significantly improved its predictive power ($p < .001$), increasing r^2 to .364. The fact that treatment condition was a significant predictor in this model probably reflects the statistically significant differences that were observed between treatment conditions on the journal tasks. It reaffirms that fact that treatment condition explained a significant portion of variability in journal tasks scores beyond that explained by differences in procedural skill.

Together, the two regression analyses revealed that average performance on the first three skills-oriented exams was a significant predictor of performance on the journal task assessments of understanding, and, conversely, the performance on the journal task assessments of understanding was a significant predictor of final exam performance. Students who scored higher on the skills assessments scored higher on the understanding

assessments, and vice-versa. These results are consistent with the results of Rittle-Johnson et al. (2001) that suggest procedural skill and procedural understanding develop in an iterative fashion.

Does Treatment Effectiveness Vary with Prior Math Achievement?

An informal qualitative investigation with high school students has led some to suggest that high achieving students naturally develop Framework-oriented ways of thinking, whereas low achieving students do not (M. Burke, personal communication, March 2006). If this is true, then the Framework-oriented treatment may have differential effects on low and high ability students. While the present study was not explicitly designed to explore this question, it was possible to conduct some relevant analyses.

To that end, the distribution of students' prior math achievement level was estimated by separating the max(ACT, SAT) scores into three groups (low, medium, and high) using standardized z-scores of –0.5 and +0.5 to separate the groups. Analysis of covariance was then conducted using attendance rate as a covariate[4] and treatment condition and prior math achievement level as fixed factors. As before, the model was weighted by journal task completion rates.

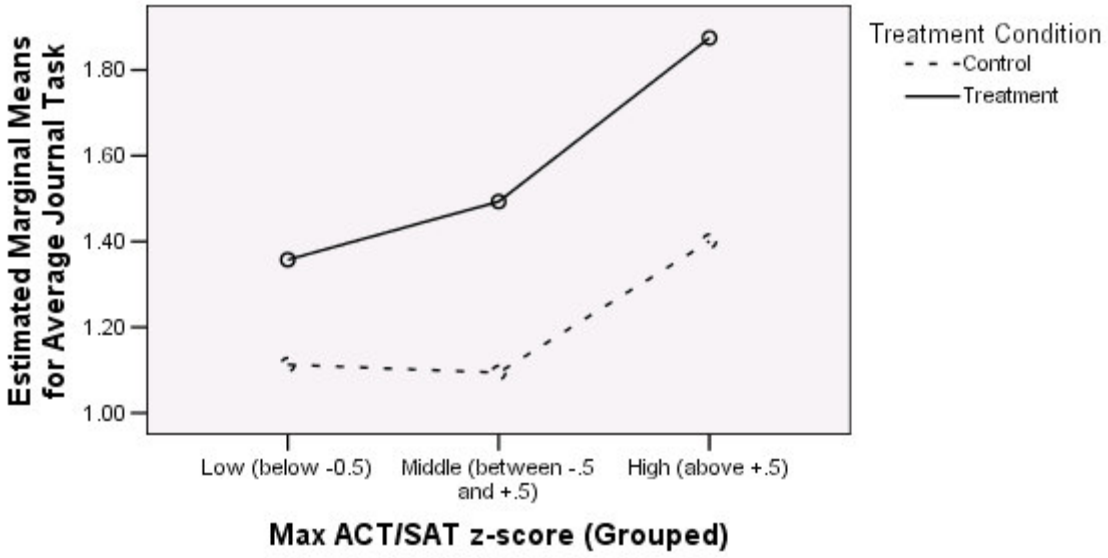
The resulting model had an $r^2 = .237$. The ANCOVA showed a nearly significant main effect for the attendance rate covariate ($F(1, 104) = 3.583$, $p = .061$) and showed significant main effects for both the treatment condition ($F(1, 104) = 11.793$, $p = .001$) and achievement level factors ($F(2, 104) = 5.608$, $p = .005$). Importantly, the interaction

[4] Of course, max(ACT, SAT) was not used as a covariate so as not to violate independence assumptions.

between treatment condition and achievement level was non-significant, ($F(2, 104) = 0.313$, $p = .732$). Follow-up analyses using the Bonferroni adjustment for multiple comparisons revealed that the achievement level main effect was due to significant differences between the highest group of students and each of the other two groups; the difference between the middle and lowest group was not significant.

Because there was no significant interaction between treatment condition and achievement level, it was concluded that there is no reliable statistical evidence in support of claims that Framework-oriented instruction has differential affects on low and high achieving students in college algebra.

Figure 7 – Average Journal Task Scores, Adjusted for Attendance Rate



Instructor Perceptions of the Framework-oriented Treatment

Table 16 shows the set of codes and sub codes, as well as the frequency of coded responses in each category, that emerged as analysis of the interview transcripts

proceeded. Together, the analysis led the researcher to identify three themes: (1) Instructors' comfort with the Framework, (2) Students' reactions to the treatment, and (3) Implementation issues. Each theme is discussed separately below.

Table 16 – Emergent Codes from Interview Analysis

Code	*n*	Code	*n*
"My comfort with Framework"	26	"What I'd do differently"	17
…In the beginning	(10)	Related to Writing questions	17
…How I adjusted	(8)	Student resistance to treatment	8
…In the end	(8)	Time constraints	7
		Weekly training sessions	7
Impact on "C-students"	19	"My growth as a teacher"	6

Instructors' Comfort with the Framework

All three instructors reported that they initially felt some difficulty trying to incorporate the Framework-oriented examples into their daily lessons. After noting her initial "hesitancy," the Highly Experienced instructor said, "I looked at the language that was provided for me, and I tried that language, … and then I thought, 'That felt really awkward.'" The Moderately Experienced instructor supported this, saying, "At first, I think it really changed me a lot [as a teacher]. I was really uncomfortable doing it." The Moderately Experienced and Less Experienced instructors both felt that incorporating the Framework initially took more preparation time than during previous semesters they had taught. The Highly Experienced instructor, who had extensive high school teaching experience but no prior college teaching experience, felt that the Framework would have been much easier to incorporate in a high school setting because the "because the material is so much more jam-packed" in college algebra.

Sometime after the first exam, the instructors agreed, they began to feel more comfortable with the Framework oriented instruction. The Highly Experienced instructor said that what it took for her was to "sit down with myself" and think about how Framework-oriented ideas might naturally arise in the classroom. She added, "maybe it was just me beginning to relate to what you were sharing with us in our Monday morning meetings." The Moderately Experienced instructor added, "kind of like [the Highly Experienced instructor], I had to learn how to just kind of put it in my own words and just to think that way." The Highly Experienced instructor also said that the examples given during the weekly training sessions became "more tangible in the classroom as the semester progressed," and she found the training sessions became more useful to her after she started familiarizing herself with the material beforehand.

During the initial adjustment period, each of the instructors reported trying (unsuccessfully) to incorporate all of the provided examples into their lectures. As they adjusted, however, the instructors reported a change in the way they used the materials. The Moderately Experienced instructor said, "Well at first I guess I would try to do every example and that sometimes wasn't very good," but added, "as I got used to it … I would just write a normal lecture, but then as I was lecturing I would incorporate it that way." The other instructors reinforced her comments by adding similar statements of their own. She then added that once she became more comfortable teaching with the Framework in mind, "I would just go at the pace of the students as well as for me, and try to do some of my own stuff too." She added later, "I was able to let my own teaching style and the new teaching style kind of mix."

The Less Experienced instructor reported that even though this was just her second semester teaching, she frequently asks herself Framework-oriented questions in her own studies. “I always like to teach kind of leaning toward that anyway,” she said, and participating in the study this semester gave her a little extra push to make understanding a priority in her teaching. “But I actually felt it was good for me because I like to teach that way,” she added. The Highly Experienced instructor also remarked that her participation in the study had helped her grow as a teacher. “I think teachers really teach inside their own box,” she said, “that is defined by their own teaching style,” and the treatment forced her to move outside of her own box. That was part of the reason for the discomfort she initially felt, she said, but she felt “proud of being able to go outside of that teaching-style box … and be able to add the component of deeper-level thinking and teaching for understanding.” Later, she added,

> I did feel a good a good deal of success with this type of teaching. And because of the way that kids responded positively sometimes, I would think to myself, “Wow,” you know, “I kind of missed out in high school not really totally emphasizing this.

In summary, it took several weeks before the instructors began to feel comfortable with the Framework-oriented instruction. Sometime after the first exam, however, they began to see ways the Framework could naturally supplement their own lessons and their own teaching styles. One instructor found it useful to write a skeleton lesson before the weekly meeting so that she was already familiar with the material to be covered. By the end of the semester, the instructors said they were able to use the Framework more naturally during class, and they were able to find ways to adapt it to fit the needs of the

students. Two of the instructors expressly stated that they found it beneficial to have participated in the study because it helped them grow as teachers.

How Their Students Reacted to the Treatment

When the instructors were asked to comment on how the students reacted to the Framework-oriented instruction, they all agreed there was an expectation from students of a "procedural mode of instruction" that emphasized procedural skill. The Highly Experienced instructor said she noticed her students "kind of leaned forward in their seats when I'd say, 'Ok, this is how you do it.'" She thought that perhaps a skill-oriented approach was "all they ever knew in the past." The Less Experienced instructor also noticed some resistance to the deeper questions of the treatment, saying, "sometimes they didn't want to know why, … and that was kind of hard to try to explain it when they didn't really want to know." The Moderately Experienced instructor noted that the top students seemed to be the ones who responded well, by which she meant they could answer the questions when prompted during class. She said,

> When I would ask 'what do you expect' or any one of those questions you had for us, [the stronger students] would know the answers. Where the other students would have to actually work through the problems to come up with a response.

During the interview, a theory was posed to the instructors that stronger students might naturally develop deeper knowledge than weaker students, and that Framework-oriented instruction might guide weaker students to develop similarly deep knowledge. The instructors were asked to comment on that theory. The Highly Experienced instructor said, "I do think that some of the kids who were lower, they just weren't quite ready," adding, "they just had whatever coping mechanisms they had dealt with all through their

life." She said some of the "conscientious students" got "a little glimmer in their eyes" when they were answering some of the Framework-oriented questions. When asked which students she was thinking of, she said, "they were the ones that always showed up for class, they'd always turn in homework, they're always there for the quizzes. Just the—I hate to say it—but the dutiful ones." The other two instructors voiced their agreement. "Yeah, the ones that cared," the Moderately Experienced instructor said. Each of the three instructors said they could point to one or two of their weaker students who by the end of the study had asked Framework-oriented questions either during or after class, but otherwise none was able to say that their weaker students had been helped any more or less than their stronger students overall.

In summary, the instructors reported that many students resisted the push toward understanding, and they suggested that these students had probably never been asked to think about Framework-oriented ideas before. They also said it was the conscientious, hard working students that responded best to the treatment, and they did not see any solid evidence to suggest that their weaker students were helped by the Framework-oriented instruction any more or less than their stronger students were.

Implementation Issues

The instructors all cited time constraints as one barrier to implementation imposed by the accelerated multi-section college algebra course. The Less Experienced Instructor said it best:

> Sometimes I felt like, if there wasn't enough time to cover everything, I better just at least show them how to do it, and you know, if they don't know why, then

> they'll have to figure it out on their own kind of thing. Because I had to cover the material. I felt like there was a lot to cover.

If she had it to do over again, she said she "would definitely take more time—to make sure the students really understood." She and the Highly Experienced instructor both thought that it would be beneficial to get the students working in groups more as well. The Less Experienced instructor had incorporated group work into her class during the study period:

> When they did that, I heard a lot of good discussion going on, and they were kind of asking the question like, "Why does it work this way," or, "How can we do that," and I felt like it was really beneficial for them.

In addition, the Highly Experienced instructor thought that, given more time, she would include more application problems that would "bring more relevance into their life, and therefore it would deepen their level of understanding of what's going on."

The instructors were also asked to comment on how much emphasis they gave to the Writing in Mathematics homework questions during class. They said that the students would often ask about the skill-oriented questions in class, but rarely did they ask about the Writing questions. The Moderately Experienced instructor observed that once they realized they weren't being tested on the Writing questions, "a lot of students gave up on the Writing questions in their homework." The Highly Experienced instructor agreed, saying that they asked about the Writing questions at the start of the semester but "not as much towards the end." She said she typically didn't write the answer on the board when they asked, "because I didn't want to have them get the feeling there's a right or wrong answer." Instead, she would just discuss the question with the class.

The course supervisor required all college algebra instructors to collect homework every week. Based on those experiences, the treatment instructors were asked whether or not the majority of students put much time and effort into the Writing questions. “I’d say the majority didn’t,” the Moderately Experienced instructor said. “Just by their answers, a lot of them didn’t make sense.” The Highly Experienced instructor said that about half the students “put some time into answering it, but the others were just, they wrote like a one-liner and just moved on.” The Less Experienced instructor agreed: “Yeah, they just put some general statement. It didn’t really answer the question.”

In summary, the instructors cited time constraints as a significant hindrance to implementing Framework-oriented instruction in college algebra. Also, the students apparently realized that the Writing questions were not going to appear on the exams, and the instructors said completion rates and effort trailed off for some students after that. The instructors all agreed that “the majority” of students did not put much effort into the Writing question, citing the generally low quality of responses as evidence of this observation.

Summary of Important Results

In addition to describing the level of implementation of the treatment and the reliability of the data collection methods, this chapter has presented several important results.

1. Overall, no statistically reliable differences were found between treatment and control groups on tests of procedural skill. The small differences that were

observed generally favored the treatment students, especially on exams given later in the semester. Temporal trends in exam performances, presented in Table 13 (p. 80), revealed that the gap favoring treatment students widened over time. Overall, it was concluded that treatment students did no worse than control students on tests of procedural skill despite an 18% reduction in the number of skill-oriented homework problems that were assigned.

2. A statistically significant difference in favor of treatment students was found on the journal tasks that measured procedural understanding. On average, the treatment students scored nearly a half point higher than control students on the 0 to 3 scale. Interestingly, the Less Experienced instructors tended to score higher on these tasks than those with more experience. Overall, it was concluded that treatment students realized considerable gains in procedural understanding compared to control students.
3. Students' average score on the first three exams was found to be a significant predictor of performance on the journal tasks. In addition, students' average procedural understanding task score was found to be a significant predictor of their final exam score in a linear model that also included attendance rate and max(ACT, SAT) as independent variables. Interestingly, when treatment condition was added to the latter model, it did not explain a significant portion of the remaining variability in final exam score. It was concluded that students who developed greater procedural skill tended to score higher on the journal tasks, and students who developed greater procedural understanding tended to score higher

on the skills-oriented final exam regardless of which treatment condition was initially assigned to them.

4. Exploratory analyses revealed that the interaction effect between treatment condition and max(ACT, SAT) score was insignificant. It was concluded that students with lower max(ACT, SAT) were not helped any more or less by the treatment than students with higher max(ACT, SAT) scores. This result has implications for theories that suggest an instructional emphasis on procedural understanding might help weaker students develop deep knowledge that stronger students naturally develop on their own; it is revisited in Chapter 5 (p. 107).

CHAPTER 5

CONCLUSIONS

Introduction

This chapter offers a concise overview of the study, summarizing the purpose and significance of the study, reiterating the research questions, and reviewing the basic methodology. Thereafter, results from different components of the study are discussed in light of the three research questions. Finally, conclusions and inferences drawn by the researcher are presented and implications for future research are discussed.

Overview of the Study

The purpose of this study was to test an instructional method designed to help students develop a deeper understanding of college algebra procedures. The significance of the study is derived from data on increasing rates of remediation on college campuses (see, for example: Kranz, 2004; Schultz, 2000) and from recent research on the importance of teaching mathematics for understanding. When a unit of mathematical knowledge is understood, that knowledge is more easily remembered (Carpenter & Lehrer, 1999; Hiebert & Carpenter, 1992; Van Hiele, 1986) and it can be applied more readily in a variety of situations (Hiebert & Carpenter, 1992; Kieran, 1992). When knowledge is part of a well-connected network, parts of the network can facilitate recall of other pieces of knowledge that are connected to the network, thereby making it easier to recreate knowledge that may have been forgotten. Finally, when knowledge is

understood it becomes easier to incorporate new knowledge into the existing structure, so that current understanding facilitates future learning (Hiebert & Carpenter, 1992). It is therefore important to develop teaching methods that help students develop understanding of the mathematics procedures they are learning.

This research project examined whether replacing a portion of the skill-oriented college algebra curriculum with examples, homework tasks, and quiz questions derived from a Framework for Procedural Understanding would help students develop deeper understanding, thereby improving their performance on tests of procedural skill. Specifically, the research questions for the study were:

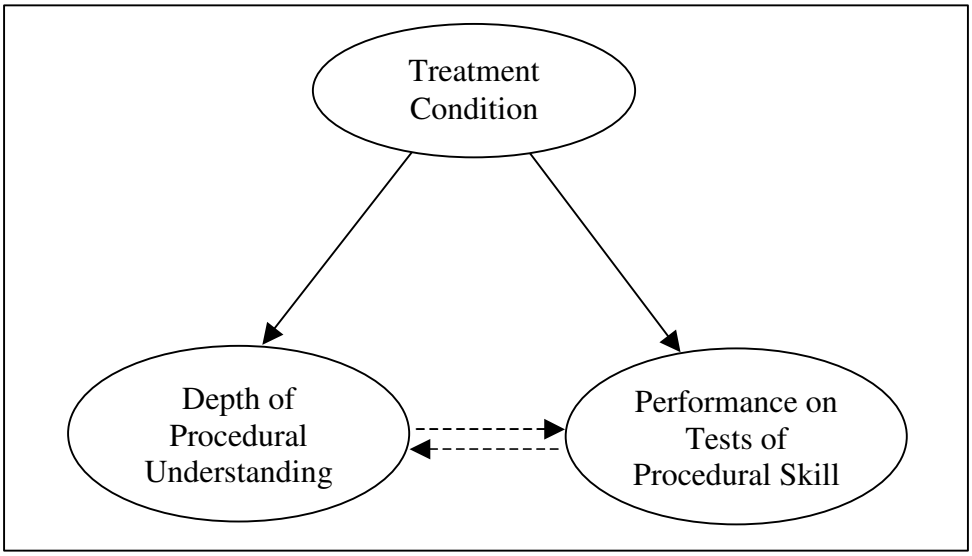
1. Were there significant differences in students' performance on tests of procedural skill in college algebra between treatment and control students?
2. Were there significant differences in the depth of students' procedural knowledge of college algebra between treatment and control students?
3. What were the treatment instructors' perceptions of the overall effectiveness of the instructional treatment?

The first two research questions were addressed by implementing a quasi-experimental research design by which the treatment condition was assigned to three of the six college algebra instructors in Fall 2005 who had prior teaching experience. The six instructors were matched into three pairs based on the amount of prior teaching experience they had, and the treatment condition was assigned to one member of each pair. Four common hour exams (including a cumulative final exam) were used to assess

procedural skill, and six journal tasks were given as three 10-minute assessments during the last three weeks of school to assess procedural understanding.

Figure 8 gives a schematic representation of the relationships of interest to the study. The two solid arrows represent the first two research questions, and the dotted lines represents a secondary relationship that the study was not originally intended to address, but which was a natural extension of the basic research design. The third research question is not represented in the schematic diagram. It was addressed by conducting an end-of-semester interview with the three treatment instructors to gain insight into the results of the first two research questions and to help inform future research.

Figure 8 – Relationships of Interest to the Study



Equivalence of the treatment and control groups was established by comparing the average max(ACT, SAT) scores from among the 85% of students for whom at least one score was available. Classroom observations were used to assess the implementation

of the lecture portion of the treatment, and homework samples were used to assess students' effort and completion rates for the homework portion of the treatment. The results showed that, compared to the more experienced instructors, the Less Experienced treatment instructor incorporated relatively modest levels of Framework-oriented instruction during the classroom observations. This was balanced, however, by her students' comparatively higher completion rates and comparatively greater effort put forth on the Writing in Mathematics homework task samples. Overall, it was concluded that the treatment was implemented successfully.

Analysis of covariance was the primary statistical tool used to address the first two research questions to examine differences in both skill and understanding between treatment groups. Attendance rates and max(ACT, SAT) scores were used as covariates to account for variability due to factors unrelated to the treatment condition. When appropriate, weighted least squares regression was used to assign greater weight to journal task scores derived from students who completed all six tasks.

The third research question was addressed using an end-of-semester interview with treatment instructors. The instructors were not made aware of any of the results from the study prior to or during the interviews. The analysis followed an emergent design in which themes were identified from repeated readings of the interview transcript (see Glesne, 1999). Codes were assigned to snippets of text, and revisions to the coding scheme were made as new themes emerged. The three themes that emerged were discussed in detail in Chapter 4 (p. 89).

Addressing the Research Questions

Research Question 1 – Comparable Procedural Skill

Overall, the results of the study did not reveal statistically reliable differences between treatment and control groups on the tests of procedural skill. Comparisons of group means showed that the treatment students scored slightly higher on the common hour exams, particularly those given later in the semester, but the high variability among exam scores prevented the mean differences from meeting the usual standards for statistical significance (alpha = .05) except on Exam 3. Overall, therefore, the answer to the first research question is that no consistently significant differences in students' performance on tests of procedural skill in college algebra between treatment and control students were found.

It is worth noting, however, that the instructional treatment was designed to produce deeper understanding, not necessarily to improve procedural skill. In fact, treatment students were assigned approximately 18% fewer drill exercises over the course of the semester, and they were assigned 8% fewer problems (drill and writing) overall. That the reduction in drill exercises did not hurt students performance on the skill-oriented exams is an important result, but it is not new. Results from cognitive psychology have clearly demonstrated a sort of "law of diminishing returns" for practicing a new skill, whereby students eventually reach a point beyond which additional practice produces relatively little gain (e.g., see Bower, 2000). Developing understanding takes time and reflection (Carpenter & Lehrer, 1999), and the results of

this study suggest that one way the requisite time might be obtained is to reduce the numbers of drill exercises that students are asked to complete.

Research Question 2 – Differences in Procedural Understanding

On average, the treatment students scored nearly a half point higher than control students on the journal tasks designed to assess students' Framework-oriented procedural understanding. The mean difference of nearly a half point indicated a large and statistically significant difference between the two groups. It was concluded that the combination of Framework-oriented instruction, homework tasks, and quiz questions was effective at helping treatment students developed comparatively greater procedural understanding than control students. Therefore, the answer to the second research question is that there were significant differences in the depth of students' procedural knowledge of college algebra between treatment and control students, and those differences favored the treatment students.

It is important to note that the students in the treatment sections were exposed to Framework-oriented content on a daily basis, and so they should be expected to outperform the control students on the Framework-oriented journal tasks. Strictly interpreted, the positive result on the second research question simply reveals that instruction focused on Framework-oriented content can lead students to develop Framework-oriented knowledge. The observed differences on the procedural understanding assessments also suggest that students do not necessarily develop deep procedural understanding without explicit instruction.

For the purposes of this study, Framework-oriented knowledge has been associated with the term "procedural understanding," so that a student who can answer Framework-oriented questions is said to have developed a significant measure of understanding of that procedure. However, the term understanding has historically been difficult to quantify. In light of recent models mathematical knowledge, it is apparent that knowledge of a procedure may be either shallow or deep (de Jong & Ferguson-Hessler, 1996; Hasenbank, 2005; Star, 2000, 2005) in addition to being either tentative or well-practiced (Hasenbank, 2005). Thus, perhaps the most appropriate inference to be made regarding the second research question is that, after Framework-oriented instruction and practice, the treatment students' responses revealed deeper knowledge of college algebra procedures than was revealed by responses from control students who did not participate in the Framework-oriented instruction.

Research Question 3 – Instructor Perceptions of the Treatment

The follow-up interview used to answer the third research question provided insight into the treatment instructors' experiences with the Framework-oriented instruction. Unlike the first two research questions, the third research question cannot be answered in a single sentence. The treatment instructors' perceptions of the overall effectiveness of the instructional treatment have been divided into three themes: (1) Instructors' comfort with the Framework, (2) Students' reactions to the treatment, and (3) Implementation issues. Each theme will be summarized below.

The instructors reported initially feeling uncomfortable teaching under the guidance of the Framework, but "sometime after the first exam" they were able to allow

their own teaching style to merge with the Framework-oriented lessons in a more natural way. It took time before the language of the Framework felt natural, and they said they had to learn to put the Framework-oriented questions into their own words and into their own examples before they felt comfortable using it. They reported that the Framework-oriented examples presented during the weekly meetings became "more tangible in the classroom" as the semester progressed, and two of the instructors said they thought their participation in the study had helped them grow as teachers (the third did not comment).

According to the instructors, there was some student resistance to the treatment. They felt the students had come to the study expecting a skills-oriented mode of instruction, and some students would "kind of lean forward in their seats" when the instructors finished a Framework-oriented example and began showing how to work a problem. The instructors also observed that some students became reluctant to put in substantial time and effort on the Writing questions once they noticed that understanding would not be assessed on the exams. None had any specific examples of extreme student resistance, however. They simply noted that the most diligent students were the ones who seemed to respond best to the treatment. Nonetheless, it appears that student apathy can be a barrier to successful implementation of Framework-oriented instruction.

The most common implementation issue noted by the instructors was a shortage of time. They felt that if the time constraints imposed by the fast-paced college algebra curriculum were removed, they would have been able to help their students develop greater understanding through increased use of group work and greater use of application problems. The Highly Experienced instructor noted that the treatment would have been

easier to implement in the high school classes she had previously taught because of the additional time that is available in that setting.

Additional Conclusions and Inferences

The primary research questions have been answered, and the results show that treatment students were able to develop deeper understanding of college algebra procedures without reductions in performance on tests of procedural skill, even when 18% fewer drill exercises were assigned. The remainder of this chapter will be devoted to discussing a number of additional implications that have grown out of this study.

Revisiting the Relationship Between Understanding and Skill

The results of the study revealed that students who scored higher on the first three exams tended to score higher on the journal tasks, and those who scored higher on the journal tasks tended to score higher on the subsequent final exam. Of course, these two analyses contain a fair amount of overlap: students who scored well on the first three exams had demonstrated a level of skill that should be reflected on the final exam. Moreover, correlation need not imply causation. However, while this study was not originally designed to assess the relationship between understanding and skill, the links that have been demonstrated are consistent with findings that suggest the two knowledge types develop in an iterative fashion (Rittle-Johnson & Koedinger, 2002; Rittle-Johnson et al., 2001).

The Results Took Time to Emerge

The treatment instructors were able to merge their own teaching style with the language and style induced by the Framework for Procedural Understanding, but it took several weeks before they became comfortable in doing so. Moreover, the evidence from the first exam favored the control instructors, suggesting that perhaps the instructors' discomfort may have influenced their students as well. Alternatively, the students themselves may have needed time to adapt to an instructional approach intended to help them develop habits of mind that would help them to execute procedures more "intelligently" (Star, 2002b). Regardless, the semester long duration of the treatment eventually produced positive results. There is an underlying implication for future research: understanding takes time to develop, and the results may take time to emerge. Therefore, short-term studies seeking to minimize the presence of confounding variables through brief implementations of novel instructional approaches are unlikely to produce meaningful results.

A Description of Baseline College Algebra Instruction

Classroom observations were conducted of all six instructors during each of Weeks 2, 6, and 10 of the study period for the purpose of assessing fidelity of implementation. The observations also provide a rare glimpse into the day-to-day classroom activities that characterize college algebra instruction at state universities like the one where the study was conducted. In addition to the more progressive notions of procedural knowledge that are captured by the Framework (e.g., "The student

understands the overall goal of the algebraic process and knows how to predict or estimate the outcome" (NCTM, 2001, p. 31)), the observation form also contains a dimension for recording classroom episodes where the focus is on how to execute the procedure. This objective is represented as Objective 2a on the classroom observation forms (Appendix G).

The results of the observations displayed in Figure 2, Figure 3, and Figure 4 show that the instructors consistently emphasized Objective 2a to a far greater extent than any of the other objectives. Table 17 summarizes the results for Objective 2a and the other objectives for the treatment and control groups. The fact that the common hour exams were explicitly designed to assess procedural skill helps explain the emphasis on Objective 2a. Nonetheless, Table 17 reveals that, at least on the days observed, the control instructors rarely focused attention on the Framework-oriented questions that have been associated with deep procedural knowledge. The treatment instructors, who were coached to include Framework-oriented content in their lessons, managed to implement a more balanced lesson that was at least partially responsible for treatment students developing deeper procedural knowledge without sacrificing performance on procedural skill.

Table 17 – Classroom observations revealed high emphasis on Objective 2a

	Objective 2a (% coded "6 – Pervasive")	All other objectives (% coded higher than "2 – Infrequent")
Control	94%	9%
Treatment	72%	18%

Deeper Lectures and Deeper Practice

Analysis of the classroom observation data revealed that the Less Experienced instructor had a difficult time incorporating the Framework-oriented content into her examples. In most cases, her matched control instructor had higher marks on the dimensions of the classroom observation form. However, as Table 18 shows, the Less Experienced treatment instructor's students had the highest mean scores on every exam after Exam 1, and they had the highest average journal score as well.

Table 18 – Mean Exam Scores and Journal Task Scores for each Instructor

	Highly Experienced		Moderately Experienced		Less Experienced	
	Control	Treatment	Control	Treatment	Control	Treatment
Exam 1	65.70	64.82	69.44	65.27	61.94	62.25
Exam 2	75.52	73.83	72.25	75.76	71.85	76.14
Exam 3	75.48	78.45	68.03	76.18	74.80	80.29
Final	66.37	66.75	66.06	72.56	66.80	70.73
Mean Exam	70.77	71.00	68.75	71.80	68.85	72.35
Mean Journal	1.05	1.30	1.12	1.67	1.28	1.70

Examination of the max(ACT, SAT) scores in Table 3 and attendance rates in Table 4 does not help explain her students' top performance, because those measures were not exceptional compared to the other groups. However, a clue to the Less Experienced treatment students' success can be found by examining the results of the Writing in Mathematics homework sample analyses. Table 9 shows that after the first exam the completion rates on the Writing questions were highest for the Less Experienced treatment instructor. Moreover, Table 10 shows that the Less Experienced treatment instructor's students were more "verbose" in their responses to the sampled Writing tasks. Therefore, the most likely explanation for the Less Experienced treatment

instructor's students' top scores is that they worked more diligently on the out-of-class portion of the treatment. This suggests that it was not merely the deeper lecture content that helped treatment students perform better. On the contrary, and consistent with the constructivist view of learning, the active homework component appears to have been an integral part of the treatment's success.

Recommendations for Future Research

Revisiting the Interaction Between Treatment and Ability

The numerical results revealed no interaction between prior math ability level and the treatment's effectiveness, suggesting that weaker students were not helped any more or less than stronger students by the treatment. It should be noted that math ability level is a relative term, and students who enroll in college algebra are typically those who were not successful in their previous algebra courses. The 90th percentile among the college algebra students' max(ACT, SAT) z-scores was 0.67. Using the national means for ACT and SAT, it was calculated that $z = 0.67$ corresponds to an ACT math score of 24 or an SAT math score of 590 (ACT, 2005; College Board, 2005). These scores are in approximately the 75th percentile nationally.

Therefore, the students in the sample were not top students, and the fact that gains were made in their procedural understanding is encouraging. Nonetheless, the treatment instructors reported feeling that many of their students were not prepared for the type of thinking required by the Framework, and that they often fell back on their old coping skills (such as rote memorization) to "get by." The question remains: does the treatment

help weaker students develop understanding comparable to that which stronger students develop on their own? Additional research is needed to examine that question, particularly in high school algebra classes and other groups that are more representative of the full range of student ability levels.

Structuring Classroom Observations using Time-line Graphs

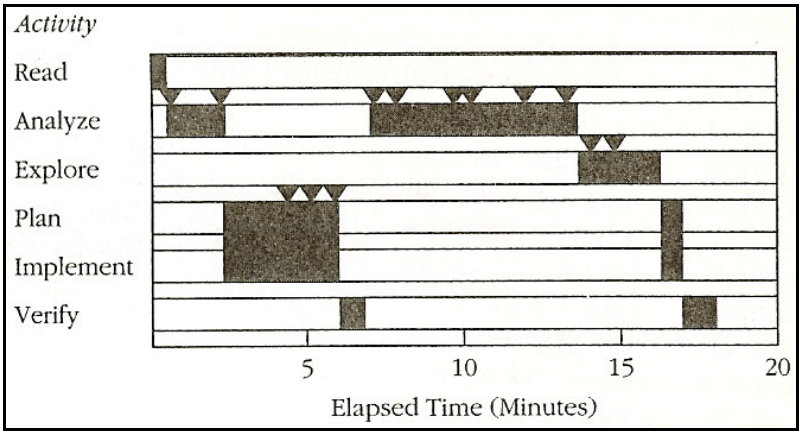
The results obtained from the classroom observations conducted in this study were limited to some extent by the design of the observation instrument. The two-part instrument began with a page where observers kept a tally of "classroom events" that they observed for each of the eight Framework-oriented dimensions. Immediately after the class period was over, the researchers then used the tally marks as a guide to record their impression how frequently each of the eight Framework objectives had been attended to in the lesson. The initial tally sheet was not analyzed further, and only the overall emphasis sheet was used to record fidelity of implementation.

As the observations unfolded, however, it became apparent to the researcher that the relative granularity of the overall emphasis sheet was not ideal. It was noted that while the tally marks allowed the researcher to distinguish between active (group work, seat work, group discussion) and passive (lecture) classroom episodes, they did not provide a means for recording how the instruction unfolded over time. Consequently, the results could not address questions such as: Did certain Framework objectives tend to precede (or follow) worked examples? What was the typical duration of a Framework-oriented "classroom event?" What percentage of time was spent on Objective 2a (Performing the Procedure), as compared to other objectives? These questions, and many

others like them, cannot be answered unless a method is devised that associates the classroom events with some record of the chronological progression of the lesson.

Fortunately, a similar method has been used for many years by Schoenfeld in his research on students' problem solving strategies. Figure 9 shows an example that "traces a mathematics faculty member's attempt to solve a difficult two-part problem" (Schoenfeld, 1992, p. 356). It illustrates how time-line graphs can help delineate the mathematician's progress through the six problem solving dimensions Schoenfeld identified. By replacing Schoenfeld's six problem solving dimensions with the eight dimensions of the Framework for Procedural Understanding, the resulting observation instrument would provide a more accurate picture of the actual classroom implementation of the Framework.

Figure 9 – An Example of Schoenfeld's Time-line Graph Method



Embedding the Framework into the Curriculum

The instructional treatment employed in the present study was used as a supplement to an existing skills-oriented curriculum, and yet it still produced large and statistically significant gains on the tests of procedural understanding without causing

declines on the tests of procedural skill. Moreover, these results were realized despite an 18% reduction in the number of drill exercises assigned. The positive results suggest the Framework has a promising role to play in curriculum development. Future research is needed to examine whether the benefits of the Framework can be expanded by embedding the treatment into the curriculum.

Curriculum developers and textbook authors should note that the researcher found it easy to build the Framework-oriented questions into existing skill-oriented examples. Many of the examples used in the treatment were simply adaptations of the examples provided in the Blitzer (2004) algebra text. Still others were developed by the researcher as he reflected on the eight Framework questions and asked himself which would be the most appropriate for the specific procedures under consideration. The Framework-oriented questions were used to augment skills-oriented questions in the curriculum with the intention of producing a more balanced curriculum emphasizing both understanding and skill. In this way, skill-oriented questions can become vehicles for helping students develop their own procedural understanding. Seamlessly building the Framework into an existing curriculum and examining its effectiveness is a natural next step toward helping more students develop deep procedural knowledge.

Closing Remarks

The National Council of Teachers of Mathematics has called for all students to learn mathematics with understanding (NCTM, 2000), and the results of the present study are especially pertinent to those who have come to realize the important role that

understanding plays in making knowledge robust, flexible, and long lasting. This study has demonstrated that students enrolled in a remedial college algebra course can be led to develop deeper procedural knowledge when instruction and practice are refocused on questions designed to elicit understanding. Moreover, these benefits can be realized without reductions in procedural skill, even when many skill-oriented homework problems are replaced with questions that promote reflection and understanding. As existing models of mathematical knowledge continue to be refined through research, reflection, and practice, new instructional models must also be developed. The results of this study have demonstrated that the Framework for Procedural Understanding (NCTM, 2001) is a promising approach for helping students develop deeper understanding of mathematical procedures.

REFERENCES

ACT. (2005). National distributions of "percent at or below" for ACT test scores ACT-tested high school graduates of 2002, 2003, and 2004. Retrieved Feb. 22, 2006, from http://www.actstudent.org/pdf/norms.pdf

Becker, L. A. (April 12, 2006). Lecture notes: Effect size. Retrieved April 12, 2006, from http://web.uccs.edu/lbecker/Psy590/es.htm

Bereiter, C. (1991). Implications of connectionism for thinking about rules. *Educational Researcher, 20*(3), 10-16.

Bereiter, C. (2000). Keeping the brain in mind. *Australian Journal of Education, 44*(3), 226-238.

Blitzer, R. (2004). *College algebra* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Bloch, D. A., & Kraemer, H. C. (1989). 2 x 2 kappa coefficients: Measures of agreement or association. *Biometrics, 45*(1), 269-287.

Bower, G. H. (2000). A brief history of memory research. In E. Tulving & F. Craik (Eds.), *Oxford handbook of memory* (pp. 3-32). Oxford: Oxford University Press.

Brozo, W. G., & Simpson, M. L. (1999). *Readers, teachers, learners: Expanding literacy across the content areas* (3rd ed.). Upper Saddle River, NJ: Prentice-Hall, Inc.

Carpenter, T. P., & Lehrer, R. (1999). Teaching and learning mathematics with understanding. In E. Fennema & T. A. Romberg (Eds.), *Mathematics classrooms that promote understanding* (pp. 19-32). Mahwah, N.J.: Lawrence Erlbaum Associates.

College Board. (2005). 2005 college-bound seniors total group profile report. Retrieved Feb. 22, 2006, from http://www.collegeboard.com/prod_downloads/about/news_info/cbsenior/yr2005/2005-college-bound-seniors.pdf

de Jong, T., & Ferguson-Hessler, M. G. M. (1996). Types and qualities of knowledge. *Educational Psychologist, 31*(2), 105-113.

Dougherty, B. J. (1996). The write way: A look at journal writing in first-year algebra. *The Mathematics Teacher, 89*(7), 556-560.

Erlwanger, S. H. (1973). Benny's conception of rules and answers in mathematics. *The Journal of Children's Mathematical Behavior, 1*(2), 7-26.

Glesne, C. (1999). *Becoming qualitative researchers - an introduction* (2nd ed.). New York: Addison Wesley Longman, Inc.

Gliner, J. A., & Morgan, G. A. (2000). *Research methods in applied settings: An integrated approach to design and analysis*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Greeno, J. G. (1978). Understanding and procedural knowledge in mathematics instruction. *Educational Psychologist, 12*(3), 262-283.

Hasenbank, J. (2005, October 20-23). *A vision of a three-dimensional re-conceptualization of mathematical knowledge.* Paper presented at the 27th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education, Roanoke, VA.

Henningsen, M., & Stein, M. K. (1997). Mathematical tasks and student cognition: Classroom-based factors that support and inhibit high-level mathematical thinking and reasoning. *Journal for Research in Mathematics Education, 28*(5), 524-549.

Hiebert, J. (1986). *Conceptual and procedural knowledge: The case of mathematics*. Hillsdale, NJ: L. Erlbaum Associates.

Hiebert, J., & Carpenter, T. P. (1992). Learning and teaching with understanding. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 65-97). New York: Macmillan.

Hiebert, J., & LeFevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 1-27). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kaput, J. J. (1995). A research base supporting long term algebra reform? *Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Columbus, OH.

Kieran, C. (1992). The learning and teaching of school algebra. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 390-419). New York: Macmillan.

Kranz, C. (2004). Regents upset over number in remedial college classes. Retrieved May 30, 2005, from http://www.enquirer.com/editions/2004/10/22/loc_regents22.html

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174.

NCTM. (2000). *Principles and standard for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.

NCTM. (2001). *Navigating through algebra in grades 9-12*. Reston, VA: National Council of Teachers of Mathematics.

Norman, D. A. (1993). *Things that make us smart: Defending human attributes in the age of the machine*. Reading, MA: Perseus Books.

Norušis, M. J. (2005). *SPSS® 13.0 guide to data analysis*. Upper Saddle River, NJ: Prentice Hall, Inc.

Rittle-Johnson, B., & Koedinger, K. R. (2002, October 26-29). *Comparing instructional strategies for integrating conceptual and procedural knowledge.* Paper presented at the 24th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education, Athens, GA.

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology, 93*(2), 346-362.

Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 334-370). New York: Macmillan.

Schultz, V. H. (2000). Education presentation to the Los Altos - Mountain View League of Women Voters. Retrieved May 30, 2005, from http://www.hobel.org/lwved/id24.htm

Slavin, R. E. (1997). *Educational psychology: Theory and practice* (5th ed.). Needham Heights, MA: Allyn & Bacon.

Star, J. R. (2000). *On the relationship between knowing and doing in procedural learning.* Paper presented at the Fourth International Conference of the Learning Sciences.

Star, J. R. (2002a). *Re-"conceptualizing" procedural knowledge in mathematics*: U.S.; Michigan.

Star, J. R. (2002b, October 26-29). *Re-conceptualizing procedural knowledge: The emergence of "intelligent" performances among equation solvers.* Paper presented at the Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education, Athens, GA.

Star, J. R. (2005). Reconceptualizing procedural knowledge. *Journal for Research in Mathematics Education, 36*(5), 404-411.

Thompson, B. (April 12, 2006). A suggested revision to the forthcoming 5th edition of the APA publication manual. Retrieved April 12, 2006, from http://www.coe.tamu.edu/~bthompson/apaeffec.htm

Montana State University (2004). *Bulletin: Montana State University - Bozeman* (Vol. LVIII, No. 3). Bozeman, MT.

Van Hiele, P. M. (1986). *Structure and insight: A theory of mathematics education*. London: Academic Press, Inc.

Wiggins, G., & McTighe, J. (2001). *Understanding by design*. Upper Saddle River, NJ: Prentice Hall, Inc.

Wu, H. (1999). Basic skills versus conceptual understanding: A bogus dichotomy in mathematics education. *American Educator, 23*(3), 14-19, 50-52.

APPENDICES

APPENDIX A:

INSTRUCTOR NOTES FOR WEEK 5

Monday of Week 5

1.5 – Quadratic Equations (Writing: 119 – 121, 123 [from Friday])

1. Solve quadratic equations by factoring.
2. Solve quadratic equations by the square root method.
3. Solve quadratic equations by completing the square.
4. Solve quadratic equations using the quadratic formula.
5. Use the discriminant to determine the number and type of solutions.
6. Determine the most efficient method to use when solving a quadratic equation.
7. Solve problems modeled by quadratic equations.

Framework Objectives to Emphasize:

- "Why does this work?" (Framework Objective 3)
- "Which procedure is the most efficient one to use?" (Framework Objective 5)

Course Supervisor Notes:

- Two days on this but you can spend most of day on completing the square and then applications on the second so try to teach some on the 1.4 day.
- Explain when the answer is $\pm$ a and when it is just the positive root.
- We are teaching two of four primary methods of solving algebraic problems. MA160 spends a lot of time on ZPR (zero product rule) and QF (quadratic formula) so please use that terminology. Even completing the square is really ZPR. (Other 2 methods are Inverse-Reverse and Guess and Check).
- Completing the square comes up when we work with circles and vertex of a parabola. We will only do problems with leading coefficient of 1 so we need to change some homework.
- Leave answers in simplified radical form except application problems.
- Skip example 6
- When you write out the quadratic formula, be sure to include "x = " and emphasize it because we will using the QF to find x^2 or (x + 3) or sinx later.
- They can program QF into calculator (on my webpage) so it will be given on exams.
- Homework changes: Elim 111, 115. Move problems to day 2 according to what you teach day 1.

Framework-oriented examples:

By now, your students should have had some practice working problems: they are beginning to develop a sense of which method(s) they prefer. Now is a good time to remind them of the merits of each approach.

Write the following list on the board to help them identify. Better yet, have them generate the italicized portions – deep learning works best when the participants are actively engaged.

- Factoring & Zero Product Rule

  Best when factoring is easy to do.

Needs to be set equal to zero first!
Examples: $x^2 + 5x + 4 = 0$; $3x^2 - 6x = 0$.
Why does this work? Because if a product of two numbers is zero, then at least one of the numbers must be zero.

- Square Root Method
  Best when you have a perfect square on one (or both!) sides
  Examples: $3(x - 1)^2 = 12$; $(2x + 3)^2 - 9 = 0$
  Why does this work? Definition of square-root! If "something-squared" equals a number, then that "something" is either the positive or negative square root of that number.

- Completing the Square
  Always works, but is not always efficient.
  This is easiest when the lead coefficient is equal to one
  Relies on the square root method!
  Example: $x^2 - 6x - 11 = 0 \Rightarrow (x - 3)^2 = 20 \Rightarrow x = 3 \pm \text{sqrt}(20)$
  Why does this work? We manipulate the expression (following the rules of algebra) to make it a perfect square. Then we apply the square root method.

- Quadratic Formula
  Always works, but takes a lot of time to simplify
  This is a good default choice if you have plenty of time
  Especially good if you can't factor and it has a lead coef. that is not 1
  Example: Redo the previous example (on completing the square)
  (They are roughly equal in difficulty!)
  Example: $3x^2 - 12x - 5 = 0$
  Why does this work? The quadratic formula is derived by completing the square of a general quadratic equation. It is in their book on page 125. They don't need to know the derivation, but it would be nice if they knew the quadratic formula is nothing more than a "completing the square" problem done once for all.

Wednesday of Week 5

1.6 – Other types of Equations (Writing: 95 – 98)

1. Solve polynomial equations by factoring.
2. Solve radical equations.
3. Solve equations with rational exponents.
4. Solve equations that are quadratic in form.
5. Solve equations involving absolute value.

Note: There is a lot to do in this section. Don't spend too much time on solving polynomial equations (which are just done by factoring anyways) or you'll run out of time for the harder examples.

Framework Objectives to Emphasize:

- How is this related to other procedures we have learned? (Framework Obj. 5)
- How can we verify our answer?

Course Supervisor Notes (from Tom's email):

- Extraneous solutions from squaring both sides can be detected by checking your work. This is something we can always stress with these students.
- Omit problems like example 4, involving two radicals.
- Point out why we can't cancel x^2 before solving $3x^4 = 27x^2$ (if x = 0, we divided by 0).
- Tell them solving absolute value problems using the method in example 8 is a bad idea. It won't work for $|x+2| < 2$.

Framework-oriented examples:

There are several classes of examples to work from:

1. Polynomial equations – we will solve these using factoring.
   Example: *$4x^4 = 12x^2$ (Remind them not to cancel the x^2! Factor instead… See the Study Tip on page 132.*
   Framework connection: This is much like the process for solving quadratic *equations by factoring… and has the same limitations. This only works if you can find a way to factor the polynomial.)*
   Example: *$x+1 = 9x^3 + 9x^2$ (This requires factoring by grouping, but is very similar to problems 3 and 5 in the homework. Better do one like this…)*

2. Radical Equations – isolate the radical and then square both sides.
   Note that we must check our answers because we could get extraneous solutions using this method.
   How can we verify our answer? 1) Plug them both into the original equation, or 2) set the original equation equal to zero and graph to look for roots.

3. Rational Equations – See p. 137 for a hard-to-remember procedure.
The basic idea is to raise both sides to the *reciprocal power*.
Examples: $3x^{3/4} - 6 = 0$; $x^{2/3} - 3/4 = -1/2$. *(This is Example 5 on p. 137: work these out before class!)*
They have a hard time remembering when to use the ± and when not to. The only way to really keep this straight is to remember that denominators are roots, and numerators are powers.
Ask them questions like, "Are we taking an even root here?" If so, then you have to put a ± in there.
Remember to check all proposed solutions in the original equation (see above).

4. Quadratic in form (Substitutions) – These are a little tricky, but they are important for calculus-bound students.
The idea is to use a substitution to make it look like a quadratic. The difficult parts for students are 1) recognizing the pattern and 2) remembering to substitute back at the end and continue solving.
Examples: Solve $x^4 - 5x^2 + 6 = 0$; Solve $3x^{2/3} - 11x^{1/3} - 4 = 0$.
Both of the examples above are solved by substitution. In the first case, let $t = x^2$. You'll then get two roots for t, and each of those is equal to x^2, which gives still two more roots (a total of four). Check your answers at the end!
In the second case, let $t = x^{1/3}$ and you'll find two solutions for t. You then have to cube both sides in each of them to find out what x equals. Work these out before class if you haven't done problems like this before! (There are homework problems assigned that require this).

5. Absolute Value Equations – Tom is right: don't do these using the method in Example 8. Instead, write: " '|X| = c' is equivalent to 'X = c OR -X = c' ". That method will generalize more readily to the inequalities in the next section.

Friday of Week 5

1.7 – Linear Inequalities

1. Graph an inequality's solution set.
2. Use set-builder and interval notations.
3. Use properties of inequalities to solve inequalities.
4. Solve compound inequalities.
5. Solve inequalities involving absolute values.

Framework Objectives to Emphasize:

- Why does this work? (Framework Objective 3).
- What types of problems can we solve with this technique? (Objective 6).

Course Supervisor Notes (From Tom's Email)

- Be sure to use book notation here. Number lines with "[" and "(" instead of open and closed circles.
- Interval notation may be new to some so extra practice and stress notation as a language (believe me, language is important in 160). We will differ slightly from the book. Use $x \in [3, 9)$ rather than set notation (though students can use either).
- $2 < x < 5$ uses "and", while $x < 2$, $x > 5$ uses "or"
- Compound inequalities must be written from smallest to largest (e.g. $-3 < x < 5$ rather than $5 > x > -3$ because they may then write $-3 > x - 1 > 3$ when they solve $|x-1| > 3$).
- Study tip and blue box on bottom of pg 150 are fie for memorization. Encourage them to write two equations to understand absolute value inequalities. You can tell from the graph if is "and" or "or".

Framework-oriented examples:

There are two classes of problems: "$|X| < c$" and "$|X| > c$". Do plenty of examples that will help them distinguish between these methods. In each case, set it up using the pattern:

$|X| < c \quad \Leftrightarrow \quad X < c \text{ AND } -X < c$

$|X| > c \quad \Leftrightarrow \quad X > c \text{ OR } -X > c$

If you try to use the book's method (cf. p. 150 at the bottom), about half of the students will solve "$|X| > c$" using exactly the same method as they would for "$|X| < c$". I'd love to help them understand the difference, and perhaps you should make this observation in passing (that "$-X > c \Leftrightarrow X < -c$", which matches the book). However, when you do examples, please do it the way I spelled out above.

Incorporating the first theme: There are many opportunities here to help the students visualize what they are doing. As you solve problems, draw a lot of number lines to illustrate what is going on. Remember that that absolute values can be interpreted as distances. For example, $|x - 5| < 3$ can be read, "the distance between x and 5 is less than 3."

	<u>Application Examples:</u> Try to work a few applications into your examples. One of the strengths of inequalities is that they are used to model the real world behavior. Here are some from the un-assigned homework. #96 – The formula $T = 0.01x + 56.7$ models the mean global temperature, T, in degrees Fahrenheit, of Earth x years after 1905. For which range of years was the global mean temperature between 56.7^{o} and 57.2^{o}? #97 – The inequality $\lvert x - 60.2 \rvert \leq 1.6$ describes the actual viewing population of M*A*S*H in the US. Solve the inequality and interpret the solution. Explain why the survey's *margin of error* is ± 1.6%.

Looking Ahead to Week 6

- Week 6, Monday: More 1.7 – Linear Inequalities
- Week 6, Wednesday: 2.1 – Lines and Slope
- Week 6, Friday: 2.2 – Distance and Midpoint Formulas; Circles

Week 5 Quiz Questions:

Choose at least two of the following to put on your quiz over week 5. (You should also create a couple of your own questions to complete the quiz).

Remember, the goal of putting these questions on quizzes is to drive home the point that these types of questions are important. They must be seen as an *emphasis* of your section of college algebra, not an add-on.

1.5 – Identify which method would be the most efficient for solving each of the following quadratic equations. Choose from (i) Factoring & Zero Product Rule, (ii) Square root method, (iii) Completing the Square, (iv) Quadratic Formula.

a) $(3x - 4)^2 = 16$ (correct answer: square root method)
b) $x^2 - 6x + 7 = 0$ (correct answers: either completing the square or the quadratic formula)
c) $9x^2 - 24x = 0$ (correct answer: factoring)
d) $3x^2 - 4x - 5 = 0$ (correct answer: quadratic formula – the lead coefficient is not 1, so completing the square is not wise).

Taken straight from the lecture examples

1.5 – What is the discriminant and what does information does it provide about a quadratic equation?

This is #123 from 1.5

1.6 – When you take the square root of both sides of an equation, it is necessary to insert the ± symbol. How does that apply to problems with rational exponents. Use "$x^{3/4} = 2$" and "$x^{6/5} = 2$" in your explanation.

Taken from lecture.

1.6 – Explain how to recognize an equation that is quadratic in form. Give an example to illustrate.

Adapted from Writing in Mathematics #96

APPENDIX B:

HOMEWORK LISTS

Control Homework List

Date	Monday	Date	Wednesday	Date	Friday
Aug 29	Lecture: Class policy, P1, P2, HW: pg 11: 18-48 every 3rd problem, 60–72 every 3rd, 81–90 every 3rd, pg 22: 3-18 odd, 29, 39-63 every 3rd, 107-109 all	Aug 31	Lecture: P3, HW: pg 34: 1-36 every 3rd, 49-103 every 3rd, 109, 113, 119-123 odd	Sept 2	Lecture: P4, P5 HW:, pg 46: 41-53 odd, 65-81 odd, 85, 87-90, 104
Sept 5	**Labor Day Holiday**	Sept 7	Lecture: P5, HW: pg 57: 3-15 odd, 21-43 odd, 51-78 every 3rd, 81, 87, 89, 99, 100, 108-110	Sept 9	Lecture: P6, HW: pg 68: 9-30 mult of 3, 31, 36-63 every 3rd, (**33 typo in ans.*), 69, 80, 81
Sept 12	Last day to add/drop online Lecture: P6, 1.1, HW: pg 81: 29-32, 33-38, 45-48, 56-59 60-64	Sept 14	Lecture: 1.2, HW: pg 92: 6-48 every 3rd, 61-67, 73, 81, 90, 91, 94, 95	Sept 16	Lecture: 1.3, HW: pg 103: 5-20 mult of 5, 27, 30, 31, 35, 40, 41, 43, 46, 47, 51, 53, 54, 68-75, 82, 84, 86
Sept 19	Last day to add/drop w/o w Review/Help	Sept 21	Lecture: 1.4, HW: pg 113: 37-40, 51	Sept 23	Lecture: 1.5, HW: pg 128: 3-48 every 3rd, 57-96 every 3rd
Sept 26	Lecture: 1.5, HW: pg 128: 101-104, 107, 108, 111, 113, 115, 129, 130	Sept 29	Lecture: 1.6, HW: pg 141: 1-19 odd, 31, 34, 35, 39, 47, 51, 53, 60-72 mult of 3, 75-85 odd, 88, 95-97, 102, 103	Sept 30	Lecture: 1.7, HW: pg 153: 3-45 mult of 3, 53-79odd, 93, 95, 98, 99,
Oct 3	Lecture: 1.7, HW: pg 153: 103, 104, 108, 109, 118, 119, 122, 126, 127,	Oct 5	Lecture: 2.1, HW: pg 188: 5-35 mult of 5, 37, 38, 45-66 every 3rd, 69, 72-75, 78-80, 90, 97, 99	Oct 7	Lecture: 2.2, HW: pg 199: 5-30 mult of 5, 35-47 odd
Oct 10	Lecture: 2.2, HW: pg 199: 49-60, 74-76	Oct 12	Lecture: 2.3, HW: pg 211: 1-19 odd, 21-69 every 3rd	Oct 14	Lecture: 2.3, HW: pg 211: 73, 79, 89, 93, 104-107, 109
Oct 17	Review/Help	Oct 19	Lecture: 2.4, HW: pg 228: 15-81 every 3rd, 91-93, 102, 110, 111, 113, 117, 118	Oct 21	Lecture: 2.5, HW: pg 244: 3-51 every 3rd (*omit 18)
Oct 24	Lecture: 2.5, HW: pg 244: 53-56, 66-74	Oct 26	Lecture: 2.6, HW: pg 257: 3-45 every 3rd, (*33 & 37 wrong ans. in book)	Oct 28	Lecture: 2.6, HW: pg 257: 47-58, 68, 69
Oct 31	Lecture: 2.7, HW: pg 267: 3-24 every 3rd, 27(a only), 30-40	Nov 2	Lecture: P5, 2.7: pg 267: 43-45, 47, 52, 57, 58, 64-66	Nov 4	Lecture: 3.1, HW: pg 288: 1-8, 9-27 every 3rd, 33-43 odd, 47, 49, 50, 59, 63, 66-70
Nov 7	Lecture: 3.2, HW: pg 300: 1-20, 21-45 every 3rd, 51, 54, 80-82	Nov 9	Lecture: 3.2 Handout, HW: worksheet	Nov 11	**Veterans Day Holiday**
Nov 14	Review/Help	Nov 16	Lecture: 3.3, HW: pg 313: 3-11 odd, 12, 27, 29	Nov 18	Last day to drop with a "W" Lecture: 4.1, HW: pg 382: 12, 13, 19-24, 29-43 odd, 44, 46, 47, 49, 55-57, 66, 68-70
Nov 21	Lecture: 4.2, HW: pg 395: 3-42 every 3rd, 43-48, 51-90 every 3rd, 102 (a & b only), 104-107	Nov 23	Lecture: 4.3, HW: pg 405: 3-36 every 3rd, 42-81 every 3rd, 84, (93, 94), 102-104	Nov 25	**Thanksgiving Holiday**
Nov 28	Lecture: 4.4, HW: pg 415: 7-43 odd, 47-52	Nov 30	Lecture: 4.4, HW: , 53-63 odd, 69, 79, 87, 88, 91	Dec 2	Lecture: 4.5, HW: pg 428: 1-9, 14, 17, 19, 21, 24, 25, 26
Dec 5	Lecture: 5.1, HW: pg 452: 1-11 odd, 12	Dec 7	Lecture: 5.1, HW: pg 452: 19-25 odd, 53	Dec 9	Review
Dec 12		Dec 14	Comprehensive Final 2:00 – 3:50 PM Dec 14	Dec 16	

Treatment Homework List

Date	Monday	Date	Wednesday	Date	Friday
Aug 29	Lecture: Class policy, P1, P2, HW: pg 11: [W: 84-87], [18-36 every 3rd, 41, 48, 60, 63, 65, 72, 81, 90], pg 22: [W: 99-103], [7, 9, 11, 15, 29, 39, 45, 49, 55, 58, 60, 63, 107-109 all]	Aug 31	Lecture: P3, HW: pg 34: [W: 111-115, 116], [1, 3, 5, 9, 11, 23, 25, 29, 31, 33, 36, 51, 53, 61, 64, 67, 77, 80, 82, 85, 89, 91, 97, 99, 103, 109, 119-123]	Sept 2	Lecture: P4, P5, HW: pg 46: [W: 96], [41, 43, 47, 49, 53, 65, 67, 71, 73, 77, 79, 81, 85, 87-90, 104]
Sept 5	**Labor Day Holiday**	Sept 7	Lecture: P5, HW: pg 57: [W: 102-104], [3-11 odd, 15, 21, 23, 29, 31, 33, 39, 41, 43, 51, 54, 57, 60, 63, 69, 72, 75, 78, 81, 87, 89, 100, 108-110]	Sept 9	Lecture: P6, HW: pg 68: [W: 71, 73, 74, 75, 76, 80, 81], [9, 12, 15, 24, 31, 39, 45, 48, 54, 57, 60, 63, 69] (FYI: #33 has typo in answer book)
Sept 12	Last day to add/drop online Lecture: P6, 1.1, HW: pg 81: [W: 49, 51, 53], [29-33, 35-37, 45-48, 56, 57, 59, 60-64]	Sept 14	Lecture: 1.2, HW: pg 92: [W: 79, 82], [6, 9, 12, 15, 21, 24, 27, 33, 39, 42, 48, 61, 67, 73, 90, 91, 94, 95]	Sept 16	Lecture: 1.3, HW: pg 103: [W: 79], [5, 10, 15, 20, 27, 30, 31, 35, 40, 41, 43, 46, 47, 51, 53, 54, 68, 69, 71-75, 82, 84, 86]
Sept 19	Last day to add/drop w/o w Review/Help	Sept 21	Lecture: 1.4, HW: pg 113: [W: 45], [37, 39, 40, 51]	Sept 23	Lecture: 1.5, HW: pg 128: [W: 119-121, 123], [3, 12, 18, 21, 27, 33, 39, 45, 60, 63, 69, 75, 78, 81, 87, 93, 96; 101-104, 107-108, 111, 113, 115, 129-130]
Sept 26	Lecture: 1.5, HW: pg 128: [W: none], [101-104, 107, 108, 111, 113, 115, 129, 130]	Sept 29	Lecture: 1.6, HW: pg 141: [W: 95-98], [3, 5, 9, 11, 15, 17, 19, 31, 34-35; 39, 47, 51, 53, 60-75 every 3rd, 77-85 odd, 88, 102-103]	Sept 30	Lecture: 1.7, HW: pg 153: [W: 111, 113-114], [3, 6, 9, 15, 21, 24, 27, 36, 39, 45; 53-63 odd, 69, 71, 75, 77, 79, 93, 95, 98-99]
Oct 3	Lecture: 1.7, HW: pg 153: [W: 118, 119], [103, 104, 108, 109, 122, 126, 127]	Oct 5	Lecture: 2.1, HW: pg 188: [W: 82-83, 85-90 (#89 is mis-num!)], [10, 20, 25, 35, 37-38, 45, 51, 57-72 every 3rd, 73, 78-80, 97, 99]	Oct 7	Lecture: 2.2, HW: pg 199: [W: 61, 62, 66, 67], [5, 15, 20, 30, 35, 37, 41, 45]
Oct 10	Lecture: 2.2, HW: pg 199: [W: none], [49, 51-52, 56-60; 74-76]	Oct 12	Lecture: 2.3, HW: pg 211: [W: 95-97, 101], [1, 5, 9, 13-21 odd, 24-57* every 3rd (*except 33), 60, 66, 69]	Oct 14	Lecture: 2.3, HW: pg 211: [W: none], 73, 79, 89, 93, 104-107, 109
Oct 17	Review/Help	Oct 19	Lecture: 2.4, HW: pg 228: [W: 98-100, 102, 105], [18-39 every 3rd, 45-63 every 3rd, 69-81 every 3rd, 91-93, 110-111, 113, 117-118]	Oct 21	Lecture: 2.5, HW: pg 244: [W: 59-61, 63], [3, 6, 12, 15, 21, 27, 36-51 every 3rd]
Oct 24	Lecture: 2.5, HW: pg 244: [W: none], [53-56, 66-74]	Oct 26	Lecture: 2.6, HW: pg 257: [W: 72, 75], [3, 9-21 every 3rd, 27-45 every 3rd]	Oct 28	Lecture: 2.6, HW: pg 257: [W: none], [47-58, 68, 69]
Oct 31	Lecture: 2.7, HW: pg 267: [W: 47-50], [3, 6, 9, 15, 18, 21, 24, 27a, 30-31, 33-35, 38-39]	Nov 2	Lecture: P5, 2.7: pg 267: [W: none], [43-45, 47, 52, 57-58, 64-66]	Nov 4	Lecture: 3.1, HW: pg 288: [W: 52-53, 57], [1-9, 12-21 every 3rd, 27; 33-37 odd; 41, 43; 47, 49-50, 59, 63, 66-70]
Nov 7	Lecture: 3.2, HW: pg 300: [W: 57-58, 60, 62], [3-9 odd, 11-15, 18, 20-21, 24-36 every 3rd, 42, 45, 51, 54, 80-82]	Nov 9	Lecture: 3.2 Handout, HW: worksheet	Nov 11	**Veterans Day Holiday**

Treatment Homework List – Continued

Nov 14	Review/Help	Nov 16	Lecture: 3.3, HW: pg 313: [W: none], [3-11 odd, 12, 27, 29]	Nov 18	Last day to drop with a w Lecture: 4.1, HW: pg 382: [W: none], [13, 19-20, 23-24, 29-43 odd, 44, 46-47, 49, 55-57, 66, 68-70]
Nov 21	Lecture: 4.2, HW: pg 395: [W: 89-90, 92], [3-21 every 3rd, 27-39 every 3rd, 43-48; 51-69 every 3rd, 75-87 every 3rd; 102ab, 104-107]	Nov 23	Lecture: 4.3, HW: pg 405: [W: 85-87, 90], [3, 6, 9, 21, 24, 30, 33, 36, 42, 48, 51, 54, 60, 63-75 every 3rd, 81, 84, (93, 94), 102-104]	Nov 25	**Thanksgiving Holiday**
Nov 28	Lecture: 4.4, HW: pg 415: [W: 71, 72], [9, 11, 13, 19-27 odd, 31, 35-43 odd, 47-52]	Nov 30	Lecture: 4.4, HW: [W: none], [53-63 odd; 69, 79, 87-88, 91]	Dec 2	Lecture: 4.5, HW: pg 428: [W: 36, 38, 39], [1-3, 5-6, 8-9; 14, 17, 19, 21, 24, 26]
Dec 5	Lecture: 5.1, HW: pg 452: [W: 69-70], [3, 5, 9, 11-12]	Dec 7	Lecture: 5.1, HW: pg 452: [W: 71-72], [19, 23, 25, 53]	Dec 9	Review
Dec 12		Dec 14	Comprehensive Final 2:00 – 3:50 PM Dec 14	Dec 16	

APPENDIX C:

PROCEDURAL UNDERSTANDING ASSESSMENTS

Journaling Task 1 Name: _____

Section: _____

You will receive 100% on this assignment as long as you make an honest effort to complete *both parts*. There is a second question on the back of this sheet. You have 10 minutes to finish.

Think about the following procedures for solving quadratic equations:

(i) the square root method,
(ii) factoring & the "Zero Product Principle,"
(iii) completing the square, and
(iv) the quadratic formula

Question 1:

In the space below, explain (in general) what *you* look for in a quadratic equation to decide which is going to be the best procedure to use to solve it. It may help to reflect on how you might solve the five equations listed below.

1. $6x^2 - 3x - 2 = 0$
2. $9x^2 - 24x = 0$
3. $(3x - 4)^2 = 16$
4. $x^2 - 6x - 7 = 0$
5. $x^2 - 6x + 7 = 0$

Please be as complete as possible in your response.
You may wish to use one of the following prompts to get started:

The first thing I look for to decide how I'll solve a quadratic equation is…
Before I start solving, I usually look at…
I remember from our class discussions that…

Question 2:

The "Zero Product Principle" is one method for solving quadratic equations. Explain how the Zero Product Principle (ZPP) empowers you as a mathematical problem solver.

In other words, what are the different types of problems you can solve once you have mastered the ZPP?

Please be as complete as possible in your response.
You may wish to use one of the following prompts to get started:

In class, we always used the Zero Product Principle (ZPP) to…
I think the ZPP could also be used for problems like…
An example of where I would use the ZPP is…

Journaling Task 2 Name: _____

Section: _____

You will receive 100% on this assignment as long as you make an honest effort to complete *both parts*. There is a second question on the back of this sheet. You have 10 minutes to finish.

Question 1:

The first step in the procedures for adding, subtracting, multiplying, and dividing rational expressions (like the examples shown below) is always to factor the numerators and denominators.

Ex. 1: $\dfrac{x^2-6x-16}{x^2-9}\cdot\dfrac{x+3}{x^2-2x-8}$ Ex. 2: $\dfrac{4}{x^2+6x+9}-\dfrac{x-7}{x^2-9}$

What is the advantage of factoring the expressions first, before attempting to add, subtract, multiply, or divide?

Please be as complete as possible in your response.
You may wish to use one of the following prompts to get started.

If we didn't factor first we would have to…
By factoring first, we are able to…
The whole point of factoring first is so we can…

Question 2:

There are several related procedures for finding the inverse of a function.

Briefly describe as many methods as you can for finding the inverse.
You may wish to use one of the following prompts to get started.

Sometimes we found the inverse by… but other times we…
Depending the problem, I would either find the inverse by… or by…
The book's method for finding the inverse is to… but we did it these other ways too…

Journaling Task 3 Name: _____

Section: _____

You will receive 100% on this assignment as long as you make an honest effort to complete *both parts*. There is a second question on the back of this sheet. You have 10 minutes to finish.

Question 1:

Here is a typical "polynomial long division" problem:

$$(x^3 + 2x^2 - 3x - 6) \div (x + 2)$$

Describe some ways you can verify that you have done the long division procedure correctly. Don't just work the problem provided. Instead, try to list any ideas you could use to verify that your answer is correct.

Please be as complete as possible in your response.
You may wish to use one of the following prompts to get started:

One sure sign that I've made a mistake is…
I've noticed that all of the correct answers have the form…
If the answer is correct, then I should be able to…

Question 2:

One of the rules for logarithms states that if b, M, and N are positive numbers with $b \neq 1$, then

$$\log_b(MN) = \log_b(M) + \log_b(N)$$

Do your best to think of a way you could convince yourself (and me!) that the rule is correct. (It does not have to be a proof!). Explain your reasoning in the space below.

Please be as complete as possible in your response.
You may wish to use one of the following prompts to get started.

You can tell that the rule is correct because…
This isn't a proof, but this is why I believe the rule is correct…
I remember my teacher justifying the rule by showing us…

APPENDIX D:

PROCEDURAL UNDERSTANDING ASSESSMENT SCORING RUBRICS

Journaling Task 1 Name: _____

Think about the following procedures for solving quadratic equations:

(i) the square root method,
(ii) factoring & the "Zero Product Principle,"
(iii) completing the square, and
(iv) the quadratic formula

Question 1:

In the space below, explain (in general) what *you* look for in a quadratic equation to decide which is going to be the best procedure to use to solve it. It may help to reflect on how you might solve the five equations listed below.

1. $6x^2 - 3x - 2 = 0$ (quadratic formula)
2. $9x^2 - 24x = 0$ (factoring)
3. $(3x - 4)^2 = 16$ (square root method)
4. $x^2 - 6x - 7 = 0$ (factoring)
5. $x^2 - 6x + 7 = 0$ (not factorable; use completing the square or quadratic formula)

Sample Response

- First, try to take the square root of both sides, like in #3. That's the square root method.
- Then, if the expression is easy to factor (it helps if the lead coefficient is 1), then use the zero product principle. Make sure it equals zero first!
- Completing the square works best when the lead coefficient is 1.
- Finally, use the quadratic formula. It works for all quadratic equations, but it is sometimes hard to simplify the result.

Note: By itself, a correct example DOES NOT represent sufficient evidence for any of the items in the sample response. See Level 1.

Level 3 Response: High degree of understanding.

- Adequately addresses three or more of the items in the sample response.

Level 2 Response: Moderate degree of understanding.

- Adequately addresses two of the items in the sample response.
- ALSO, responses that address three or more of the items but lack detail or contain errors should be coded Level 2.
- ALSO, responses that list one method and then say, "if it won't work, I just use quad. form."

Level 1 Response: Low degree of understanding.

- Addresses only one of the items listed in the sample response, possibly using wording such as "it's the only way I know how."
- ALSO, responses that are substantially but not entirely incorrect, or those that only state which method to use for each of the five examples without any explanation should be coded Level 1.
- ALSO, responses which state, "start with quad. form.," and then "if that won't work, use…"

Level 0 Response: No understanding demonstrated.

- Response is missing, vague, off-task, or otherwise fails to demonstrate any understanding of how to choose from among the four available methods.

Question 2:

The "Zero Product Principle" is one method for solving quadratic equations. Explain how the Zero Product Principle (ZPP) empowers you as a mathematical problem solver.

In other words, what are the different types of problems you can solve once you have mastered the ZPP?

Sample Response

- Solving factorable polynomial equation of all degrees (not just quadratics).
- Finding the x-intercepts (roots) of the graph of a polynomial.
- Solving specific application problems (e.g., how far does a projectile travel when fired from a cannon). A reasonable example must be provided in the response.
- Finding the roots of rational expressions (where the numerator equals zero).
- Finding restrictions to the domain of rational expressions (where the denominator equals zero).
- Using the zeros of a graph of a polynomial to write an equation of the polynomial.

Note: A correct example DOES suffice as evidence of any of the items in the sample response.

Level 3 Response: High degree of understanding.

- The response adequately describes two or more of the items in the sample response in addition to at least implicitly acknowledging the ZPP's use in solving or finding zeros of quadratics.

Level 2 Response: Moderate degree of understanding.

- The response adequately describes one of the items in the sample response.

Level 1 Response: Low degree of understanding.

- The response only discusses the use of ZPP for solving or finding zeros of quadratics (even lengthy discussions relating ZPP to other methods should be classified Level 1).
- OR the response provides an example (not necessarily correct) of using ZPP to solve a quadratic equation.
- A response which satisfies any of the preceding Level 1 criteria should be coded Level 1, even those that also reveal significant misconceptions as to the meaning or possible uses of the ZPP.

Level 0 Response: No understanding demonstrated.

- Response is missing, vague, off-task, or otherwise fails to demonstrate any understanding of the how the ZPP can be used to solve mathematics problems.

Journaling Task 2 Name: _____

Question 1:

The first step in the procedures for adding, subtracting, multiplying, and dividing rational expressions (like the examples shown below) is always to factor the numerators and denominators.

Ex. 1: $\dfrac{x^2-6x-16}{x^2-9}\cdot\dfrac{x+3}{x^2-2x-8}$ Ex. 2: $\dfrac{4}{x^2+6x+9}-\dfrac{x-7}{x^2-9}$

What is the advantage of factoring the expressions first, before attempting to add, subtract, multiply, or divide?

Sample Response

- In multiplication or division problems, factoring allows us to cancel any factors that might appear in the numerator of one and the denominator of another. This makes for a shorter simplification process.
- In addition or subtraction problems, we need to find a common denominator. By factoring first, it is easier to find the least common denominator because all factors are clearly visible. Using the least common denominator greatly simplifies the overall problem.
- Factoring also helps us identify the domain restrictions. These should be noted in the final answer.

Note: By itself, a correct example DOES NOT represent sufficient evidence for any of the items in the sample response.

Level 3 Response: High degree of understanding.
- Adequately addresses at least two of the items in the sample response.

Level 2 Response: Moderate degree of understanding.
- Adequately addresses one of the items in the sample response.

Level 1 Response: Low degree of understanding.
- Response states that by factoring first, we are able to "get a simpler equation," or "get rid of extra terms," or "make it easier to solve," or "break down the expression," but the response does not elaborate on why factoring is the key to this process.
- ALSO, responses which briefly mention "finding common factors" without discussing that they can be cancelled or used to find the least common denominator should be classified as Level 1.

Level 0 Response: No understanding demonstrated.
- Response is missing, vague, off-task, or otherwise fails to demonstrate any understanding of why factoring is appropriate.

Question 2:

There are several related procedures for finding the inverse of a function.

Briefly describe as many methods as you can for finding the inverse.
You may wish to use one of the following prompts to get started.

Sample Response

- The book's method is to 1) replace $f(x)$ with y, 2) switch x and y, 3) solve for y, and 4) replace y with $f^{-1}(x)$.
- A variation is to 1) replace $f(x)$ with y, 2) solve for x, and then 3) switch x and y, and 4) replace y with $f^{-1}(x)$. This doesn't really save you any time, but is just another way of thinking about the method.
- If you are given a graph, you can sketch the inverse by reflecting across the line $y = x$.
- If you are given individual points, you just switch the x's and y's.
- You can sometimes just put together the inverse operations in the reverse order, but that only works for functions with just one appearance of x.

Note: A correct example DOES suffice as evidence of any of the items in the sample response.
Note: Verifying that two functions are inverses (e.g., showing that $f(g(x)) = x$) is *not* a method for *finding* an inverse.

Level 3 Response: High degree of understanding.
- The response adequately addresses at least three of the items in the sample response.

Level 2 Response: Moderate degree of understanding.
- The response adequately addresses two of the items in the sample response.

Level 1 Response: Low degree of understanding.
- The response adequately addresses one of the items in the sample response.
- ALSO, responses which address two or more of the items in the sample response BUT which also confuse finding the inverse of a function with finding the multiplicative inverse of a number or finding the slopes of perpendicular lines (negative reciprocals) should be coded Level 1.

Level 0 Response: No understanding demonstrated.
- Response is missing, vague, off-task, or otherwise fails to demonstrate any understanding of the various methods for finding an inverse function.

Journaling Task 3 Name: _____

Question 1:

Here is a typical "polynomial long division" problem:

$$(x^3 + 2x^2 - 3x - 6) \div (x + 2)$$

Describe some ways you can verify that you have done the long division procedure correctly. Don't just work the problem provided. Instead, try to list any ideas you could use to verify that your answer is correct.

Sample Response

- We can verify the answer by multiplying it by (x+2). If the work is correct, we should get the original cubic polynomial. [This is just the division algorithm: $f(x) = g(x)\, d(x) + r(x)$]
- We can verify the answer by plotting it along with the rational function $f(x) = (x^3 + 2x^2 - 3x - 6) / (x + 2)$. If the answer is correct, the graphs will match everywhere on their domains.
- We can verify the answer by computing using a different method such as synthetic division.
- The answer should have degree less than that of the dividend (in this case, it will be 2nd degree), and its terms should appear in descending powers of x.
- The remainder should have degree less than that of the divisor (x+2). In this case, that means it must be a constant (possibly zero).
-

Note: By itself, a correct example DOES NOT represent sufficient evidence for any of the items in the sample response.

Level 3 Response: High degree of understanding.
- The response addresses at least two of the items from the sample response.

Level 2 Response: Moderate degree of understanding.
- The response adequately addresses one of the items in the sample response.

Level 1 Response: Low degree of understanding.
- The response does not adequately address any of the items from the sample response, but suggests some understanding of the form a correct answer might take.
    - For example, a response that correctly states that the answer should "have an x^2, an x, and a number" but that does not extend the response to consider more general cases should be coded Level 1.
- ALSO, responses which adequately address one or more of the items in the sample response but which also reveal significant misconceptions (e.g., asserting the answer is a number that can be "plugged in" to the original to check it, or asserting that a correct answer will be factorable or must not have a remainder) should be coded Level 1.

Level 0 Response: No understanding demonstrated.
- Response is missing, vague, off-task, or otherwise fails to demonstrate any understanding of how one might verify the answer.

Question 2:

One of the rules for logarithms states that if b, M, and N are positive numbers with $b \neq 1$, then

$$\log_b(MN) = \log_b(M) + \log_b(N)$$

Do your best to think of a way you could convince yourself (and me!) that the rule is correct. (It does not have to be a proof!). Explain your reasoning in the space below.

Sample Response

- One way to check the rule is to choose "easy numbers" for b, M, and N.
  - For example, let $b = 10$, $M = 100$, and $N = 1000$.
  - Then we get $\log_{10}(100 * 1000) = \log_{10}(100) + \log_{10}(1000)$.
  - The left hand side is $\log_{10}(10^5)$, or just 5. Similarly, the right hand side is 2 + 3.
- You don't even have to use "easy numbers."
  - If you have a calculator handy you can just do something like ln(5*8) vs. ln(5)+ln(8).
  - (Log_{10} also works.) You'll get the same answer both ways.
  - You can use a calculator to graph both sides of the equation for either ln or $\log_{10}$.
  - For instance, graph $y = \log(x(x+1))$ as well as $y = \log(x) + \log(x+1)$.
  - They will agree on their domains.
- Finally, you can exponentiate both sides and use properties of exponents. Specifically:
  - Take b to both sides.
  - The left hand side ($b^{\log(MN)}$) simplifies to MN.
  - The right hand sides becomes $b^{\log(M) + \log(N)}$.
  - Simplifying the right hand side using the law of exponents gives $b^{\log(M)}\, b^{\log(N)}$.
  - That simplifies to just MN, which matches the left hand side.

Note: A correct example DOES suffice as evidence of any of the items in the sample response.

Level 3 Response: High degree of understanding.

- The response correctly justifies the logarithm rule by using at least one of the techniques listed in the sample response.

Level 2 Response: Moderate degree of understanding.

- The response either
  - correctly demonstrates the inverse relationship between exponents and logarithms,
  - OR the response correctly states some relevant properties of exponents,
  - OR the response alludes to one of the items from the sample response,
  - BUT the response ultimately fails to clearly and correctly justify the logarithm rule.

Level 1 Response: Low degree of understanding.

- The response attempts to justify the logarithm rule using the properties of exponents, but does so using incorrect or irrelevant properties, or is very vague.
  - For example, a response that uses the property "$b^x + b^y = b^{xy}$" (an incorrect property of exponents) to justify the logarithm rule should be coded Level 1.

Level 0 Response: No understanding demonstrated.

- Response is missing, vague, off-task, or otherwise fails to demonstrate any understanding of how one might justify the property.

APPENDIX E:

PROCEDURAL UNDERSTANDING ASSESSMENT CODING NOTES

Coding Notes – Theoretical Perspective

1. These instruments are designed to assess the depth and diversity of students' knowledge of Framework-oriented ideas, which served as the basis for the instructional treatment.

2. In light of Note 1, poor wording or incorrect language should generally not influence coding decisions. Many of the responses are written using improper mathematical language. Because the treatment students had more practice writing mathematics during the semester, it would give them an unfair advantage to penalize for incorrect use of mathematical language. Remember, it is the depth and diversity of ideas that the instrument is intended to capture.

   So on Task 1, Question 1, suppose a student writes the following:

   > First, I look to see if I can take the *square* of both sides evenly. If I can, then I'll use the Square Root Method.

   Aside from the incorrect language, the response suggests that the student knows when to apply the procedure, and it should be coded accordingly.

3. In some cases, students really stretched to come up with a diversity of ideas to answer the question as fully as possible. As a result, *portions* of their answers are often incorrect. Unless otherwise noted on the scoring rubric, scores should not be reduced solely for those reasons.

4. "Level 0 – No Understanding Demonstrated" should be assigned to responses that essentially avoid the question or that are so vague that the student appears to be trying to talk him/herself into an answer that never comes.

5. "Level 0" should also be used for responses that answer a different question, or have a different focus, than the question that was posed. For this reason, some lengthy responses will be coded "Level 0."

Other Notes About The Tasks

6. The tasks were given in pairs at the beginning of class on three separate occasions. Students had 10 minutes to complete each pair. Question 1 was on the front and Question 2 was on the back.

7. Because these tasks were given as three separate pairs, some students will have completed Task 1 but not Task 2 or Task 3. Therefore, when all responses for a particular Question have been coded, several blank entries in the list should be expected. These correspond to students who were absent.

APPENDIX F:

CLASSROOM OBSERVATION TRAINING SCENARIO

Example of a Sequence of Classroom "Events"

Classroom Example – Solving a Quadratic Equation.

The instructor writes the following example on the board:

Solve: $x^2 + 2x - 4 = -1$

He then turns to the class and says, "Why can't I apply the quadratic formula yet?"

A student in the front row answers, "it's not equal to zero."

"Good!" he says, "we have to subtract negative one from both sides first to set it equal to zero. Then we can apply the quadratic formula." After simplifying, he writes:

$$x^2 + 2x - 3 = 0$$

"Now we have a quadratic equation in the proper form." He pauses for a second and notices that this quadratic is easy to factor. He says, "we could just go ahead and apply the quadratic formula (you know, 'x equals negative b plus or minus the square root of…' and so on), but what if we try to factor it instead? Sometimes that's quicker. Let's see if that works in this case." He writes:

$$(x \quad)(x \quad)$$

Then he asks, "what do I look at first?" When no one answers, he rephrases his question and asks, "what do I look for to figure out what the signs need to be?"

A student in the front row answers, "the negative sign in front of the three."

"Yep, and since it's negative, these signs are different," he confirms. "So we need two factors of three that differ by two," he says, pointing at the relevant coefficients of the quadratic he is factoring.

A student in the front row says, "plus three and minus one."

The instructor writes:

$$(x + 3)(x - 1)$$

"Good, I think that does it," the instructor agrees. "Now, are we finished with the problem yet?" He pauses. "Hmm… nope, we aren't. We are supposed to solve the equation, but all we've done is factored part of it." He writes,

$$(x + 3)(x - 1) = 0$$

"Now we can apply the zero product principle to finish." He writes,

$$(x + 3) = 0 \quad \text{or} \quad (x - 1) = 0 \qquad \text{(by zero product principle)}$$

"So we have x = -3 or x = 1." He writes the answer on the board. "Notice that we got two answers," he says. "That's because we started with a quadratic: quadratics always have zero, one, or two roots."

Then he adds, "What if we had used the quadratic formula instead, way back up here," he says, pointing to the line:

$$x^2 + 2x - 3 = 0$$

"We'd have: negative b plus or minus the square root of b squared minus four a c, all over two a," he says slowly while writing and pointing to the relevant coefficients:

$$x = \frac{-2 \pm \sqrt{(2)^2 - 4(1)(-3)}}{2(1)}$$

"Alright, now we simplify," he says, and proceeds to walk through the process of simplification, ultimately showing that the *x* values obtained are, indeed, -3 and 1.

Then, he compares the two approaches. "Which method would you say was easier? The factoring method or the quadratic formula method?" There is a pause as he lets the class think. "Mark, what do you think? Which method would you have done?"

Mark answers, "Factoring. It took less steps."

The instructor responds, "uh huh. It does look like it took fewer steps in this case. Can anyone think of a situation where the quadratic formula would be the preferred? Instead of factoring?"

A student raises her hand and says, "Not everything can be factored."

"That's good. And even when an expression can be factored, it might not be quite as easy as this one was. Good. What else?"

After a long pause, another student says, "My high school math teacher told us to always just use the quadratic formula, because it always works." The instructor nods. The student pauses, and then adds, "Do we have to know how to use factoring on the exam?"

The instructor says, “You have to know how to do it both ways. Sometimes factoring is a lot quicker. Using the quadratic formula on problems like the one we just finished is like using a sledgehammer to push in a thumbtack: you might end up using a lot more effort than you need.

“Alright, so we solved the same quadratic equation two different ways.” The instructor does a quick summary, saying, “notice that we had to set it equal to zero first, no matter which method we used. Don’t forget to set it equal to zero first.” He then writes it on the board:

Always set quadratic equations equal to zero before solving.

“Good, let’s do another one,” he says, and proceeds to write the following on the board:

Solve: $3x^2 - 4x + 3 = 2$

“Everyone work through this example on your own. Check your answer with your neighbor when you finish. If you get stuck, see if your neighbor can help you get going again. I’ll circulate the room to see how everyone is doing.” The instructor begins to circulate around the room, as most students work independently. A few are off task, and a few occasionally whisper to their neighbor to ask for help.

The instructor pauses near a student who has not yet begun working, and then continues walking and addresses the class, “what do you always do first?”

“Set it equal to zero,” someone says from the front row.

“Good,” the instructor says, and he continues circulating. Once he sees that about half the students are finished, he says, “alright, let’s see how you did,” and proceeds to work the problem quickly on the front board.

APPENDIX G:

CLASSROOM OBSERVATION FORM

Checklist

Brief Instructions**:** An *event* refers to a classroom episode in which a Framework objective is addressed. As you observe the lesson, place a tick mark in the appropriate box for each event that you observe.

- An event in which a majority of students are engaged (either through dialog, discussion, or writing) should be marked under the *Active* heading.
- An event in which the instructor is lecturing, having a dialog with just one or two students, or answering his or her own question, should be marked under the *Passive* heading.

Framework Objective	**Related Questions**	Passive **(Q&A / Monolog / Lecture)**		Active **(Class Discussion / Dialog / Tasks)**	
		Emphasis (focus of instructor's lecture / monolog; esp. when written on the board)	**Aside** (mentioned in passing; many students may have missed it)	**Emphasis** (focus of class discussion, dialog, writing, or calculator task)	**Aside** (secondary outcome of class discussion, dialog, writing, or calculator task)
1a. The **Overall Goal** of the procedure.	1a. "What are we trying to accomplish?"				
1b. **Predicting** & **Estimating**.	1b. "What sort of answer should we expect?"				
2a. **Performing** the procedure.	2a. "How do we carry out this procedure? What are the steps?"				
2b. **Alternate Methods / Representations**.	2b. "How else could we have done this?"				
3. Why the procedure is **Effective** & **Valid**.	3. "Why does this work? Why is it valid?"				
4. **Evaluate Results** by using context, other procedures, etc.	4. "How can we verify the answer? Does it make sense?"				
5. Assess relative **Efficiency** & **Accuracy**.	5. "What is the most efficient method to use?"				
6. **Empowerment** as a problem solver.	6. "What types of problems can we solve with this?"				

Classroom Observations – Overall Emphasis on Procedural Understanding

Brief Instructions:

- Complete this section immediately after you have observed the lesson.
- Base your responses on your overall impression of the lesson, guided by the events you recorded on the checklist.
- Remember that "infrequent" and "frequent" should also reflect the duration of the events you observed.
- See the sample below to note how to mark your response. **Please use whole numbers (0 – 6) only.**

	Absent 0	1	Infrequent 2	3	Frequent 4	5	Pervasive 6
0a. SAMPLE – This is marked a "2 - Infrequent" (Use an "x" to mark the spot)			X				
0b. SAMPLE – This is marked a "5 - Frequent to Pervasive" (Use an "x" to mark the spot)						X	

Mark the locations on each of the 8 continua below that correspond to your independent evaluation of the extent to which each Framework objective was addressed during the lesson.

	Absent 0	1	Infrequent 2	3	Frequent 4	5	Pervasive 6
1a. Overall Goal of Procedure ("What are we trying to accomplish?")							
1b. Predicting & Estimating ("What sort of answer should we expect?")							
2a. Performing the Procedure ("How do we carry out this procedure?")							
2b. Alternate Methods & Representations. ("How else could we have done this?")							
3. Why the procedure is Effective & Valid ("Why does this work? Why is it valid?")							
4. Evaluating the Results using context, etc. ("How can we verify our answer? Does it make sense?")							
5. Efficiency & Accuracy ("What is the most efficient method to use?")							
6. Empowerment as a problem solver ("What types of problems can we solve with this?")							

APPENDIX H:

SUPPLEMENTAL COURSE EVALUATION

Supplemental Course Evaluation Questions

Instructions: Please circle the appropriate number to indicate your response.

Prompt: *Compared with previous algebra courses you have taken, how effective would you say this course has been in developing…*

…your knowledge of the overall goals of various College Algebra procedures? (1 = Much better than prev. courses, 5 = Much worse than prev. courses)	1 2 3 4 5 much better … much worse
…your knowledge of how to predict or estimate the outcome of various College Algebra procedures? (1 = Much better than prev. courses, 5 = Much worse than prev. courses)	1 2 3 4 5 much better … much worse
…your knowledge of how to correctly execute various College Algebra procedures? (1 = Much better than prev. courses, 5 = Much worse than prev. courses)	1 2 3 4 5 much better … much worse
…your knowledge of alternate methods for solving various College Algebra problems? (1 = Much better than prev. courses, 5 = Much worse than prev. courses)	1 2 3 4 5 much better … much worse
…your knowledge of why various College Algebra procedures lead to valid results? (1 = Much better than prev. courses, 5 = Much worse than prev. courses)	1 2 3 4 5 much better … much worse
…your knowledge of ways you can verify your answer after using various College Algebra procedures? (1 = Much better than prev. courses, 5 = Much worse than prev. courses)	1 2 3 4 5 much better … much worse
…your knowledge of how to determine when certain College Algebra procedures are more effective than others? (1 = Much better than prev. courses, 5 = Much worse than prev. courses)	1 2 3 4 5 much better … much worse
…your knowledge of the ways that various College Algebra procedures can be used to solve problems? (1 = Much better than prev. courses, 5 = Much worse than prev. courses)	1 2 3 4 5 much better … much worse

How would you rank the effectiveness of the Writing in Mathematics homework tasks in terms of deepening your understanding of Math105 topics? (1 = Very helpful, 5 = Not at all helpful)	1 2 3 4 5 very helpful … not at all helpful

Open Response: Please comment on the following questions. Please be as specific as possible. (You do not need to address each question in turn.)

THANK YOU FOR YOUR HONEST FEEDBACK!

This course has had an increased emphasis on *understanding* of College Algebra procedures.

- Did you notice? How so?
- Do you think it has been effective? Why or why not?
- Have you developed a better understanding of mathematics than you would in a more traditional course? Explain.

APPENDIX I:

INTERVIEW SCRIPT

Guide for Interview with Treatment Instructors

How difficult was it to incorporate Framework-oriented topics in your lessons?

Was there sufficient time to cover the material in the class while trying to emphasize understanding?

How did your students respond when you tried to incorporate Framework-oriented topics into your lessons?

How receptive were the students to your attempts to teach for understanding?

Compared with other semesters you have taught, how much time would you say it took you to prepare your lessons this semester?

How did you use the materials I handed out at each meeting?

To what extent did you address the Writing in Mathematics questions in your class?

How would you characterize your students' responses to the Writing in Mathematics questions?

What advice would you give to future teachers who might be asked to teach in this way?

What would you do differently if you had to teach the same curriculum again next semester?

Do you have any suggestions for improvement of this teaching method?