SUSIE MAE LANIER
Students' Understanding of Linear Modeling in a College Mathematical Modeling Course
(Under the direction of JAMES W. WILSON)

The purpose of this study was to investigate college students' understanding of linear modeling when using a spreadsheet template to model population data in a mathematical modeling course. Schoenfeld's (1992) "framework for exploring mathematical cognition" was used to examine the students' mathematical thinking and problem solving during the course. The framework consisted of five categories: the knowledge base, problem-solving strategies, monitoring and control, beliefs and affects, and mathematical practices. These categories provided an organized structure for decomposing the students' understanding of linear modeling into manageable parts and analyzing these parts. Because of the coherent nature of the categories, they also provided a lens for looking at a students' understanding of linear modeling as a whole.

The study was conducted during fall semester of 1998. A qualitative case study approach was used for this research. Data were collected from observations, interviews, and written documents. The data were then analyzed according to the qualitative method of constant comparison that was described by Corbin and Strauss (1990).

Four main themes emerged from the data analysis. First, the students were procedurally oriented. They seemed obsessed with their procedures for finding the optimal linear model. Second, the students treated the spreadsheet template as a "black box," and hence, failed to make effective use of available representations of the linear modeling situation. Third, the students' life experiences influenced their interpretation and sense-making (mathematical practices) of the modeling situation. Finally, the students formed opinions, made decisions, and adequately communicated their ideas about linear modeling when pressed to do so.

INDEX WORDS: Mathematical Understanding, Mathematical Thinking, , Linear Modeling, Mathematical Modeling, Spreadsheet, Spreadsheet Template, Problem Solving Strategies, College, Qualitative Research

STUDENTS' UNDERSTANDING OF LINEAR MODELING IN A COLLEGE MATHEMATICAL MODELING COURSE

by

SUSIE MAE LANIER

B.S.E.D., Georgia Southern College, 1981

M.S.T., Georgia Southern College, 1983

A Dissertation Submitted to the Graduate Faculty

of The University of Georgia in Partial Fulfillment

of the

Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

1999

STUDENTS' UNDERSTANDING OF LINEAR MODELING

IN A COLLEGE MATHEMATICAL MODELING COURSE

by

SUSIE MAE LANIER

Approved:

____
Major Professor

____
Date

Approved:

____
Dean of the Graduate School

____
Date

In loving memory of

my father, H. G. Lanier, Jr.; my grandfather, H. G. Lanier, Sr.;
and my Georgia Southern Advisor, Dr. Malcolm Smith.

I know they are celebrating with me
in spirit.

ACKNOWLEDGEMENTS

This dissertation would certainly not exist without the support, guidance, and encouragement I received from many people in my life.

I cannot thank Adam, Cindy, and Kaitlyn enough for their time, patience, and willingness to help me with this study. I would also like to thank Dr. Kirk for allowing me in his classroom. Without the four of you, there would be no dissertation.

My deepest thanks goes to my major professor, Jim Wilson, for his guidance, wisdom, understanding, and encouragement. You are a wonderful teacher and role model for aspiring professors of mathematics education.

I would also like to thank my doctoral committee, Ed Davis, Pat Wilson, Denise Mewborn, and Sybilla Beckman for all their time and advice. You are a great committee, and I am fortunate to have been able to work with each of you. A special thank you goes to Denise and Sybilla for listening to my ramblings as I gathered my thoughts for this study and for believing in my ability to reach this goal.

I would like to say a special thank you to Jeremy Kilpatrick for his guidance as a teacher and our Wednesday morning talks fall semester. I have learned so much from you.

I would like to say thank you to all my colleagues at Georgia Southern University for their encouragement and support. I am especially grateful to the administration at Georgia Southern for recommending me for the Faculty Development in Georgia (FDIG) assistanship and allowing me to be on leave the past three years. A special thank you goes to my former department chairs, Curtis Ricker and Arthur Sparks, and my current department chairs, Janet O'Brien and Don Faucett.

I would like to thank Libby Morris for her leadership in the FDIG seminars held in the Institute of Higher Education at UGA the past three years.

There are so many family members to thank who have supported and encouraged me in this endeavor: my mother, Millie; my brothers, Jim and Bubba; my niece, Gracy, who has brought tremendous joy to an often stressful period in my life; my sister-in-law, Jennifer; and the remainder of my family, who are too numerous to mention. I love you all.

There are also many friends that I must thank: all the gang in Statesboro, Garfield, and Athens. I would especially like to thank Joyce and Ilene for their unwavering support and encouragement. You "kept me going" over the past three years.

Most important, I thank God for his many blessings. From my broken ankle to the completion of this study, He has walked beside me.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE

CHAPTER ONE

THE PROBLEM AND ITS BACKGROUND

From my recent study of mathematics curriculums, as well as my experiences as a student and a teacher, it became apparent that technology has greatly influenced current mathematics classrooms. The National Council of Teachers of Mathematics (NCTM) *Curriculum and Evaluation Standards* (1989) was only one of many sources that addressed the uses of technology in mathematics classrooms. The *Standards* claimed that "new technology not only has made calculations and graphing easier, it has changed the very nature of the problems important to mathematics and the methods mathematicians use to investigate them" (p. 8). Graphing calculators and computers have become prevailing forms of this technology. According to Demana and Waits (1993):

> Computer generated numerical, graphical, and symbolic mathematics is revolutionizing the teaching and learning of mathematics. The computer can be a desktop computer with a computer algebra system or a pocket computer with software built-in (a graphing calculator). The content of mathematics is changing. Reduced time is spent on paper and pencil methods and increased time is spent on application, problem solving, and concept development. Instructional methods are also rapidly changing. Investigative, exploratory methods are becoming more common in mathematics courses. (p. 1)

Kilpatrick and Davis (1993) also alluded to this change in mathematics content and instruction. They suggested that computers are "changing the ways in which mathematics is done by stimulating the use of numerical methods and modeling and by promoting the study of algorithms" (p. 203). Most recently, the draft of NCTM's *Principles and Standards for School Mathematics* (1998) included technology as one of six principles. This principle is stated below:

> **The Technology Principle**: Mathematics instructional programs should use technology to help all students understand mathematics and should prepare them to use mathematics in an increasingly technological world. (p. 22)

The discussion of this principle contained the following suggestion:

> Like any tools, technological tools can be used well or poorly. They should not be used as replacements for basic understandings and intuitions; rather, they can and should be used to foster those understandings and intuitions. Within mathematics instructional programs, technology should be used responsibly, with the goal of enriching student learning of mathematics. (p. 40)

I agree that technological tools should not replace basic understandings and intuitions; however, I do think they can and should be used to build these understandings and intuitions. Further comments concerning technology in the draft focused on how technology might best support mathematics learning and how the presence of technology implies shifts in mathematical content and the ways in which students' thinking might be different (p. 17).

Along with technology, mathematical modeling has become a common topic in mathematics education discussions. Dossey (1990) defined mathematical modeling as "the process by which real-world situations are represented in mathematical terms" (p. 3). He further remarked that many problems can be solved by "creating a mathematical model, manipulating the model, interpreting the possible solutions, and validating them in the original problem situation" (pp. 3–4). According to Hilke (1995), "Mathematical modeling is part of a growing reform movement in mathematics instruction" (p. 8). Koss and Marks (1994) claimed that this reform effort "fosters growth in each student's mathematical thinking, through active exploration, communication of ideas, and reflection over an extended period" (p. 616). Mathematical modeling is a common theme throughout the *Principles and Standards* (1998) draft. Data analysis was emphasized in the fifth standard. Also, one component of the tenth standard (representation) suggested that students "use representation to model and interpret physical, social, and mathematical phenomena" (p. 94). This modeling included "not only representation, but also acting upon the representation and interpreting the meanings of one's actions within the mathematical model and with respect to the phenomenon being modeled" (pp. 98-99).

Advancements in technology, new instructional methods, and encouragement from groups such as NCTM are among the many factors that have motivated post-secondary, as well as secondary schools, to re-examine their mathematics courses. Colleges and universities, such as the University of Georgia (UGA), have incorporated technology into already existing mathematics courses and developed new mathematics courses that require students to use technology to model problem situations that may occur in everyday life.

UGA'S Mathematical Modeling Course

Students at the University of Georgia may enroll in a mathematical modeling course that uses current technology to solve "real world" problems (Edwards, 1997). Primarily, non-science majors and other students whose major requirements do not demand specific preparation for pre-calculus or calculus enroll in this course. The course focus is on mathematical modeling and the use of elementary mathematics -- numbers and measurement, algebra, geometry, and data exploration -- to investigate real-world problems and questions. According to the course developer, Henry Edwards (1997), a primary objective of the course is to develop the "quantitative literacy and savvy that graduates need to function effectively in society and the workplace." More specifically, the objectives are to motivate students

- to combine necessary skill development with the ability to reason and communicate mathematically,
- to use elementary mathematics to solve applied problems, and
- to make connections between mathematics and the surrounding world.

The course is divided into three units consisting of (1) iteration and natural growth processes, (2) linear and quadratic models of data and phenomena, and (3) optimization. Applications and the ability to construct useful mathematical models, to analyze them critically, and to communicate quantitative concepts effectively are emphasized throughout the course.

Technology is an essential element in the design of the course. Students use graphing calculators for work in class, on tests, and on assigned laboratory projects. However, students complete the majority of computation and exploration for a required project with spreadsheet technology (in particular, spreadsheet templates created by Edwards). They also use word processors to prepare the project reports. Thus, the students write about mathematics and, according to Edwards (1997), gain experience with the real world's "principal technology for numerical calculation." The course materials include an on-line text, the slide presentation for each class lecture, spreadsheets illustrating computations and explorations, a course syllabus, project assignments, and a class log. The course materials are available in electronic form (UGA Department of Mathematics, 1998) with the on-line text also available in an optional printed and bound form. Students have access to these materials by using designated computer laboratories, by accessing the world wide web from other campus computers or personal home computers, and by copying files from laboratory computers for use with personal home computers. Freeware viewers are available to students who use home computers but do not have the necessary software for reading and printing the material. In addition, the use of e-mail is encouraged among students and the instructor.

Statement of the Problem

The impact of current changes in curriculum and instruction on students' mathematical understanding is uncertain. Will the goals and objectives of these courses be attained? Will students enrolling in these courses be mathematically literate and productive in society? Dugdale, Thompson, Harvey, Demana, Waits, Kieran, McConnell, and Christmas (1995) suggested:

> We consider that education research will be an essential component of reform efforts; it is essential that we evolve deep understanding of the potential and actual consequences of changes we propose or implement. (p. 347)

Mathematics educators, as well as other researchers, must conduct studies to determine the success of reaching course goals and the impact of these courses on students' mathematical thinking and understanding.

UGA's mathematical modeling course was an attempt at encouraging students to use current technology to solve problems, to organize and communicate their thoughts about the problems, and to make connections between the problems and the real world. Because the majority of the students enrolled in this course were not seeking mathematically related occupations, the course provided a rich environment for studying the mathematical thinking and understanding of students who represent a specific population of individuals entering college. How do these students make sense of a mathematical situation such as linear modeling? How do they organize, store, retrieve, and use their knowledge? Also, how does technology influence the connections they make between the given situation and the real world? These were some of the questions that provided the stimuli for this study.

Purpose of the Study

The purpose of this study was to investigate college students' understanding of linear modeling when using a spreadsheet template to model data in a mathematical modeling course.

Research Question

The research question was:

How are the students'

- knowledge base,
- problem-solving strategies,
- monitoring and control processes,
- beliefs and affects, and
- practices

manifested in their learning of linear modeling?

Definitions

Understanding. To understand a subject is to be able to use knowledge "wisely, fluently, flexibly, and aptly in particular and diverse contexts" (Wiggins, 1993, p. 207). For this study, how students come to know and understand a subject referred to the experiences students had and how they (1) accepted or rejected ideas evolving from the experiences, (2) integrated and stored these ideas with existing knowledge, and (3) retrieved and used these ideas in other contexts or situations. A student's understanding of a linear modeling situation was characterized by the student's knowledge base, use of problem-solving strategies, effective use of resources (monitoring and control), beliefs and affects, and engagement in mathematical practices.

Linear Modeling. For this study, linear modeling was the process by which students used a spreadsheet template [Appendix A] to determine the equation and graph of a line that best fit a set of population data. Students then used this equation and graph to interpret situations.

Theoretical Framework

To study college students' understanding of linear modeling, it was necessary to examine their mathematical thinking and problem solving behaviors in an organized manner. Schoenfeld's (1992) "framework for exploring mathematical cognition" provided such organization. This theoretical framework was an overarching structure that provided a "coherent and relatively comprehensive near decomposition of mathematical thinking (or at least, mathematical behavior)" (p. 363). It was composed of five categories: the knowledge base, problem-solving strategies (heuristics), monitoring and control (self-regulation), beliefs and affects, and practices.

The Knowledge Base. Schoenfeld (1992) referred to the knowledge base as the "mathematical tools an individual has at his or her disposal" (p. 349). He identified two issues related to this knowledge base. The first issue concerned determining the information relevant to a given mathematical situation that an individual possesses. That is, what are the individual's memory contents. The memory contents may include informal and intuitive knowledge about the mathematical situation, facts and definitions, algorithmic procedures, routine procedures, relevant competencies, and knowledge about the rules of discourse in mathematics. The second issue concerned the manner in which the information contained in the memory contents is organized, accessed, and processed.

Schoenfeld cautioned that the knowledge base may contain information that is not true. Students may have misconceptions that they bring to the problem situation. However, this false information cannot be ignored. It is part of the individual's mathematical tools.

Problem-Solving Strategies. Schoenfeld (1983) described problem-solving strategies or heuristics as general suggestions that "help problem solvers approach and understand a problem and efficiently marshall their resources to solve it" (p. 9). Polya's (1945) book, *How to Solve It*, is one of the best known sources of information on problem-solving strategies. Polya suggested that students could use many different strategies for solving problems. Some of these strategies included examine special cases, look for a related problem, look for patterns, and work backward.

Monitoring and Control. According to Schoenfeld (1992), monitoring and control or self-regulation is "one of three broad arenas encompassed under the umbrella term *metacognition*" (p. 354). The core components of self-regulation are monitoring and assessing progress as you work on a problem and acting in response to the assessments of progress (p. 355). Schoenfeld identified resource allocation as the important issue in this category. When confronted with a mathematical situation, an individual may engage in activities such as reading, analyzing, exploring, planning, implementing, and verifying. These activities may occur more than once and in varying orders as the mathematical situation is resolved. Thus, the issue concerns not just what students know, but how, when and whether they use what they know. That is, did the students make effective use of their resources.

Beliefs and Affects. Schoenfeld (1992) defined beliefs as "an individual's understandings and feelings that shape the ways that the individual conceptualizes and engages in mathematical behavior" (p. 358). These beliefs are "abstracted from one's experiences and from the culture in which one is embedded" (p. 360). Lester and Kroll (1990) described affects as an individual's feelings, attitudes, and emotions that may dominate the individual's thinking and actions in solving problems. Examining a student's beliefs, attitudes, and emotions can help determine the student's mathematical perspective or point of view.

Practices. Students' mathematical practices include their ways of interpreting and making sense of situations and ideas. Resnick (1988) commented on the importance of interpretation and sense-making:

> Becoming a good mathematical problem solver -- becoming a good thinker in any domain -- may be as much a matter of acquiring the habits and dispositions of interpretation and sense-making as of acquiring any particular set of skills, strategies, or knowledge. (p. 58)

She further proposed that these practices are socially developed:

> If we want students to treat mathematics as an ill-structured discipline -- making sense of it, arguing about it, and creating it, rather than merely doing it according to prescribed rules -- we will have to socialize them as much as instruct them. (p. 58)

Schoenfeld (1987) also alluded to the social aspect of mathematical practices when he suggested that the "practice of mathematics is a human endeavor and very much a cultural one." These practices of interpretation and sense-making can have a strong influence on what students learn and understand about a mathematical situation.

Rationale for the Study

Schoenfeld (1983) asked, "What do we want our students to get out of the mathematics courses they take?" (p. 7). His reply to this question was:

> The real service we can offer our students, both our majors and the ones we will never see again, is to provide them with thinking skills that they can use after they take our final exams. (p. 7)

This should be a goal of all mathematics educators. As Schoenfeld suggested, we want students to be able to think and to apply their thinking skills in other situations. To what degree are we accomplishing this goal? One way to measure our success is to conduct studies about students' mathematical thinking and how they come to understand mathematics. This study provided a means of assessing how well we were meeting this goal in UGA's mathematical modeling course.

Schoenfeld (1983, 1985a, 1992), Lester (1985), Lester and Kroll (1990), Mayer and Hegarty (1996), and Searcy (1997) are among the researchers who have attempted to characterize the problem solving processes and mathematical thinking of students. This study allowed me to compare and contrast my findings with their established findings. The results of this study provided a "picture" of how students enrolled in this type of mathematical modeling course

- integrated new information into their existing knowledge,
- organized and stored this information,
- made decisions (with regard to strategies and solutions), and
- used the mathematical tools available to them (such as calculators and

computers).

This information added to the existing research on problem solving and mathematical thinking.

In addition, very little of the established research concerned the use of spreadsheet technology, especially spreadsheet templates, in problem solving. I investigated the influence of a spreadsheet template on students' mathematical understanding of linear modeling. Thus, this study enhanced and extended the current knowledge in this area of mathematics education.

Finally, this study contributed to the needed research suggested by Dugdale et al. (1995). It provided a measure of how well the goals and objectives of the mathematical modeling course were being met. The results of this study may persuade teachers to examine and possibly reshape their methods of instruction in mathematics courses. The more we learn about our students' mathematical thinking and problem solving behaviors, the better we will become at finding suitable methods of instruction that will encourage desirable thinking and behaviors. Studies, such as this one, help us understand and educate our students better.

CHAPTER TWO

REVIEW OF RELATED LITERATURE

The purpose of this study was to investigate college students' understanding of linear modeling when using a spreadsheet template to model data in a mathematical modeling course. To accomplish this investigation I focused on students' mathematical thinking and problem-solving behavior in the modeling situation. Along with the framework for exploring mathematical cognition (Shoenfeld, 1992) used in this study, I reviewed other literature pertaining to problem solving, mathematical thinking, and modeling. This chapter presents that literature.

Problem Solving and Mathematical Thinking

As I began this study, I assumed that the students enrolled in the mathematical modeling course would be engaged in problem solving. Hence, I searched for the meaning of a problem and problem solving.

According to Duncker (1945), a problem exists when a person "has a goal but does not know how this goal is to be reached" (p. 1). Mayer and Hegarty (1996) rephrased Duncker's description as "you have a problem when a situation is in a given state, you want the situation to be in a goal state, and there is no obvious way of moving from the given state to the goal state" (p. 31). In his book, *Mathematical Discovery (Volume 2)*, Polya (1965) added, "an essential ingredient of a problem is the desire, the will, and the resolution to solve it" (p. 63). In these interpretations of a problem, there is a situation, a desired goal, and a lack of a path to that desired goal.

Few of the "problems" in the mathematical modeling course fit the above interpretation. There existed a situation (population data) and a desired goal (find the optimal model), but the students were given a path (spreadsheet template) to the desired goal. According to Schoenfeld (1985a), if an individual has “ready access to a solution schema for a mathematical task, that task is an exercise, not a problem” (p. 74). In a later article, he described these problems as “exercises organized to provide practice on a particular mathematical technique that, typically, has just been demonstrated to the student” (Schoenfeld, 1992, p. 337). It became clear during my class observations that most of the “problems” assigned to the students were more like Schoenfeld’s “exercises.” Mayer and Hegarty (1996) referred to such exercises as “routine problems.” Their terminology is not new. Polya (1945) also defined “routine problems”:

> In general, a problem is a “routine problem” if it can be solved either by substituting special data into a formerly solved general problem, or by following step by step, without any trace of originality, some well-worn conspicuous example” (p. 158).

Along with the multiple meanings of “problem," there are different meanings for “problem solving.” Stanic and Kilpatrick (1988) provided a historical perspective on problem solving and hinted at some of these different meanings. They suggested that problem solving encompasses “different views of what education is, of what schooling is, of what mathematics is, and of why we should teach mathematics in general and problem solving in particular” (p. 1).

Problem solving has been linked to mathematical thinking by numerous people. Duncker (1945) described problem solving in relation to thinking:

> Whenever one cannot go from the given situation to the desired situation simply by action, then there has to be recourse to thinking. Such thinking has the task of devising some action which will mediate between the existing and desired situations. (p. 61)

In his writing, Duncker pointed out that we can think of problem solving as consisting of successive reformulations of an initial problem. We encounter a problem. We think about the problem, organize the information contained in the problem in some meaningful way, and choose a strategy for solving the problem. Therefore, we no longer view the problem in its original state. We have reformulated the problem into a more understandable form. We repeat this process until our problem is in the desired goal state.

Mayer and Hegarty (1996) also connected problem solving to thinking. In fact, they defined problem solving as thinking:

> Problem solving (or thinking) refers to the cognitive processes enabling a problem solver to move from a state of not knowing how to solve a problem to a state of knowing how to solve it. (p. 31)

Mayer and Hegarty differentiated between two major kinds of cognitive processes, representation and solution. Representation occurs when a problem solver seeks to understand the problem, and solution occurs when a problem solver actually carries out actions needed to solve the problem. Mayer (1992, 1994) proffered four main component processes for mathematical problem solving: translating, integrating, planning, and executing. The first two main component processes, translating and integrating, are involved in problem representation. Translating involves constructing mental representations of the statements in the problem. Integrating involves constructing a mental representation of the situation described in the problem. A natural product of problem representation is planning. The main component process involved in the problem solution is executing or carrying out the plan (Mayer and Hegarty, 1996, pp. 33-34). Taken as a whole, the processes of representation and solution are similar to Duncker's (1945) "reformulations" of a problem. They are also reflected in Polya's (1945) approach to problem solving.

Polya's (1945) book, *How to Solve It*, is possibly the most famous discussion of problem solving. Poyla suggested a field of study for discovering how to solve problems. His book focused on four areas: understanding the problem, devising a plan for the solution, carrying out the plan, and looking back. These four areas or stages were not intended to be followed in a rigid, linear pattern, but in more of a "back-and-forth" cycle. Wilson, Fernandez, and Hadaway (1993) developed the diagram in Figure 1 to illustrate this "dynamic, cyclic interpretation of Polya's stages" (p. 61).

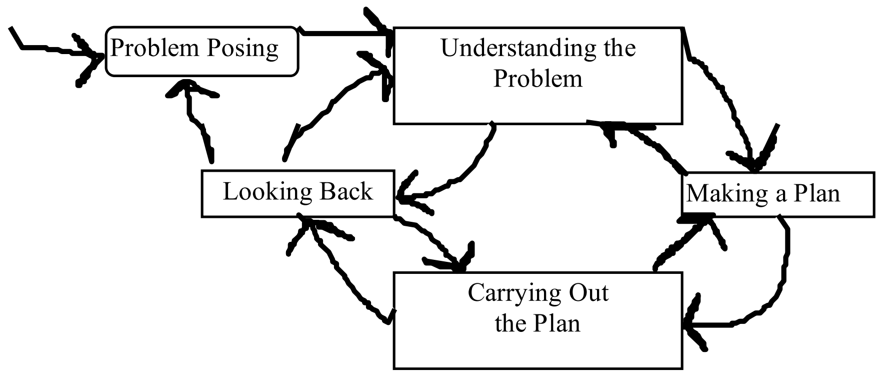


Figure 1: Interpretation of Polya's stages (Wilson, Fernandez, and Hadaway, 1993)

Polya's stage of *understanding the problem* relates to Mayer and Hegarty's (1996) *representation* phase. His *carrying out the plan* stage connects to their *solution* phase. Both, Polya (1945) and Mayer and Hegarty (1996) include a *planning* stage. However, unlike Mayer and Hegarty, Polya also emphasizes a *looking back* stage.

Polya (1945) also encouraged teachers and students to think and ask appropriate questions at each phase of the problem solving process. For instance, while devising a plan for a given problem situation a student may ask: *Do I know a related problem?* By presenting students with routine problems, Polya suggested that teachers are giving students an immediate and decisive answer to this question (p. 171).

Charles, Lester, and O'Daffer (1987) and Lester and Kroll (1990) incorporated Polya's (1945) ideas about problem solving into their own definitions. Charles et al. (1987) presented the following description of problem solving:

> Problem solving is an extremely complex activity. It involves the recall of facts, the use of a variety of skills and procedures, the ability to evaluate one's own thinking and progress while solving problems, and many other capabilities. Furthermore, success in problem solving very much depends

> on the student's interest, motivation, and self-confidence. In short, solving problems involves the coordination of knowledge, previous experience, intuition, attitudes, beliefs, and various abilities. (p. 7)

Lester and Kroll (1990) agreed with and added to this description. They suggested that problem solving involves the "process of coordinating previous experiences, knowledge, and intuition in an effort to determine an outcome of a situation for which a procedure for determining the outcome is not known" (p. 56). They also stated that the ability to solve mathematics problems must develop slowly over a long period. They concluded that problem-solving performance appears to be a function of five interdependent categories: *knowledge acquisition and utilization, control, beliefs, affects, and socio-cultural contexts*. Each of these categories can be connected to Schoenfeld's (1992) framework for explaining mathematical cognition.

Knowledge acquisition and utilization contains a wide range of formal and informal resources that assists an individual's mathematical performance. These resources include facts and definitions, algorithms, heuristics, problem schemas, and other routine procedures. It is important to realize that individuals understand, organize, represent, and ultimately utilize these resources in very different ways (Lester & Kroll, 1990, p. 56). This category of *knowledge acquisition and utilization* can be divided into two parts: Schoenfeld's categories of *the knowledge base* and *problem-solving strategies (heuristics)*.

Both, Lester and Kroll (1990) and Schoenfeld (1992) referred to the issues of knowledge content and the organization, representation, and access of that content. According to Schoenfeld (1992), the knowledge base not only consists of memory contents (knowledge), but also consists of the memory structure (how knowledge is organized, accessed, and processed). Silver (1987) described the structure of memory. Norman (1970) and Anderson (1983) provided general discussions of this topic.

I have already listed the possible contents of memory in this chapter and the previous one. I now give a brief discussion of the structure of memory. This discussion is based upon Schoenfeld (1992) and Silver (1987). Cognitive theorists have identified at least three kinds of memory registers. Figure 2, taken from Silver (1987), illustrates these registers and indicates the flow of information among them.

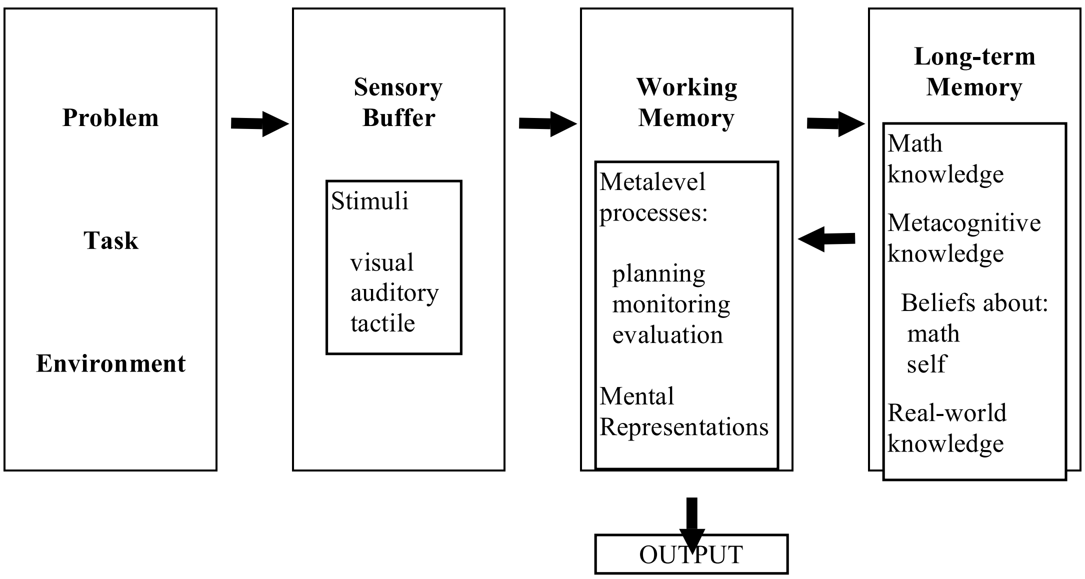


Figure 2: The structure of memory (Silver, 1987).

External information enters the structure through the problem task environment. Human beings then act as information processors. One's experiences (visual, auditory, and tactile) are registered in sensory buffers and then converted into forms in which they are used in working and long-term memory. Because much of its content is in the form of images, the sensory buffer is also called iconic memory. The sensory buffer can register a large amount of information, but it can hold it only briefly. Some of the information will be lost, and some will be transmitted to working memory.

The working memory, also known as short-term memory, receives its contents from the sensory buffer and long-term memory. This short-term memory has limited capacity. Miller (1956) indicated that individuals can only keep and operate on seven "chunks" of information in this working memory.

Long-term memory is an individual's permanent knowledge storehouse. This storehouse contains different types of knowledge. *Declarative* and *procedural* (Anderson, 1976) are two types of knowledge mentioned often in the literature. Greeno (1973) used the terms *propositional* and *algorithmic*. Ryle (1949) characterized these types as "knowing that" and "knowing how." Hiebert (1985) edited a book that explored the connections between these two types of knowledge.

In conclusion, the organization and access of information in the memory structure is dependent upon an individual's abilities to abstract and classify his or her experiences. These classifications shape what the individual sees and how he or she behaves when encountering new situations related to the ones that have been abstracted and classified.

Control refers to an individual's decisions about planning, evaluating, monitoring, and regulating. These processes of monitoring and regulating an individual's behavior are components of metacognition (Lester & Kroll, 1990, p. 57). Flavell (1979) indicated that metacognition refers to an individual's knowledge of the cognitive processes and products of the individual and others. Silver (1987) also implied that it refers to the individual's "self-monitoring, regulation, and evaluation of cognitive activity" (p. 49). Brown (1987) provided a broad historical review of this concept, and Schoenfeld (1985a, 1985b, 1987, 1989, 1992) has written much about the importance of the topic in problem solving. This category corresponds to Schoenfeld's category of *monitoring and control.*

Beliefs refer to the individual's world view. They consist of the "individual's subjective knowledge about self, mathematics, the environment, and the topics dealt with in particular mathematical tasks" (Lester & Kroll, 1990, p. 57). *Affects* refer to an individual's feelings, attitudes, and emotions. These affective factors may dominate students' thinking and actions in solving problems (p. 57). McLeod (1992) suggested that students will develop both positive and negative emotions as they struggle to learn mathematics. He also suggested that positive and negative attitudes toward mathematics will be developed when students are faced with the same or similar situations repeatedly. Schoenfeld combined these two categories into one category, *beliefs and affects*.

Finally, the social and cultural environments of individuals greatly influence their development, understanding, and use of mathematical ideas and techniques. Hence, these socio-cultural factors play an important role in an individual's success in mathematics (Lester & Kroll, 1990, p. 58). Lester and Kroll's *socio-cultural* category relates to Schoenfeld's *mathematical practices* category.

Schoenfeld has written much about problem solving. He suggested that problem solving is a "personal experience" (Schoenfeld, 1983, p. 63). Problem solvers must actively seek paths to solutions of problems. He also suggested that problem solving includes considering different approaches to solving a problem, thinking independently, and using the knowledge at our disposal effectively. Problem solving consists of false starts, reversals, and blind alleys, as well as successful steps to an appropriate solution. It means knowing when to explore, making choices about paths to pursue, and pursuing those paths to determine if they lead to a desired solution. It also means examining a solution to determine if it is appropriate. (Schoenfeld, 1983, 1985a, 1985b, 1987, 1989, 1992)

The National Council of Teachers of Mathematics has devoted much time and energy to helping students and teachers realize the importance of problem solving and mathematical thinking in curriculums and classrooms of today. Certain views of the Council in the *Standards* (1989, 1998) documents are similar to Schoenfeld's views of problem solving and mathematical thinking.

Problem solving was a major theme of the National Council of Teachers of Mathematics *Curriculum and Evaluation Standards for School Mathematics* (1989). The first standard was "mathematics as problem solving." This document described problem solving as a process that can provide the context in which concepts and skills can be learned, a process by which students experience the power and usefulness of mathematics, a method of inquiry and application, and a process by which the "fabric of mathematics is both constructed and reinforced" (pp. 23, 25, 137).

The draft of *NCTM's Principles and Standards for School Mathematics* (1998) continued the theme of problem solving. The sixth standard of this document is stated below:

> **Standard 6: Problem Solving**
> Mathematics instructional programs should focus on solving problems as part of understanding mathematics so that all students –
> - build new mathematical knowledge through their work with problems;
> - develop a disposition to formulate, represent, abstract, and generalize in situations within and outside mathematics;
> - apply a wide variety of strategies to solve problems and adapt the strategies to new situations;
> - monitor and reflect on their mathematical thinking in solving problems. (p. 76)

The current draft suggested that as students use problem solving approaches, they "develop new mathematical understandings and strengthen their abilities to use the mathematics they know" (p. 76). These understandings and abilities fall into Schoenfeld's category of *the knowledge base* and Lester and Kroll's (1990) category of *knowledge acquisition and utilization*.

The *Principles and Standards* (1998) draft's suggestion that students develop a mathematical disposition can be linked to Schoenfeld's *beliefs and affects* category. Schoenfeld remarked that beliefs and affects help determine an individual's mathematical point of view. According to the draft, when individuals develop a mathematical disposition, they tend to act in mathematically productive ways. They analyze and explore situations to see what "makes things tick" mathematically. They abstract and generalize these situations and possibly develop new connections among mathematical ideas.

The *Principles and Standards* (1998) draft also suggested that problem solving strategies should be an important part of a student's "mathematical tool kit" (p. 78). Students should have adequate instruction and practice in using these strategies. Some of the strategies mentioned in the draft include using diagrams and other representations, looking for patterns, listing all possibilities, trying special cases, working backward, guessing and checking, creating an equivalent problem, and creating a simpler problem (p. 78).

A number of studies have been conducted on the use of problem-solving strategies (or heuristics) since Polya (1945) published *How to Solve It* (see for instance Kantowski, 1977; Kilpatrick, 1967; Lucas, 1974; Smith, 1973; Wilson, 1967). There have also been numerous books and articles written about heuristics and problem solving (see for instance Krulik, 1980; Schoenfeld, 1985a). Recently, Posamentier and Krulik (1998) completed a resource book for mathematics teachers, *Problem-Solving Strategies for Efficient and Elegant Solutions*. Ten popular strategies are described in the book, including applications to everyday problem situations and mathematics. Of these strategies, *intelligent guessing and testing* (pp. 165-186), was particularly relevant to this study. This strategy is often referred to as the trial-and-error method by some students and teachers. However, this terminology is an oversimplification because the strategy is quite involved. In using the *intelligent guessing and testing* strategy, a person makes a guess and then tests it against the conditions of the problem. Another (intelligent) guess is made based upon information from testing the previous guess. This process continues until a satisfactory solution is reached. The *intelligent guessing and testing* strategy is useful when it is necessary to restrict the values for a variable to make the solution more manageable. It is also helpful when the general case is far more complicated than a specific case, with which you can narrow down the options in an effort to focus on the correct answer (p. 165).

Like Schoenfeld's (1992) framework, the *Principles and Standards* (1998) draft also suggested that students should monitor and reflect on their mathematical thinking in solving problems (p. 79). Effective problem solvers constantly monitor and adjust what they are doing. They plan frequently and consider alternatives when not making progress in solving problems.

<u>Mathematical Modeling</u>

Rubenstein (1975) defined a model as "an abstract description of the real world; ... a simple representation of more complex forms, processes, and functions of physical phenomena or ideas" (p. 192). He stated the purpose of a model is to "facilitate understanding of relationships between elements, forms, processes, and functions, and to enhance the capacity to predict outcomes in the natural and man-made world" (p. 238). Rubenstein asserted that models evolve and change as new understanding is gained. He also admitted that models may fail to contain some elements of the real world (errors of omission) or contain elements not present in the real world (errors of commission).

The *Principles and Standards* (1998) draft provided similar definitions for a mathematical model and modeling. According to this document, a mathematical model is a "mathematical representation of the elements and relationships within an idealized version of a complex phenomenon" that can be used to "clarify understandings of the phenomenon and to solve problems" (p. 98). It further stated that the act of mathematical modeling includes not only representation, but the "acting upon the representation and interpreting the meanings of one's actions within the mathematical model and with respect to the phenomenon being modeled" (p. 99).

The *Principles and Standards* (1998) draft also stated that "the act of modeling is a complex enterprise" and provided an "oversimplified diagram representing the process" (pp. 99-100). This diagram, shown in Figure 3, suggests the "path by which conclusions are drawn about the situation being modeled" (p. 100). In the diagram, the modeling process is represented from the bottom left to the top left. Mathematical analyses are represented across the top of the diagram, and the interpretation of the findings is represented from the top right to the bottom right.

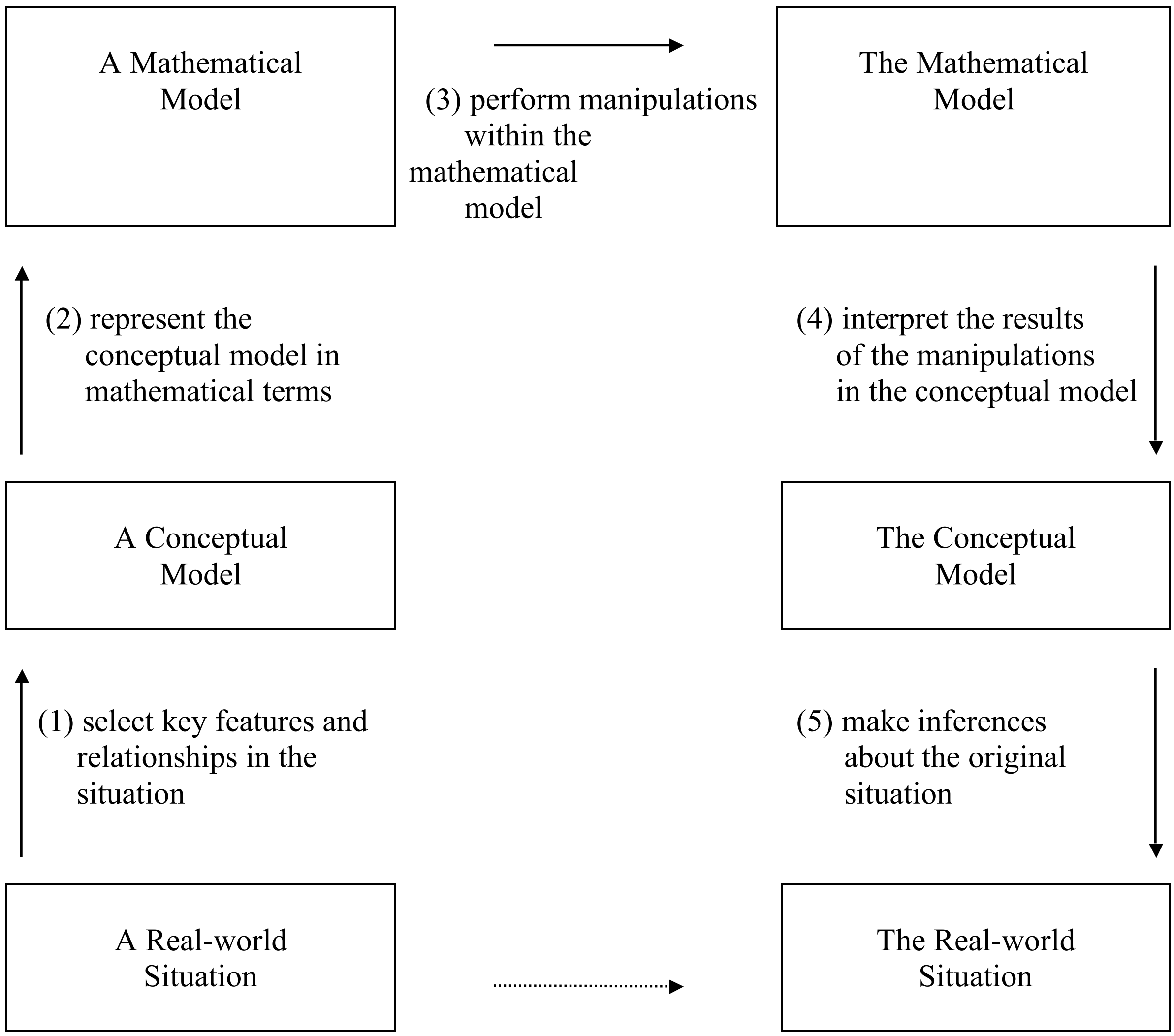


Figure 3: Aspects of the modeling process (NCTM, 1998).

During this process certain questions should be considered. For instance, does a student's mathematical model capture the appropriate relationships among the features of the situation and does the interpretation of the model make sense when mapped back to the original situation? These, as well as many other questions, must be kept in mind when students are engaged in the mathematical modeling process (NCTM, 1998, p. 100).

Representations are of great importance in modeling and mathematics. The use of the spreadsheet template to model population data in the mathematical modeling course provided the students with multiple representations of the problem situation: table of values, equation, and graph. According to Dugdale et al. (1995), such use of computers in the classroom places the curriculum focus on "reasoning with a variety of representations and understanding the relationships among the representations" (p. 330). This opinion seems to agree with the *Principles and Standards* (1998) draft's view of representations as vehicles for mathematical thought.

Recent Studies on UGA's Modeling Course

Searcy (1997) conducted an intrinsic case study on mathematical thinking in a pilot section of the University of Georgia's mathematical modeling course. The pilot section was taught by course creator, Henry Edwards, during Winter Quarter 1997 and was described as an applied college algebra course. Searcy used Schoenfeld's (1992) framework to examine the complexity of a single student's mathematical thinking in this class. She used informal interviews, classroom observations, student work, and an exit interview as methods of data collection.

Searcy found that a mixture of facts, algorithmic procedures, and informal knowledge dominated the student's knowledge base. The student used "finding a related problem" strategy to solve routine problems and basic strategies that reflected Polya's sense of heuristic reasoning to work non-routine problems. The student's approach to routine problems was little more than "checking the answer." For non-routine problems she planned, tested, and abandoned non-productive strategies. As for beliefs, the student seemed to believe in two types of mathematics: classroom mathematics and everyday life mathematics. Her beliefs about the real world strongly influenced her attempts to make sense of situations that she encountered in the course. The student's practices were considerably different from those advocated by the mathematics community. She seemed to have few social encounters to help shape her mathematical thinking. She relied on other resources, like intuitive knowledge and personal theories, for her interpretation and sense-making of mathematics.

Lanier (1997) conducted a pilot study on a different section of UGA's applied college algebra course during Winter Quarter 1997. The purpose of this case study was to identify the impact of the applied college algebra course on a student's perceptions of her ability to succeed in algebra. The use of technology seemed to have a favorable impact on the student's perceptions and attitudes. The student believed the technology alleviated the "burden" of memorizing and allowed more time for understanding concepts. Therefore, she felt she could be successful in the course. The use of multiple assessment methods (quizzes, written tests, computer assignments and projects) also had a positive effect on her perceptions of success.

Mathematical Modeling with Spreadsheets

I was unable to find another course that uses a spreadsheet template such as the one in this study to introduce modeling. However, there are schools across the country that have their own mathematical modeling courses with similar goals to UGA's mathematical modeling course. These courses also encourage the use of calculators and spreadsheets to model real world situations.

In Georgia, Clayton College and State University (CCSU) has its own version of *Math 1101: Mathematical Modeling* (CCSU, 1999). This course in applied college algebra uses graphical, numerical, symbolic, and verbal techniques to describe and explore real world data and phenomena. Investigation and analysis of applied problems are supported by appropriate technology, such as graphing calculators and spreadsheets. Like UGA's mathematical modeling course, CCSU's course is intended for non-mathematics intensive majors.

Salisbury State University (SSU) offers a course, *Math 165: Introduction to Mathematical Modeling* (SSU, 1999), for students interested in developing their ability to solve problems in mathematics or science and prospective middle school teachers in mathematics or science. This course was developed under the auspices of the *Maryland Collaborative for Teacher Preparation*. The objectives of the course are similar to those of UGA's mathematical modeling course. They include helping students

- make connections between mathematics and other disciplines,
- present and analyze real world phenomena using a variety of mathematical representations,
- develop strategies and techniques for applying mathematics to solve problems, and
- explain and justify their reasoning using appropriate terminology in both oral and written form.

Students enrolled in SSU's mathematical modeling course use calculators, computers, and microcomputer-base laboratories to collect their own data, generate graphs, and analyze their results. Students work together in groups. Assessment is based on exercises, reports, electronic journals, a portfolio, and examinations.

College algebra is not the only freshman mathematics course to be inundated by modeling. There is also a reform movement within business calculus courses that calls for a modeling approach with spreadsheet software. One such course is the *Villanova Project* (Pollack-Johnson & Borchardt, 1999). This project is an extension of one developed at Clemson University. The course is intended as a first year mathematics course for business and social science. It uses a graphing calculator for calculus and spreadsheet software for multiple regression, matrices, and linear programming. As with SSU's mathematical modeling course, Villanova's course has similar objectives to UGA's modeling course. The objectives of this course include, but are not limited to,

- emphasizing mathematical modeling and functions as they relate to the real world,
- solving realistic, interesting, and practical problems that use real world data,

- emphasizing understanding and applications of concepts, and
- using technology as a tool.

Materials from the Villanova Project have been adopted at other schools throughout the nation (see for instance Smith, 1999). Thus, many colleges and universities have recognized the need for and have attempted to provide courses that

- connect the mathematics used in the classroom with the real world,
- emphasize mathematical reasoning and communication (written and verbal) skills, and
- encourage appropriate use of technological tools.

Summary

In this chapter I presented literature related to the main areas of this study: problem solving, mathematical thinking, and mathematical modeling with technology. I attempted to describe the foundation or basis for Schoenfeld's (1992) framework that was used to analyze the data for this study. Also, I attempted to provide some insight into the shift in curriculum for introductory college mathematics courses. I did not provide much literature on spreadsheet templates. The use of spreadsheet templates in introductory college mathematics courses has yet to be well documented in the literature. Nevertheless, I hope the literature cited in this chapter provided a better understanding of this study.

CHAPTER THREE

METHODOLOGY

The purpose of this study was to investigate college students' understanding of linear modeling when using a spreadsheet template to model data in a mathematical modeling course. To conduct this investigation, I needed to focus on the students' thoughts, actions, and interpretations as they solved population problems using the linear modeling approach presented in the course. Thus, the nature of the research situation demanded that I use a qualitative case study approach such as that described by Merriam and Simpson (1995). The study was conducted during fall semester of 1998. The case study approach not only allowed me to focus on the modeling situation, but also provided the reader with a rich description of the situation. This chapter describes the participants, the methods of data collection and analysis, and the limitations of the study.

The Participants

I used a criterion based strategy combined with convenience sampling (Patton, 1990) to select participants for this study. The criteria for selection was that a chosen student must be enrolled in the mathematical modeling course taught at the University of Georgia, must be able to express his or her thoughts and actions in both oral and written form, must be an above average or excellent student in most courses, must be willing to participate, and must have the time to participate. The purpose of the study necessitated my choosing students that were enrolled in the mathematical modeling course as participants. The other criteria were used to provide the richest possible data and to ensure the completion of the study.

Participants were chosen from students at the University of Georgia enrolled in MATH 1101 (Mathematical Modeling) during fall semester of 1998. These students were not majoring in mathematics or a related field requiring extensive mathematics courses. Dr. Kirk, a mathematics professor at UGA, taught two sections of this course during fall semester, a 10 - 11 A. M. class meeting three days a week (Monday, Wednesday, and Friday) and a 5 - 6:15 P. M. class meeting two days a week (Tuesday and Thursday). The first homework assignment for the students was to send Dr. Kirk an introductory e-mail about themselves. Dr. Kirk allowed me to read these introductory e-mail assignments. As I read the e-mails, I looked for students who elaborated on their mathematical backgrounds and expressed their thoughts about mathematics. The most informative e-mails were from students in the Tuesday/Thursday class. I contacted twenty-five of these students by e-mail. In my e-mail to them, I briefly explained my study and asked if they would be interested in finding out more about the study and possibly becoming a participant. Eight of the twenty-five responded that they were interested in learning more. I talked with these students individually through e-mail, on the telephone, and in person. All of the eight were acceptable participants for this study. Eventually, the determining factor was time. Three of the eight students met the time criteria and agreed to be participants for my study. They were Adam, Cindy, and Kaitlyn.

Adam. Adam was a nontraditional college freshman. He was an above average student and attended the University of Georgia part-time. He graduated from high school in June 1987 and served in the Air Force for three years. He was married with two children and worked full-time as a firefighter. Adam's major was child and family development. His goal for the future was to seek a job in social services. In high school, Adam took Pre-algebra, Algebra I, Geometry, Consumer Mathematics, and Technical Mathematics. The mathematical modeling course was his second mathematics course at UGA. Adam took a mathematics course from the Department of Academic Assistance last year. This course was designed to prepare students for college algebra. Adam sat in the second row (right side) of the theater style classroom.

Cindy. Cindy was a traditional college freshman. She was an excellent student and attended the University of Georgia full-time. She graduated from high school in June 1998. Cindy's major was sports science. However, she had not chosen a career. She was considering changing her major. In high school, she took honors mathematics courses. These courses included Geometry, Algebra II, Advanced Algebra, Trigonometry, and Pre-calculus. The mathematical modeling course was her first mathematics course at UGA. Cindy sat in the first row (left side) of the theater style classroom.

Kaitlyn. Kaitlyn was a traditional college freshman. She was an above average student and attended the University of Georgia full-time. She graduated from high school in June 1998. Kaitlyn's major was early childhood education. Her goal for the future was to teach second grade. In high school, she took Algebra I, Geometry, Algebra II, and Trigonometry. The mathematical modeling course was her first mathematics course at UGA. Kaitlyn sat in the fourth row (middle) of the theater style classroom.

All three participants were exposed to the usual algebra topics discussed in many high school mathematics classrooms. These topics included finding slopes of lines, writing equations of lines, and graphing lines. However, Adam and Kaitlyn did not remember studying linear modeling before enrolling in this course. Cindy thought she had studied linear modeling in high school, but could not demonstrate or explain anything about the process.

Three was an appropriate number of participants for this study. Adam, Cindy, and Kaitlyn provided detailed and rich data. In addition, having a nontraditional college student (Adam) as well as two traditional students (Cindy and Kaitlyn) as participants allowed me to compare and contrast the understandings of the two different populations of college students.

<u>Methods of Data Collection</u>

Observation, interview, and document analysis were the methods of data collection
used in this study. The data from these three methods supported and enhanced each other and provided reliable answers to the research questions.

<u>Observation.</u> I conducted two types of observations for this study. First, I observed the majority (over 70%) of the class sessions. I wanted to determine the procedures and information the students were exposed to in class. I sat in the back of the classroom and took notes as the students did. I also made note of the students' behavior and reactions to the lessons during this time.

The second type of observation was made outside of class. I observed each participant as he or she used the linear modeling spreadsheet template to find the optimal model for a population data problem and a data problem that they were not exposed to in class. I encouraged the students to "talk aloud" as they worked on the problems. I took notes as well as audiotaped these observations.

<u>Interview.</u> I conducted informal interviews with the participants throughout the semester to get to know the participants and to monitor their progress in the course. These informal interviews were in the form of e-mails, phone conversations, and brief conversations before and after class. Also, I conducted a semi-structured interview [Appendix B] with each participant after the completion of the second part (the linear modeling part) of the course project [Appendix C]. These interviews occurred in my office or outside the graduate studies building. I asked each participant a set of questions pertaining to linear modeling and the processes they performed for the project. This set of questions formed an interview guide, was the same for each participant, and provided the same type of data from each participant. In addition, I asked the participants follow-up and probing questions based upon their answers to the interview guide questions. The questions focused on what the students did while solving the problems, how they did it, why they did it, and what conclusions they made. These types of questions helped me answer the research questions. Finally, I conducted an exit interview [Appendix D] with each participant at the end of the semester. All of these interviews were held in my office. During this interview, I asked the participants to use their spreadsheet template to find the optimal linear model for a familiar data set (population data) and a non-familiar set (one not experienced in class) and asked them to interpret these situations. I also asked them to summarize their linear modeling experience. Finally, I shared my data from observations, written work, and previous interviews with the participants in this interview. This provided verification of my perceptions of the students' thoughts and behavior during the study.

Document Analysis. I examined two pieces of the participants' written work. Each participant completed a required three-part project for the course. The second part of the project required students to use a spreadsheet template [Appendix A] to determine the optimal linear model for a set of population data and to interpret the findings. This template was provided in the course package given to the students at the beginning of the semester. Each participant also completed a test covering linear modeling, quadratic equations, and higher order equations. I collected the second part of the project [Appendix C] and the third test from each participant.

Time Line. This study was conducted during the fall semester of 1998. The following time line displays major course and research events.

Date	Events
Tuesday, September 15	Reviewed introductory e-mail messages from the students. Selected twenty-five students from e-mails. Sent initial e-mail to students asking about interest in study. Began observing class sessions. Review of material for Test 1.

	Course content for Test 1: Percentage increase and decrease problems, interest and iteration problems, tabulation with calculator, graphing with calculator. Cindy and Kaitlyn responded to initial e-mail.
Wednesday, September 16 - end of semester	Corresponded with Cindy and Kaitlyn through e-mails, phone conversations, and conversations outside class.
Thursday, September 17	Test 1
Tuesday, September 22	Students began section about natural growth of populations. Introduced to spreadsheet template (natural growth model) for part 1 of course project. Adam expressed interest in study.
Tuesday, September 22 - end of semester	Corresponded with Adam through e-mails and conversations outside class sessions.
Friday, October 2	Part 1 of Course Project (natural growth model) due.
Tuesday, October 6	Review of material for Test 2. Course content for Test 2: Natural growth of populations, growth and decline in the world.
Thursday, October 8	Test 2
Tuesday, October 13	Students began section about straight lines and linear growth model. Introduced to spreadsheet template (linear growth model) for part 2 of course project.
Tuesday, October 20	Students ended linear modeling section and began discussing quadratic models.
Friday, October 23	Part 2 of Course Project (linear growth model) due.
Thursday, October 29	Semi-structured interview with Cindy in my office (9:30 A.M., 45 minutes in duration). Semi-structured interview with Kaitlyn in my office (1:00 P.M., 60 minutes in duration).
Tuesday, November 3	Semi-structured interview with Adam outside graduate studies building (4:00 P.M., 60 minutes in duration). Review of material for Test 3.

	Course content for Test 3: Straight lines and linear growth model, quadratic model and equations, higher degree polynomial models.
Thursday, November 5	Test 3
Tuesday, November 10	Discussion in class about final part of course project. Students introduced to spreadsheet template (bounded growth model) for this part of the project.
Thursday, November 12	Students began section on maximum and minimum problems.
Friday, November 20	Final Course Project (bounded growth model) due.
Tuesday, December 1	Review of material for Test 4. Course content for Test 4: Maximum and minimum problems. Last class observation. Exit interview with Adam in my office (6:30 P.M., 90 minutes in duration).
Thursday, December 3	Test 4
Tuesday, December 8	Exit interview with Kaitlyn in my office (9:30 A.M., 60 minutes in duration).
Wednesday, December 9	Exit interview with Cindy in my office (1:00 P.M., 60 minutes in duration).

Method of Data Analysis

The data generated by observations, interviews, and written documents were analyzed using the constant comparative method. I transcribed and coded the data for themes and categories. Corbin and Strauss (1990) described the three basic types of coding that was used for analysis: open, axial, and selective.

I began the analysis by openly coding each piece of data as it was collected. This process involved assigning labels to blocks of data that best described that data. These labels included

- test and project procedures;
- definition and/or meaning of slope, y-intercept, census value, predicted value,

average error, sum of squared errors, and best fit;

- purpose of linear modeling;
- ideas about population and the real world;
- misconceptions about ideas and concepts;
- systematic trial-and-error and looking for a related problem strategies;
- questioned, corrected, verified computations and solutions;
- treated spreadsheet as black box;
- limited choices for slope and y-intercept;
- relied on procedures and intuitions to select the better model between linear and natural growth models; and
- views about mathematics, computers, and spreadsheet template.

As new data was collected, the codes from this data were compared to those of previous data. This was done to determine similarities and differences in the data and to allow for refining and updating the codes.

The next phase of the analysis process (axial coding) was to place the labels from open coding into categories. These categories corresponded to Schoenfeld's (1992) framework for exploring mathematical cognition. Each of the labels was placed into one of the categories. When a label seemed to fit into more than one category, I discussed this issue with colleagues and used my interpretation of Schoenfeld's framework to place the label into what I thought was the most appropriate category.

Finally, the codes were further refined and unified around the core category: mathematical understanding of linear modeling. This process is known as selective coding. The final coding scheme arrived at by this analysis is given below.

<u>Mathematical Understanding of Linear Modeling</u>

1. Knowledge Base
   1.1. Algorithmic and Routine Procedures
      1.1.1. Test Problem Procedure (graphing calculator)

1.1.2. Project Problem Procedure (spreadsheet template)

1.2. Definitions, Facts, and Meanings

1.2.1. Slope(m)
1.2.2. Y-intercept (b)
1.2.3. Census value
1.2.4. Predicted Value (P(t))
1.2.5. Average Error
1.2.6. Sum of the Squared Errors (SSE)
1.2.7. Best Fit

1.3. Informal and Intuitive Knowledge

1.3.1. Purpose of Linear Modeling
1.3.2. Population and the Real World

1.4. Misconceptions

1.4.1. Relationship between average error, census values, and predicted values
1.4.2. Number of data points through which a line of best fit must pass
1.4.3. Idea of a line “barely touching” a point
1.4.4. Comparison of average error for linear model to average error for natural growth model to determine better model

2. Problem Solving Strategies

2.1. Systematic Trial and Error (Intelligent Guess and Test)
2.2. Look for a Related Problem

3. Monitoring and Control

3.1. Realized errors in computations
3.2. Corrected errors in computations
3.3. Questioned solutions
3.4. Verified Solutions

4. Practices

4.1. Treated spreadsheet template as black box
4.2. Limited types of numbers for slope (*m*) and y-intercept (*b*)
4.3. Used informal and intuitive knowledge of real world to select better model (linear or natural growth)
4.4. Used algorithmic procedure (average error) to select better model
4.5. Used informal and intuitive knowledge of real world to decide how well the model fit the problem

5. Beliefs and Affects

5.1. Views about Mathematics

5.1.1. Cause of Anxiety, Apprehension, and Tension
5.1.2. Useless and Stupid
5.1.3. Pretty Easy
5.1.4. Something to Dislike or Tolerate

5.2. Views about Computers
   5.2.1. Cause of Anxiety, Apprehension, and Tension
   5.2.2. Easy to Use
5.3. Views about Spreadsheet Template
   5.3.1. Easy to Use
   5.3.2. Does most of the Work

Limitations

There were limitations in conducting this case study. These limitations concerned the instructions given for finding a linear model, the students' experiences with spreadsheets, my narrow focus on the material presented in the mathematical modeling course, usual limitations of data collection, and subjectivity.

Instructions for Linear Modeling. The instructions given to the students for finding a linear model may have been a limitation to this study. The students were provided with a spreadsheet template for linear modeling [Appendix A]. Dr. Kirk presented this template as a "black box" (Personal Communication, October 6, 1998). He instructed the students on how to enter their data, find initial values for *b* and *m* (the slope and y-intercept), and determine the optimal linear model. The "rule of the game" was to make the average error as low as possible by manipulating *b* and *m*. Part two of the course project [Appendix C] also included these instructions. Dr. Kirk did not elaborate on the inner workings of the spreadsheet template. He also chose to skip a section of the course material that introduced spreadsheets. Thus, the students had an already established procedure for finding a linear model. The students may have developed a different understanding of linear modeling if they were required to understand the various parts of the spreadsheet template (such as the sum of the squared errors and average error) or were asked to create their own spreadsheet.

Lack of Experience with Spreadsheets. Another limitation was that the three participants had never used a spreadsheet. This lack of experience really made the spreadsheet template appear to be a "black box." A person with experience using and creating spreadsheets may have understood the various parts of the template without formal instruction. However, all three participants found the template easy to use.

<u>Narrow Focus on Course Material.</u> My decision to focus on linear modeling was a limitation of this study. The mathematical modeling course included the natural growth, the linear, and the bounded growth models. By limiting my focus to the linear model, I did not allow for the students' comparisons of the three models. However, the participants did compare the linear model with the natural growth model. This was because the first part of the course project dealt with the natural growth model and the second part of the course project was an extension of the first.

<u>Limitations of Data Collection.</u> Collecting data through observations depended on my ability to observe and record important events properly. The record of the observations was my interpretation of what was happening, not the student's or anyone else's interpretation. I attempted to remain objective throughout the observations, and not focus on the activities I hoped to observe (an act that would have influenced the type of data recorded). During observations I focused on external behaviors (such as the what the participants wrote or what actions they performed on the computer). I could not see what was happening in the participants' minds. I attempted to overcome this limitation by asking the students to verbalize their thoughts while attempting to solve a problem and by asking them questions about my observations.

The course projects did not "show" all the students' thoughts and work that occurred during the problem solving process and the preparation of the report. I attempted to clarify their thoughts and actions in the interviews.

<u>Subjectivity.</u> Past experiences may have influenced this study. I observed a pilot section of this mathematical modeling course during the winter quarter of 1996. Also, I have taught algebra to college students for fifteen years. This teaching experience has certainly influenced how I perceive other classrooms and courses. It has also influenced how I relate to people, especially students. I was concerned that my being a teacher and my previous observations of this course could negatively affect the way I reported class observations and conducted interviews. I constantly reminded myself that I was the researcher, not the teacher. As a researcher, I strived to have minimal interference with the classroom, the students, and their problem solving processes. Of course, my presence in class and during interviews made it impossible to be completely separate from the students and the experience. Nevertheless, I continuously monitored and addressed my actions and concerns so that I did not inadvertently jeopardize the validity of the data and the results of my study. I also discussed my concerns with several of my colleagues.

CHAPTER FOUR

RESULTS

This study investigated college students' understanding of linear modeling when using a spreadsheet template to model data in a mathematical modeling course. I collected data using interviews, observations, and document analysis during the fall semester of 1998. I used the constant comparative method as described by Corbin and Strauss (1990) to analyze this data. This chapter reports my findings.

While conducting this research, I sought to answer the following question:

How are the students'

- **knowledge base,**
- **problem-solving strategies,**
- **monitoring and control processes,**
- **beliefs and affects, and**
- **practices**

manifested in their learning of linear modeling?

As I became heavily involved in the analysis process, several of Schoenfeld's (1992) categories of mathematical thinking emerged from the data. In particular, evidence of the students' knowledge base concerning linear modeling, the strategies used to find the optimal linear model, and the mathematical practices of the students emerged from the data. There was also some evidence of the students' self-regulation processes. I did not focus on the beliefs of the students in my research. However, the students' feelings about mathematics and spreadsheets did manifest themselves in some of our conversations. Thus, Schoenfeld's categories helped me answer the research question.

Knowledge Base

As mentioned in Chapter One, the knowledge base refers to the "mathematical tools an individual has at his or her disposal" (Schoenfeld, 1992, p. 349). These mathematical tools include algorithmic and routine procedures, meanings and definitions, informal and intuitive knowledge about the mathematical situation, relevant competencies, and even misconceptions.

Students in the mathematical modeling course used two types of technology and, hence, two procedures for linear modeling. In class and on homework problems, students used a graphing calculator with pencil and paper to solve linear modeling problems. The following problem from Test 3, Form C is an example of such a problem.

> The population of Math City was 88 thousand on January 1, 1992, and was 108 thousand on July 1, 1997.
>
> (a) Assuming the same rate of increase continues, write the linear population model P(t) = b + mt giving the population of Math City (in thousands) t years after January 1, 1992.
>
> (b) How many years – after 1/1/1992 – will it be until the population of Math City is 140,000 people?
>
> (c) Find the month and the calendar year when the population of Math City is 140,000.

There were three forms of Test 3 with only the numbers being changed on each form. The students practiced this type of problem in their homework assignment [Appendix E] for the linear modeling part of the course. The procedure for answering this type of problem was:

- Find the slope, *m*, by
- Let the y-intercept, *b*, equal the population of the first year given.
- Write the linear population model as $P(t) = b + mt$.
- Let $b + mt$ equal the desired population in part (b) of the problem.

- Solve this equation for *t* with paper and pencil or with a graphing calculator. The usual procedure for the calculator was to find the point of intersection of the two graphs *Y1 = b + mt* and *Y2 = the desired population*. The first coordinate of this point was the answer for *t*.
- Add the whole number part of *t* to the population of the first given year.
- Multiply the decimal part of the answer for *t* by 12 to determine the appropriate month.

This procedure was demonstrated many times in class sessions (Class Observation, October 13, 15, 20, November 3).

Adam, Kaitlyn, and Cindy were able to perform the procedure for finding the linear population model in part (a) of the problem. They also found the correct number of years for part (b). There were slight differences in the way the three participants expressed their linear equations. Adam and Kaitlyn expressed their slopes in decimal form. Cindy, however, expressed the slope as a fraction. In fact, throughout the semester, she did many calculations by hand and expressed her answers in fraction form. I questioned her about this in the exit interview on Wednesday, December 9, 1998. She stated that she felt more comfortable working with fractions. She had been encouraged to write answers in fraction form in high school. Cindy and Kaitlyn correctly answered part (c) of the problem. Adam was able to find the correct year but was one month away from the correct month.

In part two of the course project [Appendix C], the students used a spreadsheet template to find the optimal linear model for a chosen city's population data. Dr. Kirk presented the spreadsheet template as a "black box" (Personal Communication, October 6, 1998). The spreadsheet depicted the squared errors and the sum of the squared errors (SSE). However, Dr. Kirk chose not to go into detail about these numbers or the process of creating this spreadsheet. He elected to skip section 1.5 of the course that gave students the opportunity to explore the inner workings of a spreadsheet and to create their own simple spreadsheets. His reason for skipping the section was that because of computer equipment and software updates on his machine, students would observe different procedures from those needed for the template. He was concerned the students would become confused (Personal Communication, October 6, 1998). The procedure for finding the optimal linear growth model was as follows:

- Choose a U. S. city based on the last two digits of your social security number.
- Enter the name of the city on Sheet l of the spreadsheet template.
- Enter the population of the city for the years 1960, 1970, 1980, and 1990 under the census column.
- Enter the initial value of *b* as the population for the year 1960 in the appropriate cell.
- Calculate an initial value for *m* by letting

  .
- Enter this initial value for *m* in the appropriate cell.
- Copy and paste the name of the city and the values for *b*, *m*, and the census years into Sheet 2 of the spreadsheet template.
- Change *b* and *m* to make the average error as small as possible.
- Write the optimal linear growth model as $P(t) = b + mt$ with *b* and *m* the values that give the smallest average error.

This procedure was discussed in class (Class Observation, October 13, 20) and described in the project directions. All three participants were able to satisfactorily perform the procedure, and hence, find an acceptable optimal linear model for their project. Adam, Cindy, and Kaitlyn were also able to demonstrate the procedure in their exit interviews. I asked them to find a linear model for the city of Chicago, Illinois. All three students found an appropriate slope and y-intercept for the linear model.

Slope and Y-Intercept. The variables *m* and *b* played an extensive role in the linear modeling procedure. The students had definitions for *m* and *b*. Some of these definitions included algebraic words and concepts, while others pertained to the population modeling situation. During their interviews on October 29, Cindy and Kaitlyn each referred to *m* as the slope and *b* as the y-intercept. I asked them to define slope and y-intercept. Cindy responded that the y-intercept was "the place where x is zero." Kaitlyn defined the y-intercept as the "initial population ... the first population." She was referring to the census value for the year 1960. Adam did not use the algebraic terms to describe *m* and *b*. He never referred to *b* as the y-intercept and *m* as the slope. Like Kaitlyn, Adam defined *b* as the "initial population." In an interview on November 3, he stated, "... *b*, to me, is the initial population that you are computing *m* for. In other words, I guess that *b* would be like (the population for) 1960."

When asked to define slope, Cindy stated, "The slope is like rise over run. It's like the increase or decrease in your line." She could quote and appropriately use the formula for finding slope given two points. I asked her to tell me what the slope meant for this population problem. At this point she had 2.4 in the spreadsheet cell for *m*. She responded that it meant an "increase of 2.4 thousand each year ... 2.4 thousand people for each year." Adam defined *m* as "the amount of population growth per year." He further elaborated, "... *m* is 6.7 (he had a different city), which is the population would grow by 6.7 thousand people per year, each year." Both Adam and Cindy stated that the actual population may not grow by this amount. The value for *m* was an estimate for how the population might change each year.

Kaitlyn could find the slope using the above mentioned formula and defined it as an increase or decrease in the line, but she could not tell me how it related to the population situation. She defined slope as "how the line increases and decreases, like how it goes up or down (pause) like how fast it increases or decreases" (Interview, October 29). In her exit interview on December 8, she said, "It's the rise over the run, like a triangle." However, when I asked her what a slope of 6.7 meant in terms of the population problem, she responded, "I guess that it (population) increases. It's not like a percent or anything. I don't know how it tells like how much it increases. I just know it tells that it increases" (Interview, October 29).

Although Kaitlyn was not sure of how the slope influenced the population, she did realize that the slope was not a percentage. Adam and Cindy also realized this point. In several conversations, Adam emphasized that the slope was a constant number, unlike the rate *r* that he found for the natural growth model in part one of the project (Personal Communication on October 20, Interview on November 3, Exit Interview on December 1).

Thus, all three students knew that changing the slope (*m*) made the line "look different." However, they had a difficult time expressing this difference in words. They talked about the line increasing or growing. The students were attempting to describe the "steepness" of the line. Kaitlyn remarked in her exit interview (December 8) that a different value for *m* (the slope) changes "the steepness of the line", but leaves the starting point (the y-intercept) the same.

The students were instructed to choose cities with increasing populations. Thus, their calculated slopes were positive and their linear models were increasing. I asked them what would happen if the slope was negative. All three responded that the line would decrease, that is, the population would decline. I presented them with this situation in the exit interview. The population for Chicago was decreasing in the years from 1960 to 1990. I asked them to find a linear model for this set of population data. With little hesitation, each student used the spreadsheet template to carry out their "learned" linear modeling procedure.

Cindy and Kaitlyn entered their population data rounded to the nearest thousands as they had done for their project report. Interestingly, Adam entered the population data as given in the census table instead of rounding as he had done in his project. The different appearance of the numbers did not seem to be a source of conflict for him. He proceeded with his usual spreadsheet procedure. Once again, Adam and Kaitlyn expressed their value for the slope (*m*) using decimals while Cindy wrote her slope (*m*) as a fraction. All three students assigned the population value for the year 1960 to the y-intercept (*b*). Thus, the students were able to find a linear model for the Chicago data. The students then searched for the optimal linear model by trying several values for *m* and *b* and observing the change in the average error with each trial.

It was during these exit interviews that I discovered that some of the students did not have a clear "picture" of the relationship of the slope to the graph. At one point in an interview (Exit interview, December 1, 1998), Adam suggested that changing the slope caused the entire line to move "up or down." I asked him to sketch a picture of this movement. He drew several translations of the original graph. Adam was translating the graph in his mind, and thus, changing the y-intercept as well. This misconception seemed to suggest that Adam had not given much attention to the graph when using the spreadsheet during the course. When asked about the effect of changing *b* , Adam suggested that a lower value for *b* would cause the line to "incline" and a higher value for *b* would cause the line to "decline." He clarified his definitions of "incline" and "decline" by describing the line as shifting upward or downward. I was still uncertain about his intended meanings, so I asked him to explore the situation. I encouraged him to try several values for the slope (*m*) and y-intercept (*b*) and to notice the effect on the graph. The following excerpt from his exit interview (December 1) reveals his exploration and his surprising conclusion:

A: Now my *b* or population value is 3,550,404. If I change it to 3 million which is below the 3.5 million, then I'm thinking my line will adjust. It will shift upwards.

S: Ok, let's try it.

(Adam changes the value of b to 3 million.)

A: Oh, the whole line came down. Well that blows my theory. (laughs) It's like it adjusted everything. The whole line, it dropped it completely down. Basically it's the same just dropped down.

S: What do you think would happen if you put in 4 million for *b*?

A: It would probably go above the original points. (Adam tried this.) Yep.

S: How is this different from changing *m*? What does changing *m* do to the line?

A: *M* changes the rise. It changes the angle of ascending or descending whereas *b* changes the whole line as a unit.

Initially, Adam had confused the actions on the graph produced by altering *m* (the slope) and *b* (the y-intercept). His "theory" was that changing *m* caused a vertical translation of the line and changing *b* caused the steepness of the line to change. His use of the words "incline" and "decline" were attempts to describe the change in steepness of the line. My instructions to explore the situation provided an external stimulus for Adam to question his ideas and to re-analyze the situation. This analysis allowed him to disprove his "theory" and to reach a different conclusion about the effects of changing the slope and y-intercept on the graph. He was able to conclude that changing the slope did not change the y-intercept and that the graph was "rotating" about this point. He also concluded that changing the y-intercept "moved the line up or down as a whole unit." These new conclusions helped to correct some misconceptions that Adam had about the relationships between the slope and the y-intercept to the graph.

I also wanted to find out the students' definitions for "average annual rate of change." This term had been used in the directions for the project. When asked to define this term, the students responded in one of two ways. Cindy and Kaitlyn said that the average annual rate of change was the slope. Adam referred to it as the amount of population growth per year. Thus, all three students knew that *m* and the average annual rate of change were the same.

Census. Other students in the class had asked Dr. Kirk about the census column of the spreadsheet (Class Observation, September 24, 29). They were concerned that the column was not completely filled with numbers. I was curious to see what Adam, Cindy, and Kaitlyn thought of this situation. The three students stated that the given census figures were the recorded populations for the years 1960, 1970, 1980, and 1990. They also said that there would be no values in the census column for census years beyond 1990 because those years had not occurred yet.

Predicted Values. The students defined the $P(t)$ values as predictions, projections, or estimates of the population. They stated that these numbers were not the actual populations for the given dates. They also stated that these numbers came from "plugging in" the values for m, t, and b into the linear modeling equation $P(t) = b + mt$.

All three students could find an estimated population for any given year by evaluating the optimal linear model equation. In Adam's interview on November 3, I asked him to estimate the population for the city of Anchorage in the year 2007. His response to this request was

> A: 2007? I would uh, instead of going up, dividing it by, instead of dividing the census by (pause) how would I do that, uh, first of all I would have to figure out how many years there are from 1960 to 2007. That should be 47 years. And I would probably put 291 in the census column (pause) I don't know how I would do that. I thought about dividing the new or the census, 1990 minus 1960 and I would divide it by 47, but I'm not sure if that would work. So I'm not really sure on how I would do that.

It was evident from his response that Adam did not have an immediate process for finding the estimated population. He correctly computed the number of years from 1960 to 2007, but he appeared to be thinking about slope instead of predicted populations. Because the year 2007 was not a census year, it was not listed on the spreadsheet. I thought this may have influenced Adam's thinking, so I decided to pursue the situation by asking him the same question about a census year that appeared on the spreadsheet.

S: What if I asked for the population for the year 2040?
What would you tell me?

A: 2040, well I guess I could do it this way, insert 6.7 into
my, as for my *m*, and use 22.5 for *b* and P(*t*) I'd have to
make it for 80 years and I would solve for 80 years.

S: What do you think you would get?

A: (laughs) I don't know. I'd have to get my calculator.
(Adam calculates the value using his calculator)

S: Could you do a similar process for 2007?

A: Right, I'm thinking I could do the same thing as well.
Except, uh *t* would be 47 . I would say 47, *b* would
be 22.5 and *m* would be 6.7 and *t* would be 47.

S: Ok, can you say that in an equation form for me?

A: Uh, where's my *b*, 22.5 plus 6.7 multiplied by 47.

S: And that would give you what?

A: That would give me the (predicted) population over the
period of 47 years.

Adam was able to find the appropriate *t* value, $t = 2040 - 1960 = 80$, and to use this value with his slope and y-intercept values in the optimal linear equation to determine the estimated population for the year 2040. He was then able to complete the same process for the year 2007.

Adam did not appear to realize that the estimated value for the year 2040 was shown on the spreadsheet under the P(t) column. However, he did recognize this fact during his exit interview (December 1). In this interview, Adam began using his linear model equation for Chicago to estimate the population for the city in the year 2050.

As is evident in the following excerpt from that interview, he suddenly realized that the estimated value was shown on the spreadsheet and that his work was unnecessary.

S: What's your equation for the linear model?

A: P(*t*) = 3,550,404 + -25,555*t* , *t* is the amount of years.

S: So what's the estimated population in 2050?

A: My *t* would be, wait a second, (counts to himself) my *t* would be eleven. So it would be *b* plus *m* and my *t* would be eleven. So P(*t*) would be 3, 330,404 plus negative 25,555 multiplied by 11.

S: Why eleven?

A: (long pause, counts to himself) Good grief, I did it wrong. One hundred ten, wouldn't it? (long pause) No, it's one hundred years.

S: Why one hundred years?

A: That's what I'm trying to figure out. (counts to himself again). Let's say ninety years. Is that better? Ok, why ninety years? Because that's the amount of time from the year 1960 to the year 2050.

S: Ok, can you tell me what the estimated population would be in 2050?

A: The estimation, the approximate would be 1,250,373 (Adam was looking at the computer screen and reading the value for the year 2050 in the P(*t*) column). That's just an estimate.

Adam had made several mental mistakes calculating the number of years from 1960 to 2050. He became so involved in this process that it was as if he had forgotten the question. When I repeated the question, he abandoned his process and read the number from the spreadsheet. I continued the conversation by asking Adam what would happen if he had continued his process using $t = 90$.

S: What would happen if you used your equation for the optimal model and let $t = 90$?

A: I would say it would be close to this (pointing to the value for 2050 in the P(t) column).

S: How close?

A: Um, it would be right on it, very close. I guess that's what it would be equal to.

S: Why?

A: Because it's calculating this formula for the spreadsheet.

S: What is that column, P(t)? Where do those numbers come from?

A: Well, the first four numbers come from the census.

S: For P(t)?

A: Oh no. Ok, that means the population in years. These numbers come from the linear growth model, from this formula implementing the –25,555 (the number he found for the slope of the model).

S: So the estimated population for the year 2030 would be what?

A: The estimated would be 1,761,491.

S: What about for the year 2080?

A: You want me to use the calculator? I would have to use the calculator. (The table of values on the spreadsheet only went up to the year 2050.)

S: How would you use the calculator?

A: First I would get my *b* which is 3,550,404. I'm going to add my *m* which is a negative number multiplied by 120 since there are 120 years between 1960 and 2080. We are running out of folks: 483,696.

Adam was finally able to connect the linear modeling equation with the *P(t)* column of the spreadsheet. He was also able to once again verbalize his process for finding the estimated population using the equation. Both Cindy and Kaitlyn could use the P(t) column to predict populations for the years indicated on the spreadsheet template. They could also use their optimal linear equations to find the predicted population for a given year.

Average Error. The students' definitions of average error and "line of best fit" uncovered some misconceptions. None of the students could define the sum of the squared errors (SSE) even though this number appeared on the spreadsheet. Cindy and Adam explained that average error was the difference between the predicted value *P(t)* and the census value.

C: ... But average error is how much our line is off from the actual points and how much we would expect it to be off in the future, plus or minus 2.5 thousand. (Interview, October 29)

A: The average error, the average error is 24.985 and that stands for, that, that, I believe it stands for 24,985 people. That's how much, that's how off you can or cannot be from your point. I'm guessing. I don't know.

S: OK, from which point?

A: From your, from the census point. Although I'm not positive about that. (Interview, November 3)

Adam was not certain about his definition. He and Cindy appeared to overlook or ignore the word "average." I continued the interview by asking about this point. Adam was unable to explain further. However, Cindy was able to adjust her definition.

S: So why is the difference between the census and the line (P(*t*) value) 3000 for the year 1980 and 4000 for the year 1990?

C: Um. (pause) Oh, I guess it's because of average. They must have added these up (referring to the numbers under squared error at the bottom of the spreadsheet) and divided by four.

Kaitlyn's explanation of average error surprised me. She began her definition much like Adam and Cindy. However, as the interview on October 29 continued, a misconception about this number emerged.

S: What is average error?

K: It would tell me what the error was over the um (pause) I can describe it like people. Like the 25 thousand people is like a plus or minus 25 thousand people between those numbers or whatever (points to spreadsheet).

S: Between the census and the P(*t*)?

K: I was thinking it was just like in those numbers like just the census numbers. Maybe, I don't know. It could be between the census and the P(*t*). I didn't really think about that.

As you can tell from the above conversation, Kaitlyn thought the average error represented the difference between the census values. She did not consider *P(t)* to be involved.

<u>Line of Best Fit.</u> The students did not have a completely formed definition for "line of best fit." In the beginning they suggested that a line of best fit goes through as many data points as possible. Cindy suggested that the line had to go through at least two of the data points. I encouraged her to think more about this idea.

S: So it's not possible for a best fit line to go through only one point or none at all?

C: No, a line must go through two points, but (pause) it is possible for it (the line) not to go through any data points. You want the line to be as close as possible to all the points. It doesn't actually have to go through them. You know, you could have four points and the line might go between them with two above and two below it (sketches graph on paper to demonstrate her idea).

As Cindy talked she sketched the following graphs (Figure 4) on a sheet of paper.

Figure 4: Cindy's sketch of best fit lines

Cindy was able to adjust her definition. She knew she must have two points to determine a line, but she was able to conclude that those points did not have to be the original data points.

Kaitlyn thought a majority of the data points was necessary. I asked her to explain a line of best fit in our interview on October 29.

K: I guess it's not like a completely accurate line. This is a problem because they don't know what the population is going to be in the next 50 years or whatever. But um it's kind of like a guess, an estimate, like the best estimate they can give of the next years.

S: OK, let's talk about the points and the line. For a line of best fit, how many data points does the line actually have to go through?

K: I would think like the majority, at least more than half. Like there's four points here and it should at least touch three of the four of them. So if it's a best fit it should touch most of them. If not go through them at least right on the side of them, like that one (pointing to the graph).

S: So you're saying it wouldn't be possible for it to not go through any of the data points?

K: Well, I guess it (long pause)(mumbles) I don't know.

S: So what was your definition of a best fit line?

K: It's the best estimate or the best overall thing that you can get. It's not completely accurate, but it's going to show um I guess it's just the best estimate. So I guess if it didn't go through any of them it would kind of like start there and go up between them like that. The points are really what happened so if the line doesn't go through any of those then maybe they should adjust their line, make it go back through like where the points are and stuff. Maybe their best fit is wrong.

Kaitlyn's last remarks point to another of her misconceptions. Because the census points are the accepted populations, the line must go through most of them. If it did not, then the best fit line was wrong. Kaitlyn thought the best fit line could not pass between the points. The idea of being as close to all the points as possible was not part of her definition.

Adam insisted that the line must pass through the first and last data point. I challenged his idea in our interview on November 3.

A: ... Is that the most logical solution that can go in that line or is it (pause) um after trying various figures, that, that best fit figure is the one that seems to be the, works the best and not, and most applicable to that, to that year, to that area. I think that's what best fit means to me. I'm not sure.

S: OK, what happens to your graph when you change *b* and *m*?

A: I would say the line comes closer to the actual points or it can go away from the points. I think the point of changing *b* and *m* is to attempt to have the line cross through all four points or at least come closest to all four points. And that's what *b* and *m* does, just moves the line back and forth, trying to um come closest to all the four points instead of just point one and point four.

S: Are you saying then that the line doesn't have to go through all the points?

A: Right, um sometimes it can be, for example, in my case it is very hard for it to go through all four of the points, because one point is considerably low, like 1970's point was almost well it was only 4000 away from the 1960 point, 4000 above it, and then in 1980 it rose to 174 thousand which is like let's just say 125 thousand more than the 1960 one. So it was like a sharp increase, where in 1970 there was not that much of an increase, so the line is obviously going to go from 1960 across to 1980 and 90 and it will leave 1970 without a line through it, because of the slow growth during that decade.

S: How many points would you say that the line would have to go through?

A: I would say it would have to go through point one and point four, 1960 and 1990, so two points total.

S: Why?

A: Because that would give you your linear graph.

S: Does your equation or your graph on sheet two go through point one and point four?

A: Yes, it barely touches the bottom part of point one. ...

A: Yes, my line on my second sheet does cross through point one and point four, just barely touching the bottom of the point in 1960 and crossing right through the middle of um the point of 1990. And the reason I, the reason I did this is I'm trying to get down to, as close as possible to 1970, and when I did get close enough to 1970 without leaving 1960 I was, I was still, I actually went through point three and right through point four. So I had, ... I managed to touch three out of the four points with the line.

This conversation with Adam introduced another misconception. Adam described the line as "barely touching the bottom" of a point and crossing "through the middle" of a point. The graph in Figure 5 was the picture on which Adam based his conclusion.

Figure 5: Adam's graph of his optimal model for Anchorage

As you can see from the figure, it "appears" that the line is barely touching the bottom corner of the first and third data points and is passing through the fourth data point. Kaitlyn also hinted at this "touching." She chose the same city as Adam (Anchorage, Alaska), so her graph looked almost the same as Adam's. She described the line as being "right on the side" of the point. However, Kaitlyn suggested that the line was very near the point, and Adam thought the point was on the line. He appeared unaware or unconcerned that the census value and the *P(t)* value were different for the place where "the line barely touched the bottom of the point." He also did not consider that the graph could be distorted. For Adam, as long as the line appeared to "touch" the point, he considered the point to be on the line. He did not appear to know or consider the geometric notion that a point does not have size or shape.

This misconception may have been exacerbated or produced by the technology. I explored this idea with Adam in the exit interview on December 1. I instructed him to graph his linear model equation and his census points on a graphing calculator. I then asked him to change the window for x and y several times and to describe the graph. He was able to observe the distortion in the graph as the scale for x and y became larger. He realized that the line did not pass through the points even though it appeared to do so with the larger window. Adam concluded that the graph on the spreadsheet was also distorted and that the line only "appeared to touch the points." I am not convinced that Adam recognized that a point does not have size or shape because of this exploration. However, he did realize that a "picture" can be misleading.

Linear Modeling Process. I wanted to determine how the students made sense of the linear modeling process. Why bother with such a procedure? Was it useful? Thus, I asked the students to state the purpose of linear modeling. They had not been given a statement of the purpose of linear modeling in class. Hence, their statements were based upon intuition, informal knowledge, and their experiences with linear modeling during the semester.

All three students suggested that it was a method of predicting what will happen in the future. However, their responses, such as the one provided by Kaitlyn (Interview, October 29) in the excerpt below, focused on population data.

S: What is the purpose of linear modeling?

K: It gives you like an idea of population or increase or decrease in something. I'm not sure. Just gives you an idea.

S: OK, an idea about what?

K: Like for populations, gives you an idea of whether it is going to increase or decrease in the next fifty years or um like maybe how fast or how steady it's going to be. If it's steeper then it's increasing faster, I think. Gives you an idea of how fast or slow it's increasing or decreasing.

S: What would be the advantage of knowing these things?

K: Prepare you like if you need to, if you know your town is going to increase then you can go ahead and start planning for the future, like the economy, like working on buildings like schools and more houses and more malls and banks. So you can actually have jobs for people. Or if it's going to decrease, you can figure out how you can … (mumbles) I guess just figure out what's moneywise best for your town. Try to make everybody happy with the town.

Thus, their ideas about linear modeling revolved around the population problems that were presented in the course. They did not seem to consider the idea that linear modeling could help us predict the future about other types of mathematical situations.

During the exit interviews, I again asked the students to state the purpose of linear modeling. I wanted to determine if their focus remained on population problems at the end of the semester. Adam continued to cling to the population situation. He stated that the purpose was “to predict future population growth.” I asked if it had to concern population growth, and he responded that it could be “descending.” He appeared to have no thoughts of any other type of problem situations that would be appropriate for linear modeling. Kaitlyn’s response was similar. She stated that the purpose was to “maybe give a round about look at what the population is going to be.” Cindy provided a better and possibly more insightful answer. She had previously stated that the purpose was “to predict future populations” (Interview, October 29). She now gave the following explanation in her exit interview on December 9.

> C: It helps you hypothesize about the future. It helps you represent data that you have. It gives you a way of visually seeing abstract thoughts and ideas.

Even with Cindy’s insightful explanation there was still no mention of specific types of problems that could be done with linear modeling. Eventually I was able to get them to consider other problems. In the exit interviews, I asked them for some examples of other types of problems that could be modeled with the spreadsheet template. Kaitlyn’s example was that we could consider a problem involving how many miles or how long it takes to get from one place to another (Exit Interview, December 9). Cindy suggested that we could use the spreadsheet for “probably anything ... even the amount of cows on a farm” (Exit Interview, December 9). Adam came up with several possibilities in his exit interview on December 1.

> A: Probably like something in business, maybe supply and demand problems. Say if you were in the toy business and wanted to predict the demand for certain toys or the number of employees needed to make those toys. You could apply it to economics or industry, assembly type work, or airlines. Like ferrying passengers back and forth, making estimates of how many planes you need or how many mechanics you need to work on the planes.

Because I asked about the spreadsheet, I am not sure that I would have received the same responses if I had asked them to give examples of problems that are appropriate for linear modeling.

Problem-Solving Strategies

When faced with a mathematical situation, there are certain strategies that a person can use to help resolve the situation. Polya (1957) described many of these strategies in his books on problem solving. Some of these strategies included drawing pictures, considering special cases, and looking for a similar problem. The students in this study relied on the *look for a related problem* strategy for homework, quiz, and test problems. These problems were similar to those presented in class. Thus, the students had only to relate them to an appropriate class problem to determine the method for finding a solution.

The students relied on the *look for a related problem* strategy and a *systematic trial and error* strategy for their projects and exit interview problem. Adam, Cindy, and Kaitlyn used a systematic "trial and error" method with the spreadsheet template to find the optimal linear models for their population data. They focused almost solely on three cells of the spreadsheet: *b*, *m*, and *average error*. The rule of the game was to vary *b* and *m* until the average error was as small as possible. During the interviews, the students elaborated on this strategy.

Cindy provided a brief description of her strategy (Interview, October 29, 1998).

C: Well, once you get your slope and y-intercept you have your beginning equation. Then you use trial and error to change m and b so that the average error is the smallest it can be.

S: Explain the trial and error to me.

C: You just keep changing *m* and *b* until you think you've got the smallest number for average error.

S: Was there a pattern to how you changed *m* and *b*?

C: No, I pretty much just used trial and error, just picking numbers.

S: How many choices did you try for *m* and *b*?

C: I don't know. I think I had a list in my project.

S: Here's your project.

C: It must have been only four.

S: You didn't do more than these?

C: No, I recorded all the ones I did.

S: How did you know you had the best values?

C: Because that's the smallest average error I could find. When I chose 107 and 109 (for *b*) with slope 2.4 the error went back up so I knew 108 (*b*) and 2.4 (*m*) gave the smallest error.

Cindy did not go into detail on choosing *m* and *b*. She also did not choose very many values before deciding her average error was the smallest it could be. However, during the exit interview on December 9, she demonstrated a strategy similar to the ones described by Kaitlyn and Adam, and she admitted using "logic" to figure out the best choices for *m* and *b*. Kaitlyn provided more insight into the trial and error strategy in her interview on October 29.

K: I just kept playing because you had to make the average error as low as it could go. I never got as low as everybody else's. I think my numbers are just different. But um, I had to make that and that (pointing to SSE and Ave Error cells) as low as possible and so I just kind of played with it and I had to like decrease that number (*b*) but I had to increase my slope (*m*). So it took me like 15 or 16 tries actually, so I could figure out which numbers I had to increase and decrease and how far to go before, because I followed it like to 23 and that number (average error) started increasing again and I followed it to like 6.7 or something and that number started increasing again, so I just had to figure out which numbers gave me the lowest.

S: So how did you know when you had the lowest? Was there a pattern to how you plugged in the numbers?

K: Well at first I just picked, I picked like 43 and like 6 point, well I kept 6.07 just to see if 43 would make it increase or decrease, and it made it um decrease, I think, yeah. Then I messed with the *b* number until it was as low as it could, and then I went back and I played with the *m* number until I made it as high as it would go but still had the average error as low as it would go. And then I went back and I changed the *b* to see if it like made the number any smaller. And um, I didn't really have a good strategy, I just kept putting numbers in, like that were lower or higher than the ones I put in before to make my average error as low as possible.

Adam's strategy was similar to Kaitlyn's, but he persevered longer. In an interview on November 3, Adam talked about his trial and error strategy.

A: Page 2 (of the spreadsheet template) is where you take the values of sheet one and you start toying with them a little bit. Where *b* was 44 you start putting various numbers in for *b* and also you can start putting various numbers in for *m* where *m* initially was 6.066. What you try to do is get the lowest average error possible. And uh, I finally got it up to about 30 or 40 tries. At least to my knowledge it was the lowest. *b* came out to 22.5 and *m* came out to 6.7 thousand, and it reduced my average error from, the initial average error of 28.663 to the new average error of 24.985.

S: OK, you said you tried 30 or 40 possibilities?

A: Right.

S: Was there any strategy to how you chose those numbers?

A: Yeah, (pause) I would, I would lower *b* a little bit, say I would lower it by half a percent, say it was 23, I would go down to 22.5 and I would leave m constant Yeah I would leave *m* at the original and if I saw it got lower then I would go ahead and lower the percent and leave *b* alone. Try to do them both at, no, one at a time versus trying to do both at a time. When I saw I couldn't lower it any more, then um doing it that way, then I would just, that's when I would just guess, I would just kind of throw something in there a little bit radical to see what it would do and if it started going down again then I would toy with it again just little by little.

Adam's system for choosing *m* and *b* was much like Kaitlyn's. However, he tried more than twice as many values for *m* and *b*, and he added an extra element of guessing. After systematically changing *m* and *b* until he could not lower the average error any more, he would choose what he called a "radical" value for one of them. Adam's perseverance and strategy proved beneficial. His optimal model was almost identical to the one found with a graphing calculator using the linear regression key. Adam's values were: m = 6.7, b = 22.5, and average error = 24.985. The values found on the calculator were: m = 6.72, b = 22.2, and average error = 24.984.

This *systematic trial and error* strategy may seem rather simple. However, it is a complex process. Posamentier and Krulik (1998) referred to this strategy as the *intelligent guess and test* strategy. The students had to have beginning values for the slope (*m*) and y-intercept (*b*). These values were not just random numbers chosen by the students. The students were instructed to assign the census value for the year 1960 to the y-intercept (*b*) and to calculate and assign the slope of the line passing through the first and last data points to the slope (*m*) of their model. The initial values for *m* and *b* and the resulting *average error* were recorded in a table. They then selected a variable (*m* or *b*) to change and held the other constant. For instance, suppose the slope (*m*) was changed and the y-intercept (*b*) was held constant. The first "guess" for *m* was a slightly larger or smaller number than the initial value. These new values for *m*, *b*, and *average error* were added to the table. The students analyzed the effect of this change on the average error and made their next "guess" based on this analysis. They were attempting to make the average error as small as possible. Thus, each guess was an "informed or intelligent" guess. Once satisfied that they could no longer reduce the average error by altering *m*, the students held this value for the slope constant and altered the y-intercept (*b*). This process of changing *b* and analyzing the effect of the change on the average error continued until the students thought they had reduced the average error as much as possible. At this stage the process of altering the slope (*m*) began again. The students' willingness and tenacity to find the lowest average error determined the duration of the entire process. Adam became so involved in the process that it became a challenge to find a lower average error. He spent several hours altering *m* and *b*. In contrast, Cindy was satisfied with just a few (only four) "guesses."

The graph of the linear model did not appear to be important in the *guess and test* strategy. The tremendous attention devoted to the average error seemed to overshadow any benefits of using the graph to make intelligent guesses for *m* and *b*. During the exit interviews (December 1, 8, 9, 1998), I observed the students as they searched for a linear model for Chicago's population data. All three students focused on the average error and began recording their results. They did not mention the graph, and it did not appear that they even looked at the graph during this process. I questioned them about this issue. Kaitlyn responded that she did not really look at the graph except to make sure the line did not do anything "strange." Her definiton of "strange" was that the line disappeared from the graph or did not go through any of the data points. Cindy suggested that she was probably supposed to look at the graph, but that she just watched the increase and decrease in the average error. Adam also responded that he did not place much emphasis on the graph when choosing values for *m* and *b*. In his exit interview (December 1) he stated:

> I pretty much just used trial and error. I didn't pay much attention to the graph. I didn't think about that if I change *b* the graph's going to go down. I just picked a number and tried to make the average error as small as possible.

<u>Practices</u>

The students' nonuse of the graph to make decisions while using the guess and test strategy was one indication of the students' mathematical practices involving the spreadsheet template. The students treated the spreadsheet as a "black box" (the same as Dr. Kirk's presentation of the spreadsheet). The linear model was represented in three ways on the spreadsheet: an equation, a graph, and a table of values. The students almost completely ignored the graph and the table of values. Their main goal was to make the average error as small as possible. They readily admitted that they did not understand how the average error was computed. They assumed that the squared errors and the sum of the squared errors that were shown on the spreadsheet were involved, but had no idea how they were involved. Cindy confirmed this practice in our interview on October 29, "I don't know what SSE is. I didn't use it. I don't think we need it." She later said, "... I just look at the average error to see if it goes up or down" (Exit Interview, December 9). Kaitlyn admitted looking at the graph on occasion, but stated that she did not focus on it while adjusting *m* and *b*. She expressed, "I may glance at it to make sure it's not doing something really crazy, but I just concentrate on the average error" (Exit Interview, December 8).

Another indication of the spreadsheet being seen as a black box was evidenced in the students' seeming lack of connections among the representations of the linear model. I mentioned in the previous section on the students' knowledge base that the students could find a population value *P(t)* for any given year by evaluating the optimal linear model equation. In fact, Adam chose to do this for every year, including those years listed in the table of values on the spreadsheet. Recall the excerpt from the interview with Adam on November 3.

S: What if I asked for the population for the year 2040? What would you tell me?

A: 2040, well I guess I could do it this way, insert 6.7 into my, as for my m, and use 22.5 for b and P(*t*) I'd have to make it for 80 years and I would solve for 80 years.

S: What do you think you would get?

A: (laughs) I don't know. I'd have to get my calculator.

I expected Adam to look at the table of values and read the value of *P(t)* for the year 2040. Instead, he calculated the value using his optimal model equation. Adam's mathematical practice was to use the equation (not the table or the graph) to answer questions about populations. He did not appear to recognize the relationship between the *P(t)* column and the linear model equation or the graph. In an interview on November 3, Adam provided his thoughts about any connections between the graph and the (*P(t)* and *census*) columns of the spreadsheet.

A: ... the census, I don't see how they are related because I've actually changed that census, those census figures to the new P(*t*): 23, 90, 157,224. So those points do correspond with the points on the graph.

S: Ok, I want to make sure I understand what you are saying. The 23 corresponds to …

A: The 23 corresponds to the new figures for 1960 and 90 for 1970, 157 for 1980, and 224 for 1990.

S: Show me where the 90 would be for 1970. Where is that point on the graph?

A: I'm going to have to retract that, because I see that, well that's not quite 90 (pointing to individual dot on graph, actually the value for the census for 1970, laughs) I'm not sure (mumbles) No, I guess not. Well, I retract that 'cause I'm not sure what, I thought P(*t*) would do to that but obviously it's still going with the census, because the census is the one that goes, it's just below, it's below, it's around 50, so it's 44, and then it goes a little bit up to 48 and then to 174 and 226 for 1990. So I don't see the figures for P(*t*), but I do see the figures for the census column (pointing to the individual dots on the graph). So I guess I'm retracting what I said earlier.

S: So the four dots on the graph are the census?

A: Yes

S: Where does the line come from?

A: The line, the line comes from the linear growth model $P(t) = b + mt$ I would say, because that's what's moving when you adjust your, when you adjust for average error that's what's moving is the actual line, so uh, that's what I see there.

Adam was able to eventually connect the census values with the graph, the linear equation with the graph, and the linear equation with the P(t) column. However, he was unable to make the connection between the P(t) column and the graph.

Cindy and Kaitlyn's mathematical practices did include limited use of the table of values to answer questions about populations. They made use of the $P(t)$ column of the spreadsheet to answer questions about the estimated population of a city for a given census year. However, neither student referred to the graph to answer any questions. This behavior further strengthened the idea of the spreadsheet being a black box for the students.

All three students' mathematical practices pertaining to linear modeling were limited to what they had done in the course. They did not appear to seek help from resources outside the classroom environment (such as books or people not involved with the course) and were resistant to considering these outside resources. Average error became the authority in determining the best models. In the semi-structured interviews (October 29, November 3) with the students, I asked how they would find the optimal linear model for a set of data without using the spreadsheet. All three students responded that they would use paper and pencil and attempt the procedure from the spreadsheet by hand. The following are a few of their comments from these interviews.

S: How would you model your data without using this spreadsheet?

C: I don't know. Um (pause) You would do it with pencil and paper. You would have to choose your m and b and then plug in all these points and calculate the average error by hand. Then we would choose another

m and *b* and repeat the process. (Interview, October 29)

K: I guess you would try to do it by hand, use graph paper and draw it all out. (Interview, October 29)

A: I guess you would do a table of sorts. ... You'd have to figure out all this other stuff down here, the squared error with how the calculations for the, how the average error is ... found. (Interview, November 3)

In the exit interviews I asked the students to find the optimal linear model for a set of data (pulse rate reductions, [Appendix D]) using the spreadsheet and then without using the spreadsheet. Cindy and Kaitlyn were able to adapt the spreadsheet to this data. Thus, they were able to carry out their usual procedure for finding the optimal linear model with this data. Adam was unable to adapt the spreadsheet. He became very frustrated. With much coaching from me, Adam eventually wrote down the appropriate columns to replace the *years*, *t*, and *census* columns of the spreadsheet. He was then able to find values for *m* and *b* and to continue in the usual manner. When asked to find the optimal model without using the spreadsheet, each student found the slope between the first and last data point and assigned the value of the first data point to *b* (the same procedure for the spreadsheet). Then they repeated their idea of "doing the spreadsheet by hand." During his exit interview on December 1, Adam summed up the sentiments of the students.

S: How would you find the optimal linear model for this data without using the spreadsheet?

A: I don't know how to explain it, but I know it deals with this squared error down here. I'm not sure how you would do it.

S: Are you saying you would do what the spreadsheet does, but you would do it by hand?

A: Right.

S: Is there another way you could do the problem?

A: I don’t know of another way. I don’t know how you could tell how it adjusts the average error.

I wanted to determine if I had exhausted the students' abilities or if I could push them to derive another method. Because the students would not suggest or could not suggest another way of finding a linear model for the data and had ignored the graph on the spreadsheet, I suggested that they graph the pulse rate data. All three students were able to draw a graph of the data. The following conversation with Cindy during our interview on December 9 illustrates what happened.

C: ... You want me to graph it?

S: Sure, you can graph it. How would you find your initial model?

C: I guess you could find the line between the first and last points.

S: OK, draw that line. How good a representation is that line?

C: Not too good. It’s not close to all of the points.

S: Sketch a line that you think would be a better representation of the data.

At this point Cindy drew another line that was approximately equidistant from all the points and the conversation continues.

S: Can you find an equation for that line?

C: Sure, you could estimate a couple of points on the line and find your slope and equation from them. But I don’t know how I would know that this is the best line, without doing all that other stuff.

S: All that other stuff being the squared error and average error stuff on the spreadsheet? You would do that by hand if you understood it?

C: Yeah. It would take a long time, but you could do it (to find the optimal line).

Cindy found the equation for her first line using the first and last data points, (50, 15) and (70, 2). She estimated two points, (50, 16) and (70, 3), for her second line and found the equation using these estimated points.

The students were so "in tune" with their spreadsheet procedure that all three drew an initial line through the first and last data points. Their second lines represented the data better, but only Cindy and Kaitlyn could find equations for their lines. Adam could not carry the process further. He appeared tired and was ready to end the interview.

Figure 6: Cindy's graph for the pulse rate problem

Figure 7: Kaitlyn's graph for the pulse rate problem

The graphs that Cindy and Kaitlyn drew for the pulse rate problem are shown in Figures 6 and 7. Cindy's line was perhaps the best. She attempted to draw a line the same distance from all the points. I found it interesting that Cindy estimated the two points corresponding with the first and last data points. This seemed to indicate that she was unwilling to completely give up her spreadsheet procedure. Kaitlyn drew her second line through the second and fourth data points. She then estimated her initial point as (50, 17). She used this point with the second data point (55, 13) to find the equation of her line. Unfortunately, because she estimated her initial point and used it to find the equation, this was not the equation for the line passing through the second and fourth data points. Kaitlyn did not seem to realize this discrepancy.

As evidenced in Cindy's interview, the students still would not abandon the idea of calculating the average error by hand and using it to determine the optimal linear model. No other methods of determining the optimal model emerged in the interviews.

Another mathematical practice of the students that emerged from the data was to restrict the types of numbers that could be used for the y-intercept (*b*). When choosing values for *b*, Cindy and Kaitlyn selected only whole numbers. They gave their reasons in interviews conducted on October 29.

S: You only used 107, 108, and 109 for the y-intercept, *b*.

Why didn't you consider decimal numbers for it?

C: Because we rounded off the numbers to the nearest thousand (pointing to the census numbers on the spreadsheet).

S: You mean the census numbers?

C: Yeah, he (Dr. Kirk) told us to round them to thousands ... .

S: I noticed ... for all of your *b* values they were whole numbers. Was there a reason for just using whole numbers for *b*?

K: I didn't even think about using decimal places. (laughs) I guess that's why. Really, I didn't even think about using decimal points for my *b* value. I didn't think about it because it didn't have one in the original value for b. So I just thought about using the same kind of numbers.

The two students restricted their choices to the type of number used in the census column of the spreadsheet and did not consider the advantage of allowing for decimals. That is, it "made sense" to them to choose the same type of number with which they began. They were once again employing the *look for a related problem* strategy. I asked them what they thought would happen if they did allow *b* to be a decimal number. Both agreed that they may be able to lower the average error and obtain a better model.

Once the students obtained their optimal linear models they were asked to interpret their findings in both the projects and the interviews. These acts of interpretation are part of their mathematical practices. Cindy chose the city of Little Rock, Arkansas for her project. Both Adam and Kaitlyn chose Anchorage, Alaska for their projects. All three cities had increasing populations. I asked the students how realistic they thought their models were for predicting future populations for their chosen cities. Their responses were:

> C: They are probably OK for now. But, in a lot of years they're not very reasonable. The line is going to keep growing, increasing and increasing. No city is going to do that. You would have an infinite number of people which isn't possible. It isn't possible to put everybody on earth in one city. At some point the city will level off, begin to drop off. For instance, the gold rush, people leave town after the gold's gone and the town dies.
>
> S: So all towns will eventually drop off?
>
> C: Yeah, I think so. There's no way they can keep growing and growing. (Interview, October 29)
>
> K: I think it's pretty good. The population's increasing because a lot of people are moving to Alaska for some reason and my line is increasing. I don't know if it will increase by 600 thousand like it's suppose to. (Interview, October 29)

> A: I would have to say, to me it's realistic. ...
> Because you can figure that (pause) Anchorage, Alaska is kind of a far out place, far away place and it probably wouldn't have more than a quarter million. The census for 1990 is 226 thousand, so 50 or 60 years later I wouldn't expect the population to be much over three quarters of a million the way it is now. I'm not sure that many people want to move up to Alaska, so probably less population growth there. (Interview, November 3)

The students used their informal and intuitive knowledge about the real world to decide on the fit of the model. All three students appeared to believe the model fit the real world for

the near future. However, Cindy was the most adamant that it was not a good model for the long term.

During the exit interviews, I asked the same question about the students' models for Chicago, Illinois. Chicago's population was decreasing. Their responses to this question were:

> K: I guess it could be (reasonable), but you never know what's going to happen to a city. Like they could build a great mall and everybody would want to go live there. ... I wouldn't look at it long term, maybe short term kind of planning. ... I guess it's appropriate for like the times that have already happened and like the next fifty years. Then after fifty years switch to another model. (Exit Interview, December 8)

> C: ... Every year it's dropping off. So that seems about right. ... At some point it's going to be zero and keep on decreasing until your numbers are just ridiculous. So your model would need to be updated. A city's population is not going to go to zero unless you have a nuclear war or something. ... It's probably (a good predictor) when you are not going too far into the future. You don't have to extend your line very far. (Exit Interview, December 9)

> A: We are running out of folks. ... The population is still decreasing for the year 2080. If you keep decreasing, you

> are going to run out of people. ... It (Chicago) may lose population, but it will never run out of people. I don't see that happening. I don't know where it would stop at, but I don't believe it will lose all of its people. ... it's showing that Chicago is losing population and eventually after so many years Chicago will no longer have a population. I don't feel I could use this (model) for that (predicting). ... After so many years I would have to say it's unrealistic. Even at 2050 we have an estimated value of 423 thousand. I don't think Chicago would ever get that low. It's too big of an industrial city. Lot of stuff going on there. ... I don't remember the formula, but maybe the bounded would be good. It would give you a fixed figure that you would never go below. (Exit Interview, December 1)

Again the students agreed that the linear model may not fit the real world in the far off future. They felt that a zero or a negative population would not make sense for Chicago. Their sense-making and interpretations were based on their intuitive knowledge of the "real world." Adam made a good observation by suggesting that the bounded growth model may be more appropriate for Chicago. The students were asked to compare the bounded growth model with the natural growth and linear models in the third phase of the course project. However, Adam was the only participant to mention another type of mathematical model as possibly being better for Chicago's situation.

In part two of the course project each student was asked to compare the linear model with the previous natural growth model and to choose the better model for his or her chosen city. The following excerpts from the three project reports gives the students' opinions about the better model.

> **Adam's opinion** (Part two of course project, October 22):
>
> I have given you two predictions to consider, both the natural growth and linear growth models. It is my professional opinion that we should consider the findings of the linear growth model to be our best route of action in planning for our city's future. Since it would seem doubtful that our city will have a population of 7 million in the year 2050. The conservative figures of the linear growth model (626 thousand) seem to be more conceivable and realistic.
>
> With these figures we can more readily plan for the future of this city. We must consider allocating land for future landfills and housing

developments. Our infrastructure must be expanded in order to accommodate future demands. We need to recruit industry to our area in order to ensure that there will be jobs for our citizens. This is a long term solution; long term solutions are accomplished by meeting short term goals and we must work toward developing proactive plans to enhance our communities resources.

Cindy's opinion (Part two of course project, October 22):
Looking at the predicted populations for Arkansas's capital city, we notice an alarming amount of growth using both the linear and natural growth models. Because the linear growth model has only an average error of 2.5 thousand, it seems reasonable to rely more on this equation than the natural growth model which has an average error of 4.543 thousand. Unless the city limits are expanded, Little Rock will be extremely overcrowded. Plans for new high-rise apartments are being made in order to accommodate citizens. With such an enormous growth, surely the economy will be booming; however, with no place to live, people might look to the streets for shelter. Crime and pollution are likely to become serious problems as well for Little Rock if its population continues to grow at this rate. Regardless of what tomorrow may bring, Little Rock's government is dedicated to making Arkansas a better place to live.

Kaitlyn's opinion (Part two of course project, October 22):
These predictions for the city will enable us, here in Anchorage, to prepare for the increasing population growth in the next fifty years. These findings are as accurate as they can be based on predictions. The natural growth model, using the formula **P(t)=A(1+r)^t** is the most accurate model used. It's [sic] average error was 24.833, compared to the average error of the linear model, which was 25.020. This is only a .187 difference, but the lesser chance of error, the greater chance our town has of preparing for the future. According to the natural growth model our population for the year 2050 is predicted to be 6,255,000. ... These findings are very important in the future planning of our city. As we begin to build more schools, homes, and plan more jobs, these numbers will allow us as citizens of Anchorage, Alaska to adjust and strengthen our economy as needed.

Two different mathematical practices emerged from these excerpts. One practice was to base the choice of the better model on informal and intuitive knowledge about the real world. Adam based his decision about the better model on the prediction values that the models produced and whether he thought the populations could (or would) reach those values. He did not consider the average error that had been so instrumental in determining the optimal models. This is particularly significant because the average error for his linear growth model (24.985) was larger than the average error for his natural growth model (24.882). It appeared not to matter that the numbers (average errors) might suggest that the natural growth model was better. Adam did not think the natural growth model fit the real world situation. He reaffirmed this belief in an interview on November 3.

> A: The natural growth model goes all the way to 7 million people for the year 2050 and the linear growth model goes up to 626 thousand for the year 2050. To me it would be more feasible to project for 626 thousand versus 7 million because I just don't perceive 7 million people moving up there in the next 50 to 60 years. I would think everything would be easier on taxes and government to plan for 626 thousand versus 7 million.

The second practice that emerged from the project excerpts was to base the choice of the better model on algorithmic procedures used in the course. Cindy and Kaitlyn's practice of determining the better model was opposite from Adam's practice. They adhered to the practice of making decisions based on the average error. Both students chose the model with the lower average error as the better model. They thought this model fit the real world situation. Thus, the lower average error should give them the better model. After all, the goal of the entire semester was to make the average error as small as possible. The fact that they were comparing average errors from two different models did not appear to concern them. The decision to use this "method" for selecting the better model was an intuitive and logical conclusion drawn from their experiences in the class. However, it may not be a reliable method. If the population data had been fit with a polynomial model of higher degree, the average error would probably have been less than the one for the linear or natural growth models. Because the students' population data were always increasing or always decreasing, this model would not have been a good model for predicting future populations for the chosen cities.

Monitoring and Control

When confronted with a mathematical situation, students perform activities such as reading the problem, analyzing the situation, exploring the situation, planning for a possible solution, implementing the plan, and verifying the results. These activities do not occur in a particular order, and they may be performed more than once in a problem-solving situation. The students in this study exhibited all of these actions with varying degrees of frequency. Most of their time was spent implementing and verifying. Exploration and analysis usually occurred during the process of finding the optimal linear model. Planning was the least occuring action.

The students were very procedural in their actions. They would read the problem and implement the appropriate procedure (spreadsheet or class method). There was little need to plan for a solution. The plan had been given to them in the form of the spreadsheet procedure or the class method for working the homework problems. They assumed (and rightly so) that the problem would be one of these two types. They analyzed each problem to determine the appropriate related problem for the given situation. Their main exploring was done while using the *guess and test* strategy to determine the optimal models. This exploration involved making intelligent guesses and analyzing the results of the guesses to determine the next phase of the exploration. Verification was the most prominent activity of monitoring and control used by the students. They verified solutions by checking computations and occasionally glancing at a graph. Their verification of the optimal model was their eventual acceptance that they could not find a lower average error. This acceptance came in the form of loss of interest in further exploration, satisfaction with the amount of "guesses" for the slope (m) and the y-intercept (b), or inability to find a lower average error.

I was able to observe several specific occasions in the data where two of the students' monitored their actions and adjusted their behavior accordingly. Adam described an experience he had with this type of behavior in a conversation that we had a few minutes before the class session on October 22.

> A: I went upstairs to see Dr. Kirk because I was afraid I was not doing things correctly. I got 7 million for the population of my city in the year 2050 with the natural growth model. But, I got only 600,000 people with the linear growth model. That is a big difference, so I thought I was wrong. That's why I went to see Dr. Kirk. He said I was doing things correctly.

Adam read the problem, implemented the spreadsheet procedure, and attempted to verify his findings. His usual method of verification (acceptance that he had the lowest error) was not sufficient. The differences in the numbers (predicted populations) for the two models caused enough concern for Adam that he sought verification from other sources (Dr. Kirk and me).

Another example of Adam's disposition to monitor and change his actions occured during the third test. As mentioned earlier in this chapter, Adam answered the first two parts of the linear modeling question on the third test correctly. However, he missed the correct answer to the third part of the question by one month. It was clear from his test paper that Adam had written the correst response (November) originally, but had changed his mind. He erased November and gave October as the answer. Thus, Adam had carried out his plan for a solution, attempted to verify his answer, decided that he had reached the wrong conclusion, and made what he thought was the appropriate change. Unfortunately,

I failed to pursue the issue of why he thought he was wrong in his first attempt to verify his answer.

A third example of monitoring and control exhibited by the students occurred during Cindy's exit interview on December 9. I asked Cindy to use the spreadsheet to find a linear model for the city of Chicago. She calculated her slope (*m*) using the slope formula for two points. However, she did not realize that she had inverted the formula. When she entered her values for *m* and *b* in the spreadsheet, she remarked, "Something's wrong. The line doesn't go through my points." Cindy retraced her steps and realized her mistake. She then correctly calculated her value for *m* and continued to find the linear model. The graph served as verification for Cindy that she had correctly carried out her plan for a solution. When she did not obtain verification, she analyzed the situation and made appropriate adjustments.

In the previous examples, the stimuli for monitoring their actions and altering their behaviors came from within the students. There were also occasions in the study where outside stimuli "triggered" these actions in the students. One such occasion was Adam's difficulty calculating the *t* value for the city of Chicago in the year 2050 (Exit Interview, December 1). Earlier in this chapter, I described Adam's attempt to find the estimated population for the city of Chicago in the year 2050. His initial response to this task was:

> A: My *t* would be, wait a second, (counts to himself) my *t* would be eleven. So it would be *b* plus *m* and my *t* would be eleven. So P(*t*) would be 3, 330,404 plus negative 25,555 multiplied by 11.

I questioned his intial *t* value. This caused him to reflect on his answer and to attempt to adjust his calculations.

> S: Why eleven?
>
> A: (long pause, counts to himself) Good grief, I did it wrong. One hundred ten, wouldn't it? (long pause) No, it's one hundred years.

In this response, he questioned himself and re-analyzed the situation. Unfortunately, he still did not calculate the *t* value correctly and did not appear to need to verify his results. Thus, I questioned his answer once more.

S: Why one hundred years?

A: That's what I'm trying to figure out. (counts to himself again). Let's say ninety years. Is that better? Ok, why ninety years? Because that's the amount of time from the year 1960 to the year 2050.

Again, Adam analyzed the situation and re-calculated the *t* value. This time he mimicked my prompt. He sought verification for his answer and provided an appropriate reason for his result.

Beliefs and Affects

Although I did not focus on the students' mathematical beliefs and attitudes, some of their views of mathematics appeared in the data. Two of the three participants thought of mathematics as being difficult. Adam briefly described his feelings about mathematics in an e-mail on September 24.

S: How do you feel about mathematics?

A: I have lots of apprehension and anxiety.

In an interview on November 3, he expanded on his feelings.

S: What was your general opinion of mathematics in high school?

A: Um, I guess you could say I was terrified, because I knew I wasn't very strong in math. I was hesitant to take a harder math class although I could have. I just figured well I don't want to fail, so I'm going to try something a little bit easier like consumer math so I can get that credit and press on.

S: What's your opinion of math now?

A: It comes to me easier because I believe I have a determination to succeed - now that I guess that I am paying for it out of my own pocket and that I am mature and that I see that I need to make it through this course and that I don't really have a problem with it. I'm still apprehensive, but I'm not scared of it.

As you read in this last excerpt from Adam, his view of mathematics has begun to change. He is "not scared of it" and "it comes ... easier." However, the tension and apprehension are still there.

Kaitlyn expressed even stronger feelings about mathematics. In an e-mail on September 23, she said,

> I do not like math a lot, most of the time it seems useless, since I want to teach second grade, I don't feel like I need to know all the stuff I have to learn. But, it is something I tolerate since I don't have a choice. Along with being math incompetent, I am also computer illiterate. I have never worked with computers that much and I have never ever done a spreadsheet, so I'm a little nervous about this project.

She also talked about her feelings in our interview on October 29.

S: OK, what was your general opinion of math when you were in high school?

K: I didn't like math. (laughs)

S: Why?

K: Because I always didn't do very good in it. I mean I always got B's or whatever. But it was always my worst grade. I just always thought it was kind of stupid.

S: OK, has your opinion changed any since you have been in college?

K: (laughing) Not really. I'm just waiting for this semester to be over and hopefully I'll pass and then I won't have to take any more.

Kaitlyn's attitude about mathematics had not changed from high school. She did not see the point of studying mathematics or the usefulness of mathematics.

Cindy did not have strong opinions or negative opinions about mathematics. In an interview on October 29, she said, "Math was OK. I didn't think it was that bad (in high school)." When asked how she felt about mathematics now, she responded, "It's still OK. This class is pretty easy." Because mathematics was easy for her, it was an OK subject.

Adam, Cindy, and Kaitlyn also reacted to the use of spreadsheets. None of the students had any experience with spreadsheets. I asked them what they thought of the spreadsheet during an interview. Adam responded:

> I thought it was cool. ... I enjoyed manipulating my numbers and um, it's kind of a challenge to try to get it to lower the average error. It would be harder without the spreadsheet, ... it kind of does the work for you. All you do is input numbers and it kicks you back the average error automatically. I found the spreadsheet to be easy. (Interview, November 3)

Kaitlyn admitted being scared of using the spreadsheet, but felt better about it as the semester continued. She stated:

> I didn't even know what a spreadsheet was until I had to do this. I freaked out when I found out I had to use a spreadsheet because I'm like computer illiterate. This is the first year that I've ever done e-mail. At least I know what one is now and what you can use it for. I don't know if you can use it for other things or not. I don't have any idea. At least I know if I ever have to do a model I know I can use a spreadsheet. (Interview, October 29)

At the end of the semester (Exit Interview, December 8), Kaitlyn remarked that the spreadsheet does "most of the work for you." She continued by stating that "it automatically gives you the average error." Cindy echoed this response during the semester. She stated, "It would take a lot of work and a lot of time to do this modeling process by hand" (Interview, October 29). Then at the end of the semester she remarked:

> C: ... You can find the optimal model looking at the average error without really knowing how it is calculated. The spreadsheet does it for you. If you were going to do it by hand you would have to understand how to find the squared error and the average error. It could also take a long time by hand. (Exit Interview, December 9)

The apprehension, anxiety, and tension over using computers and spreadsheet software dissolved as the semester progressed. These words were replaced with “cool” and “easy.” By the end of the course the students felt that using technology (spreadsheet template, computer, and graphing calculator) had made their lives easier. They seemed to recognize and appreciate the time that was saved by using the spreadsheet template to find the optimal linear models.

The students’ overwhelmingly procedural nature seemed to imply that they believed mathematics was a procedural process. The procedural emphasis in the class allowed this belief to be a significant component for success in the mathematical modeling course.

The students attitudes toward mathematics did not appear to change during the semester. Kaitlyn still thought of it as useless and stupid. Adam mostly tolerated the subject because it was required. Cindy held to her neutral opinion of mathematics.

CHAPTER FIVE
SUMMARY AND CONCLUSIONS

The purpose of this study was to investigate college students' understanding of linear modeling when using a spreadsheet template to model population data in a mathematical modeling course. I endeavored to describe the students' understanding of linear modeling as they progressed through a mathematical modeling course. To accomplish this goal, I focused on the students' mathematical thinking and problem solving during the course. This focus in the research suggested I needed an organized structure for examining the students' mathematical thinking and problem solving behavior. Schoenfeld's (1992) framework for exploring mathematical cognition provided such a structure. The framework consisted of five categories: the knowledge base, problem-solving strategies, monitoring and control, beliefs and affects, and mathematical practices. These categories provided an organized structure for decomposing the students' understanding of linear modeling into manageable parts and analyzing these parts. Because of the coherent nature of the categories, they also provided a lens for looking at a students' understanding of linear modeling as a whole.

The study was conducted during fall semester of 1998. A qualitative case study approach as described by Merriam and Simpson (1995) was used for this research. Data were collected from observations, interviews, and written documents. The data were then analyzed according to the qualitative method of constant comparison that was described by Corbin and Strauss (1990). The results of the analysis were presented and discussed in terms of the theoretical framework and the research questions. This chapter presents conclusions drawn from the results of the data analysis, implications from the study, and recommendations for further research.

Conclusions

Four main conclusions or themes emerged from the results of the data analysis in this study:

- The students seemed notably procedurally oriented and, in particular, obsessed with their "average error procedure."
- The students treated the spreadsheet template as a "black box" and, consequently, failed to make effective use of available representations of the modeling situation.
- The students' life experiences influenced their interpretation and sense-making (mathematical practices) of the mathematical situation.
- The students formed opinions, made decisions, and communicated their ideas about linear modeling when asked to do so.

Like Schoenfeld's categories, these themes are not independent of each other. The students' procedural nature and obsession with the average error certainly encouraged their "black box" treatment of the spreadsheet template and failure to use the available representations. Also, the students' life experiences possibly contributed to the development of their procedural orientation.

Throughout the semester the students were asked to use a spreadsheet template and to consider the average error in determining optimal models. The use of this average error became a dominant procedure for them in the course. They accepted this procedure as the authority for determining the best model within a given modeling situation and even across modeling situations. This use of average error seemed to become an obsession with the students. Every conversation, observation, and written document contained numerous referrals to average error and its role in finding an optimal linear model. This obsession is possibly a result of the students' *procedural disposition*.

Searcy (1997, p. 153) described *procedural disposition* as "a habit of thought that is primarily focused on the acquisition and use of procedures." According to Silver

(1987), procedural knowledge, as labeled by Anderson (1976), is formed through the abstraction and classification of individuals' experiences and is stored in long-term memory. This knowledge, characterized by Ryle (1949) as "knowing how," shapes what individuals see and how they behave when encountering new situations. The students in this study were certainly procedurally driven when solving problems. They patterned homework problems after classroom examples and focused on their average error procedure for the course project. This procedural disposition had an extensive influence on interpreting the data from this study in terms of Schoenfeld's (1992) categories.

The students' knowledge bases were dominated by definitions, meanings, and procedures. Informal and intuitive knowledge played a secondary role to these elements. Misconceptions were also an important factor in the knowledge bases of the students. These misconceptions arose in part because of the students' obsession with the average error procedure and their apparent unwillingness to consider other alternative methods. During the course, and even at the end of the semester, none of the students were able to conceive of an alternative method of linear modeling for the given population situations. Also, some of the misconceptions were possibly exacerbated by the technology used in the course.

Most of the problems in the mathematical modeling course could be described as exercises. Polya (1945) and Mayer and Hegarty (1996) referred to these types of problems as routine problems. Schoenfeld (1985a, 1992) explained the distinguishing feature of exercises or routine problems. With routine problems, students already know a reliable path for obtaining a solution. Thus, the use of exercises in this course indicated that the students' problem solving behaviors may not exhibit the "true spirit" of problem solving such as that described by Duncker (1945), Polya (1945), Charles et al. (1987), Lester and Kroll (1990), and Schoenfeld (1983, 1985a, 1992). This type of problem in the course curriculum possibly contributed to the students' procedural orientation.

In this study, the path for obtaining a solution usually existed in the form of a known procedure. The students had been shown the procedure for using the spreadsheet template and the procedures to be used with the homework problems in class sessions. This significantly reduced their need to rely on many of the problem-solving strategies discussed in the problem-solving literature. The use of routine problems immediately determined the need for the *look for a related problem* strategy. Within the spreadsheet environment, the students relied on Posamentier and Krulik's (1998) *intelligent guess and test* strategy to determine the optimal linear model. No other problem-solving strategies appeared necessary for successful handling of the problem situations.

The students' mathematical practices involved interpreting and making sense of the linear modeling situation. Like the other categories, these practices were strongly influenced by the students' procedural disposition. Procedures, definitions, and meanings were used to interpret and make sense of the linear modeling situation. When these elements were not sufficient or created a conflict within the students, the students' informal and intuitive knowledge helped complete the sense-making process. These practices seemed to come from within the linear modeling environment. The students relied on procedures, the instructor, and their informal and intuitive knowledge about the real world to make decisions about the linear modeling situation. Outside resources such as books or people not involved in the class did not appear to be needed or desired in the interpretation and sense-making processes.

The students exhibited (to some degree) the actions of self-regulation (monitoring and control) described by Schoenfeld (1985a, 1987, 1989,1992), Brown (1987), and Silver (1987). However, their disposition to monitor and control their actions during a problem-solving situation was largely reduced to actions of implementation and verification. Because of the procedural actions of the students, especially their attention to average error, there was little need for analyzing and planning in the linear modeling situation. The students had been given plans for the desired solutions. They only needed to implement these plans and verify the solutions. Implementation of the plan for finding the optimal linear model did require that the students engage in some exploration and analysis of the mathematical situation. Verification existed in the forms of checking computations and occasionally glancing at a graph to make sure "nothing strange was happening."

The students' beliefs and attitudes about mathematics, computers, and spreadsheets contributed to their success in the mathematical modeling course. Their procedural disposition suggested that they believed mathematics was a procedural process and that computers, especially spreadsheet templates, were the best tools for completing this process. This procedural belief about mathematics was enough for the students to be successful in the mathematical modeling course. The course project required the students to interpret their findings and make predictions about the modeling situation. However, this project was only a small percentage of the course grade, and the interpretation and sense-making for the project was only a fraction of this percentage.

McLeod (1992) suggested that students will develop positive or negative attitudes about mathematics when faced with the same or similar situations repeatedly. The students in this study used spreadsheet templates for three modeling situations: natural growth, linear, and bounded. They recognized the benefits of using the spreadsheet templates for the modeling situations. They were easy to use, made calculations much easier, and saved time. These factors led to the students developing a positive attitude toward the spreadsheets, and thus, the linear modeling situation. These positive attitudes were similar to those found in Lanier's (1997) pilot study.

The second major conclusion or theme resulting from this study was the students' treatment of the spreadsheet as a "black box," and their subsequent failure to make effective use of available representations of the modeling situation. NCTM's *Principles and Standards* (1998), as well as many other sources, have emphasized the importance of multiple representations in mathematical modeling and mathematics in general. The spreadsheet template provided the students with three representations of the linear modeling situation: a linear equation, a table of values, and a graph. The students in the study did not appear to consider the significance of having these multiple representations. Perhaps this is due to their overwhelming attention to procedures, especially their average error procedure.

The students made little use of the graph or the table of values when searching for their optimal models. Their focus was on average error, which in turn, placed their focus on the linear model equation and the values for the slope (m) and the y-intercept (b). The equation became the major form of representation for the linear modeling situation. The students did not question why or how the "average error procedure" gave them their optimal model. They just "did what they were told." In fact, the students did not completely understand the computation for the average error and the relationships it had to the predicted values and the graph.

In essence, the students treated the spreadsheet template as a black box. Because the template was presented in the course as a black box, the students' treatment of it is not surprising. The average error cell of the template was the mysterious key to finding the desired solutions and being successful in the course. It appeared that the students could have carried out the procedure without concern if the census, slope, y-intercept, and average error cells were the only visible parts of the spreadsheet. Although they were successful in the course, this "nearsightedness" resulted in the students' missing connections and forming misconceptions about the linear modeling situation. The students felt the spreadsheet would "do most of the work" for them. Thus, it was unnecessary to understand how it worked. Therefore, the students developed an "incomplete" understanding of the linear modeling situation.

A third conclusion or theme from this study was that the students' life experiences influenced their interpretation and sense-making of the modeling situation. This seemed to support Resnick's (1988) ideas about the importance of society in the development of mathematical practices. These life experiences included "job-related" experiences as well as "academic" experiences. Adam used more than his "academic" experiences for making sense of the linear modeling situation. He used his informal and intuitive knowledge and personal theories about the real world for interpretation and sense-making. He based his decisions about the fit of a model to data and the choice of the better model on what he thought would occur in real life. He also used "real world language" to describe the modeling situation. He rarely used mathematical terms such as slope and y-intercept. He referred to the elements (*m*, *b*, and *t)* in the equation for the linear model as the amount of population change per year, the population of our initial year, and the number of years that have elapsed since our initial year.

Cindy and Kaitlyn used informal and intuitive knowledge and personal theories about the real world for interpretation and sense-making to a lesser degree than Adam. Their "academic" experiences seemed to have the greatest influence on their interpretation and sense-making practices. They based their decisions about the fit of a model and the better model on procedures that were used in the course and on the size of the average error. Also, they were more "algebraic" with their language. They used words such as slope, y-intercept, and linear equation to describe the modeling situation. These words are typical of the language used in most algebra classrooms. However, the connections between these words and the real world were not always clear to the students.

The fourth main conclusion or theme from this study was that the students could form opinions, make decisions, and communicate their ideas about linear modeling when asked to do so. One important aspect of the course project was to require the students to interpret their linear model and discuss its usefulness and reliability in predicting future populations for their chosen cities. The students fluently expressed their thoughts and opinions about the linear modeling situation in their written project reports. All three of the students accepted one of the mathematical models (natural or linear growth) as a fairly good predictor of populations for the near future. However, each student was unwilling to accept the models' reliability over a long time. They suggested that cities and the world are constantly changing and that the situation should be re-assessed in about fifty years to determine if the model still "fit" the situation. This suggestion indicated the students were able to go beyond their procedural orientation. They believed that the lowest average error produced the best model, but they were also aware of events in the world that influenced population. Thus, the students drew conclusions about the modeling situation not only from procedures, but also from their knowledge and personal theories about events in the world.

During the interviews, the students often gave a quick first response to many of the questions. It seemed that little thinking went into these responses. However, with further questioning, the students expressed their ideas better and even rethought and reformed some of these ideas. For example, Cindy's quick response to the question about the number of data points through which a line of best fit must pass was two. However, in exploring this topic further, she decided that a line of best fit does not have to pass through any of the data points, but must come as near to all the data points as possible. She not only expressed her views orally; she also drew pictures to represent her ideas. Thus, the students were adept at communicating their ideas in written, verbal, and graphical forms.

In summary, the students in this study could use a spreadsheet template to find an optimal linear model for a set of population data. This task was well suited for their procedural nature. This procedural nature strongly influenced their mathematical thinking and problem solving behavior. It was so ingrained in the students that they appeared to become obsessed with the average error procedure used in the course. They based their interpretations and decisions on procedures, intuitive knowledge, and personal theories about the world. They failed to make effective use of available representations and knowledge. This oversight may have been encouraged or even produced by the technology and the structure of the course. The students recognized that linear modeling was a method of representing a real world situation and predicting (or estimating) future outcomes of the situation. They accepted the mathematical models as viable predictors of the current situation, but realized that there were other factors that should be considered. When encouraged to do so, the students formed opinions about the linear modeling situation and effectively communicated their ideas using written, verbal, and graphical representations.

Implications from the Study

Several implications for the mathematics education community can be drawn from this study. These implications relate to improving student learning through better understanding of students' mathematical thinking and problem solving behavior and changes in instructional practices and curriculum. They also relate to improving and updating the theoretical frameworks used in research.

Mathematical Thinking and Problem Solving Behaviors. The results of this study indicated that students may have a strong procedural disposition that overshadows all other dispositions. This procedural nature was so ingrained in their thought and behavior that the students would modify (or at least attempt to modify) any situation so that it fit within this realm. The importance of this disposition cannot be overlooked. Mathematics educators must recognize the strength and influence of a procedural disposition on students' mathematical thinking and problem solving behavior. Only then can they begin to develop mathematical tasks and situations that will lead students beyond this realm to a more conceptualized understanding of mathematics.

A goal of any teacher should be to help students develop the ability to place the mathematics involved in a mathematical situation into the real world. After all, the purpose of mathematical modeling is to represent real world situations in mathematical terms (a model of the situation) that can be manipulated and then interpreted with respect to the real world situation (Dossey, 1990; NCTM, 1998). Adam's ability to relate the linear model to the real world situation gave us a glimpse of this goal. However, his ability was probably due more to his combined experiences in life than just those in this course. This suggests that mathematics educators need to determine the life experiences that each student brings to a mathematical situation and to provide activities that build on these experiences.

Curriculum and Instructional Practices. The spreadsheet template used in this study was an innovative method for introducing linear modeling to students. It has the potential for being used with students at several levels of the educational system. Besides college students, the spreadsheet could be used easily with secondary school students and possibly middle school students. This is particularly true if it is treated as a black box. However, the "black box" treatment should not be the most desirable approach to using the spreadsheet. Typically, students mimic out of class what they are shown in class.

Students should be encouraged to make effective use of the multiple representations presented within the spreadsheet template. This may be accomplished by spending more instructional time on average error and the sum of the squared errors (SSE). If students are taught the meaning of these numbers and how to calculate them, some of the mystery of the "black box" would disappear. Of course, spending more instructional time on a topic often requires spending less time on other topics. This is a constant issue in making decisions about curriculum and instructional practices.

Getting students to notice and use multiple representations for mathematical situations may also require explicit instruction and strategic questions from teachers about the benefits of each representation and the connections among the representations. It was evident from this study that making multiple representations available to students does not guarantee that they will use them. It has also been reported in NCTM's *Standards* (1989) and *Principles and Standards* (1998) that students may have difficulties making connections on their own. They must be encouraged and guided in this process.

Dugdale et al. (1995) encouraged researchers to assess the potential and actual consequences of implementing new curriculum and instructional practices. "What are the consequences of implementing this type of course at the college level?" Schoenfeld (1983) suggested that a desired consequence of any course that we teach should be that we provide our students with thinking skills that they can use in future situations. The students in this study were very adept at using the spreadsheet template to model population data. However, only two of the three students could adapt the spreadsheet template to problems other than population situations. Also, none of the three students could use their knowledge of the linear modeling situation to derive a different method for finding the optimal linear model. Thus, are we providing them with thinking skills that they can use in future situations? The results of this study suggest probably no. The students' thinking skills seemed to be hindered by their procedural disposition and possible "obsession" with the average error. Also, their treatment and acceptance of the spreadsheet as a "black box" seriously decreased the need for improving and strengthening their thinking skills.

A possible consequence of the course curriculum and instructional practices (including the use of technology) is that it encouraged the students' procedural disposition. Also, the technology used in the course may have inspired the students to form misconceptions (such as a line barely touching a point).

Schoenfeld's Framework. Schoenfeld's (1992) framework provided an adequate theoretical structure for examining students' mathematical thinking and understanding. It allowed for the decomposition of the data and the re-assembly of the results of the analysis to produce a complete picture of the students' mathematical understanding of linear modeling.

Considering the results of this study, as well as Searcy's (1997) study, students' attention to procedures seemed to be a major component of their mathematical thinking and problem-solving behavior. This procedural nature had an enormous influence on their mathematical understanding.

Procedures are embedded in Schoenfeld's *knowledge base* category. However, the students' procedural disposition strongly influenced all five of his categories. Perhaps procedures and procedural dispositions should have a stronger emphasis in Schoenfeld's framework and other frameworks for exploring mathematical cognition

Recommendations for Further Research

This study added to research (Lester & Kroll, 1990; Schoenfeld, 1992) on mathematical thinking and problem solving. It reaffirmed many of the conclusions that Searcy (1997) formed about mathematical thinking in a modeling course. Even though progress has been made in research in this area, many questions remain unanswered.

Some of these unanswered questions involve curriculum issues. In particular, what is the students' mathematical understanding of linear modeling if

- the students were encouraged to use the multiple representations within the spreadsheet to make decisions and connections in the linear modeling situation?
- the students were provided with a detailed explanation and exploration of average error and sums of squared errors?
- the instructional practices within the course placed more emphasis on interpretation and sense-making and less emphasis on procedures?
- the course content required students to apply linear models to situations other than population?
- the students were required to develop their own spreadsheet templates?
- the students were presented with more than one method for linear modeling?
- the students were allowed to work together on the course project?

In conducting research on these issues, mathematics educators should continue to identify the consequences of such curriculum changes (Dugdale et al., 1995). Along with these questions, research needs to address the use of this type of spreadsheet template at the secondary level. Would secondary students possess a similar understanding of linear modeling to college students? Would the secondary classroom provide a better atmosphere and more time for interpretation and sense-making?

Other important research questions that need to be investigated include learning, sense-making, and technology issues. For instance, how common is a "procedural disposition" among college students? What factors contribute to the development of this procedural disposition? How can mathematics educators de-emphasize this procedural disposition and encourage students to develop a deeper conceptual understanding of mathematics?

Regarding sense-making, what life experiences does each student bring to a mathematical situation? How can mathematics educators build on these life experiences and help students strengthen their interpretation and sense-making skills? How can these skills be effectively used to develop a deeper understanding of a mathematical situation and to make connections between this situation and the real world?

As for technology issues, to what extent are misconceptions produced by technology? How can we anticipate and correct these misconceptions? What is the impact of new technology on student understanding and learning?

These questions may only be answered by building on existing theoretical ideas, such as those presented by Resnick (1988) and Schoenfeld (1992), and developing new theoretical "lenses" for analyzing data. These frameworks must provide the necessary tools for examining and identifying a student's procedural disposition, life experiences, and reactions to technology.

A Final Thought

This study allowed me to look examine students' understanding of linear modeling when using a spreadsheet template. In this chapter, I have attempted to summarize the results and conclusions of the study. Also, I have discussed implications of the study and provided ideas for further research in the areas of curriculum, learning, and sense-making.

My experience in conducting this study has heightened my awareness of students' inclination to focus on procedures and to accept those procedures without question. It has also affirmed the importance of students' life experiences on their interpretations and sense-making, and thus, on their mathematical understanding. My teaching will forever be influenced by this study. In future college mathematics classes, I will always wonder about and question a student's understanding of the mathematics involved and my influence on this understanding.

REFERENCES

Anderson, J. R. (1976). *Language, memory, and thought*. Hillsdale, NJ: Lawrence Erlbaum.

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Brown, A. (1987). Metacognition, executive control, self-regulation, and other more mysterious mechanisms. In F. Reiner & R. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 65-116). Hillsdale, NJ: Lawrence Erlbaum.

Charles, R., Lester, F., & O'Daffer, P. (1987). *How to evaluate progress in problem solving*. Reston, VA: National Council of Teachers of Mathematics.

Clayton College and State University. (1999). *MATH 1101: Mathematical modeling* [on line]. Available: http://www.math.clayton.edu/1101

Corbin, J. & Strauss, A. (1990). Grounded theory research: Procedures, cannons, and evaluation criteria. *Qualitative Sociology, 13*(1), 3-21.

Demana, F., & Waits, B. (January 1993). *Using a numerical and graphical approach to teaching mathematics*. Unpublished manuscript, Ohio State University, Columbus.

Dossey, John. (1990). *Discrete mathematics and the secondary mathematics curriculum*. Reston, VA: National Council of Teachers of Mathematics.

Dugdale, S., Thompson, P. W., Harvey, W., Demana, F., Waits, B. K., Kieran, C., McConnell, J. W., & Christmas, P. (1995). Technology and algebra curriculum reform: Current uses, potential directions, and research questions. *Journal of Computers in Mathematics and Science Teaching, 14*(3), 325-357.

Duncker, K. (1945). *On problem solving*. Psychological Monographs 58, No. 5. (Whole # 270). Washington, DC: American Psychological Association.

Edwards, H. (1997). *Math 104: Applied College Algebra*. Unpublished manuscript, University of Georgia, Athens.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist, 34*, 906-911.

Greeno, J. G. (1973). The structure of memory and the process of solving problems. In R. Solso (Ed.), *Contemporary issues in cognitive psychology. The Loyola Symposium* (pp. 103-133). Washington, DC: Winston.

Hiebert, J. (1985). (Ed.) *Conceptual and procedural knowledge: The case of mathematics*. Hillsdale, NJ: Lawrence Erlbaum.

Hilke, E. V. (1995) *Mathematical modeling. Fastback 392*. Bloomington, IN: Phi Delta Kappa Educational Foundation.

Kantowski, M. G. (1977). Processes involved in mathematical problem solving. *Journal for Research in Mathematics Education, 8*, 163-180.

Kilpatrick, J. (1967). Analyzing the solution of word problems in mathematics: An exploratory study. *Dissertation Abstracts International, 28*, 4380A. (University Microfilms No. AAI68-5442).

Kilpatrick, J. & Davis, R. B. (1993). Computers and curriculum change in mathematics. In C. Keitel & K. Ruthven (Eds.), *Learning from computers: Mathematics education and technology* (pp. 203-221). Berlin: Springer-Verlag.

Koss, R., & Marks, R. (1994). The teacher and evaluation. *Mathematics Teacher, 87*(8), 614-617.

Krulik, S. (Ed.). (1980). *Problem solving in school mathematics*. (Yearbook of the National Council of Teachers of Mathematics). Reston, VA: NCTM.

Lanier, S. M. (1997). *The impact of an applied college algebra course on student perceptions of success*. Unpublished manuscript, University of Georgia, Athens.

Lester, F. K. (1985). Methodological consideration in research on mathematical problem-solving instruction. In E. A. Silver (Ed.), *Teaching and learning mathematical problem solving: Multiple research perspectives* (pp. 41-69). Hillsdale, NJ: Lawrence Erlbaum.

Lester, F. K. & Kroll, D. L. (1990). Assessing student growth in mathematical problem solving. In G. Kulm (Ed.), *Assessing higher order thinking in mathematics* (pp. 53-70). Washington, DC: American Association for the Advancement of Science.

Lucas, J. (1974). The teaching of heuristic problem-solving strategies in elementary calculus. *Journal for Research in Mathematics Education, 5*, 36-46.

Mayer, R. E. (1992). *Thinking, problem solving, cognition* (2nd ed.). New York: Freeman.

Mayer, R. E. (1994). Mathematical ability. In R. J. Sternberg (Ed.), *Encyclopedia of intelligence* (pp. 127-150). New York: Macmillian.

Mayer, R. E., & Hegarty, M. (1996). The process of understanding mathematical problems. In R. J. Sternberg & T. Ben-Zeev (Eds.), *The nature of mathematical thinking* (pp. 29-53). Mahwah, NJ: Lawrence Erlbaum.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81-97.

Merriam, S. B., & Simpson, E. L. (1995). *A guide to research for educators and trainers of adults* (2nd ed.). Malabar, FL: Krieger.

McLeod, D. B. (1992). Research on affect in mathematics education: A reconceptualization. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 575-596). New York: Macmillan.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics.* Reston, VA: NCTM.

National Council of Teachers of Mathematics. (1998). *Principles and standards for school mathematics.* [on line draft]. Available: http://standards-e.nctm.org.

Norman, D. (Ed.). (1970). *Models of human memory*. New York: Academic Press.

Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2nd ed.). Newbury Park, CA: Sage.

Pollack-Johnson, B., & Borchardt, A. (1999). *Using the TI-82 and Spreadsheets in a 1st-Year Math Course for Business and Social Science (The Villanova Project)* [on line]. Available: http://renoir.vill.edu/faculty/brucepj/html/fipse/project4.html

Polya, G. (1945). *How to solve it: A new aspect of mathematical method*. Princeton, NJ: Princeton University Press.

Polya, G. (1965). *Mathematical Discovery: On understanding, learning, and teaching problem solving* (Vol. 2). New York: John Wiley & Sons.

Posamentier, A. S., & Krulik, S. (1998). *Problem-solving strategies for efficient and elegant solutions: A resource for the mathematics teacher*. Thousand Oaks, CA: Corwin Press.

Resnick, L. (1988). Treating mathematics as an ill-structured discipline. In R. I. Charles & E. A. Silver (Eds.), *The teaching and assessing of mathematical problem solving* (Vol. 3, pp. 32-60). Reston, VA: The National Council of Teachers of Mathematics.

Rubenstein, M. F. (1975). *Patterns of problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Ryle, G. (1949). *The concept of mind.* London: Hutchinson.

Salisbury State University. (1999). *MATH 165: Introduction to Mathematical modeling* [on line]. Available: http://henson2.ssu.edu/~dccathea/math165

Schoenfeld, A. H. (1983). *Problem solving in the mathematics curriculum: A report, recommendation, and an annotated bibliography*. Washington, D.C.: Mathematical Association of America.

Schoenfeld, A. H. (1985a). *Mathematical problem solving*. Orlando, FL: Academic Press.

Schoenfeld, A. H. (1985b). Metacognitive and epistemological issues in mathematics understanding. In E. A. Silver (Ed.), *Teaching and learning mathematical problem solving: Multiple research perspectives* (pp. 361-380). Hillsdale, NJ: Lawrence Erlbaum.

Schoenfeld, A. H. (1987). What's all the fuss about metacognition: In A. H. Schoenfeld (Ed.), *Cognitive Science and mathematics education* (pp. 189-215). Hillsdale, NJ: Lawrence Erlbaum.

Schoenfeld, A. H. (1989). Teaching mathematical thinking and problem solving. In L. B. Resnick & B. L. Klopfer (Eds.), Toward the thinking curriculum: Current cognitive research (pp. 83-103). (1989 Yearbook of the American Society for Curriculum Development). Washington, DC: ASCD.

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. 3-38). New York: Macmillan.

Searcy, M. E. (1997). Mathematical thinking in an introductory applied college algebra course. *Dissertation Abstracts International, 58*, 3056A. (University Microfilms No. AAI98-07131).

Silver, E. A. (1987). Foundations of cognitive theory and research for mathematics problem-solving instruction. In A. H. Schoenfeld (Ed.), *Cognitive science and mathematics education* (pp. 33-60). Hillsdale, NJ: Lawrence Erlbaum.

Smith, D. (1999). *Math 1525: Calculus with matrices.* Virginia Polytechnic Institute and State University Department of Mathematics [on line]. Available: http://www.math.vt.edu/people/smithdb/class_home/1525_3552.html

Smith, J. P. (1973). The effect of general versus specific heuristics in mathematical problem solving tasks. *Dissertation Abstracts International, 34*, 2400A. (University Microfilms No. AAI73-26637).

Stanic, G. & Kilpatrick, J. (1988). Historical perspectives on problem solving in the mathematics curriculum. In R. I. Charles & E. A. Silver (Eds.), *The teaching and assessing of mathematical problem solving* (Vol. 3, pp. 1-22). Reston, VA: The National Council of Teachers of Mathematics.

University of Georgia Department of Mathematics. (1999). *MATH 1101: Mathematical modeling* [on line]. Available: http://www.math.uga.edu/1101/index.htm.

Wiggins, G. P. (1993). *Assessing student performance: Explaining the purpose and limits of testing*. San Fransisco: Josey-Bass.

Wilson, J. W. (1967). Generality of heuristics as an instructional variable. *Dissertation Abstracts International, 28*, 2575A. University Microfilms No. AAI67-17526).

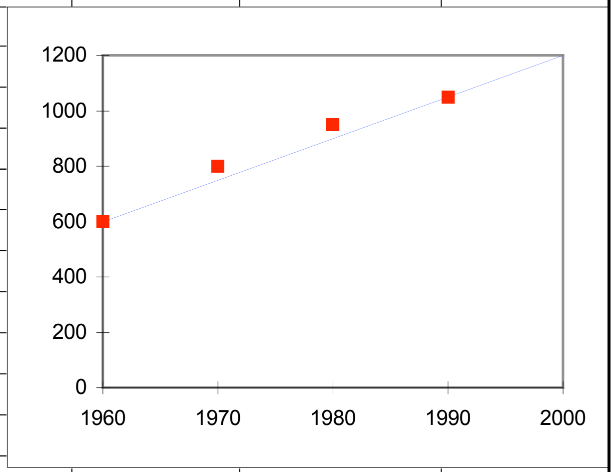
Wilson, J. W., Fernandez, M. L., & Hadaway, N. (1993). Mathematical problem solving. In P. S. Wilson (Ed.), *Research Ideas for the Classroom: High School Mathematics* (pp. 57-78). New York: MacMillan.

APPENDIX A

SPREADSHEET TEMPLATE

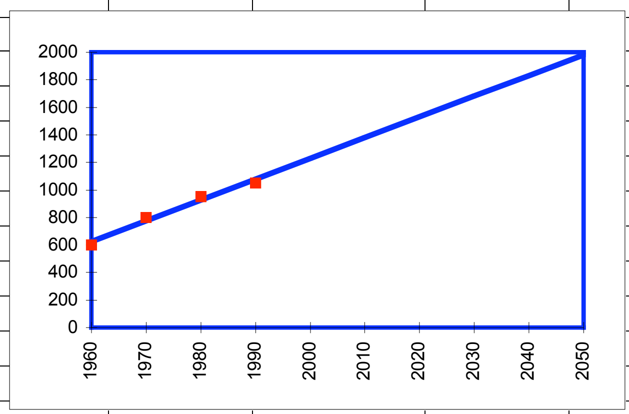
SHEET 1

PROJECT SPREADSHEET 2				**b =**	**600**	Filename: **LINEAR.XLS**	
Linear Growth Model				**m =**	**15**		
P(t) = b + m t				**SSE =**	**5000**	**Ave Error =**	**35.355**
CENTURY CITY							
Year	t	Census	P(t)				
1960	**0**	**600**	**600**				
1970	**10**	**800**	**750**				
1980	**20**	**950**	**900**				
1990	**30**	**1050**	**1050**				
2000	**40**		**1200**				
2010	**50**		**1350**				
2020	**60**		**1500**				
2030	**70**		**1650**				
2040	**80**		**1800**				
2050	**90**		**1950**				
Sqrd Err							
0				NOTE: See Sheet 2 also.			
2500							
2500							
0							
THE GAME:		Try to choose b and m to make SSE as small as possible.					



SHEET 2

PROJECT SPREADSHEET 2				**b =**	**623**			
Linear Growth Model				**m =**	**15.1**			
P(t) = b + m t				**SSE =**	**2506**	**Ave Error =**	**25.030**	
CENTURY CITY								
Year	t	Census	P(t)					
1960	**0**	**600**	**623**					
1970	**10**	**800**	**774**					
1980	**20**	**950**	**925**					
1990	**30**	**1050**	**1076**					
2000	**40**		**1227**					
2010	**50**		**1378**					
2020	**60**		**1529**					
2030	**70**		**1680**					
2040	**80**		**1831**					
2050	**90**		**1982**					
Sqrd Err								
529								
676								
625								
676								



APPENDIX B

SEMI-STRUCTURED INTERVIEW GUIDE

1. The linear growth model $P(t) = at + P0$ has two parameters: a and P0. What does each of these parameters mean to you?

2. In your project you were asked to determine the average annual rate of change for your chosen data. How did you determine this number?

3. Why is this number a reasonable first estimate for the parameter a in the linear model?

4. How did you get your first estimate for the parameter P0?

5. Why is this value a reasonable first estimate for P0?

6. Explain how you used the linear growth spreadsheet.

7. How did you determine the linear model function that best fits your data?

8. Are the parameters for this function the same as your initial parameters? Why?

9. What does "best fit" mean to you?

10. How well do you think your linear growth function represents your data?

11. What does average error mean to you? minimum average error?

12. How did you determine projected populations for years beyond 1998?

13. How realistic do you think these numbers are? Why?

14. How well do you think your linear growth function predicts the population of your chosen city for future years?

15. In your own words, explain the linear modeling process and its purpose.

APPENDIX C

PART TWO OF COURSE PROJECT: LINEAR MODELING

MAT 1101 Project

Draft 2 including Parts 1 and 2 -- Due Friday, October 23, 1998

Continuing in your role as consultant (or whatever) to the mayor (or whatever) of your city, you are now to revise Draft 1 as suggested, and also add Part 2 described below.

Part 2: Linear Growth Model

To prepare for your spreadsheet work, determine your city's average annual rate of change during the 30-year period from 1960 to 1990. This average rate is a reasonable first estimate of the long-term annual rate of change a of Century City. The recorded 1960 population is a reasonable first estimate of the initial population parameter P0 to use in the linear model **P(t) = a t + P0** for predicting the future population growth of Century City.

Next, use the linear growth spreadsheet (linear.xls) to determine the initial population parameter **P0** and rate of change a that empirically minimize the average error for the 1960-1990 census figures for your city. You will want to include in your report at least the following two spreadsheet graphs showing both the known actual data points and the growth curve obtained by plotting the linear growth function **P(t)**.

* The population line for the period 1960-2000 using your initial estimates of a and **P0**.
* The population line for the period 1960-2050 using your final "best fit" values of a and **P0**.

Your narrative should state explicitly your linear model function that best fits the population of Century City. For instance, you might say "We find that the linear model **P(t) = 15.1 t + 623** best fits the available census data for Century City, with an average error of 25,030 persons." And your report will surely include at least the following table (extracted from your spreadsheets and included in the body of your discussion of the future of Century City).

* The projected populations for the years 2000, 2010, 2020, 2030, 2040,

and 2050 using your previous best fit natural growth model (in one column) and your current best fit linear model (in another column).

How this table might look is illustrated below. Add anything you can to shed light on the situation. Certainly you will want to mention which fit enjoys the least average error over the 1960-1990 period, and hence which projection for the year 2050 seems the most likely. Remember that your salary (and perhaps your job itself) depends on your accuracy, completeness, ingenuity, and reliability. You may want to enhance your report with additional spreadsheets -- perhaps, a single spreadsheet chart showing both the best fit line and the best fit natural growth curve for the period 2000-2050.

Some More Specific Suggestions

Spreadsheets
For each type of model, you should have two spreadsheets - one for 1960-1990 representing the past (before optimizing the parameters), and one for 1960-2050 projecting the future (after optimizing). Generally it seems best for your spreadsheets to be appended (on separate pages) as "exhibits" that are referred to in the body of the report - rather than inserted in the discussion itself. But a specific chart from a spreadsheet might well be included in the body of the report. For example, you might say something like "The following chart from Exhibit 2 indicates the growth in the population of Century City that the linear model **P(t) = 15.1 t + 623** predicts."

(With the chart having been Copy/Pasted from the spreadsheet to your word-processing document.) When you include or refer to such a chart, you should mention what model it refers to (as specified by the formula quoted above) and what average error is involved. For instance: "The average error (comparing the linear model **P(t) = 15.1 t + 623** and the actual 1960-1990 census data for Century City) is about 25,000 persons. For a population in future years exceeding 1 million, this is less than 3% error, so we have some confidence in our projections."

Tables

Concise summary tables for inclusion in the body of your report should be extracted from your spreadsheets. For instance, if we combine the natural growth and linear growth best fits for Century City, we get the following table.

Year	Natural Growth Projection (thousands)	Linear Growth Projection (thousands)
2000	1287	1227
2010	1531	1378
2020	1821	1529
2030	2166	1680
2040	2576	1831
2050	3064	1982

Summary

Perhaps this is the most important part of your report, where you draw your conclusions and make your professional suggestions. For instance, looking at the table above, you might say "Our natural growth model $\mathbf{P(t) = 643(1.0175)^t}$ predicts a Century City population of 3.064 million in the year 2050, whereas our linear growth model $\mathbf{P(t) = 15.1\ t + 623}$ predicts a population of only 1.982 million in 2050. Because the average error in the natural growth model is about 38 thousand, whereas the average error in the linear growth model is only about 25 thousand, it seems reasonable to rely more on the linear model. We therefore believe it prudent to anticipate for planning purposes that by the year 2050 the population of Century City will have grown to about 2 million people."

Suggested Outline for Project #2.

(1) Introduction: Pretty much unchanged from Part1, except that you should now state that you are going to have two separate predictions of the future population: A natural growth prediction and a linear prediction. Of course, any grammatical or spelling errors from Part1 should be corrected.

(2) The original natural growth model. Be certain that you have corrected any errors in Part1.

(3) The optimal natural growth model. Be certain that you have corrected any errors in Part1.

(4) The original linear growth model. This paragraph should include an

explanation of how you computed the initial values of ***b*** and ***m*** in the model **P(t)=b+mt**.

Reference should be made to the spreadsheet graph of this model. In addition the original model should be displayed in boldface on a line of its own.

(5) The optimal linear growth model. This paragraph should include an explanation of how you computed the optimal values of ***b*** and ***m*** in the model **P(t)=b+mt**.

In particular, as with the natural growth model you should include a table of the values of b and m tried and the corresponding average errors. Reference should be made to the spreadsheet graph of this model. In addition the optimal model should be displayed in boldface on a line of its own.

(6) Conclusion: Pretty much as before, except that now you should have a table giving both predictions. Also you must state the predictions for which you have the most confidence and why.

(7) You must include all 4 spreadsheets. These may be separate at the end, or imbedded into the text.

APPENDIX D

EXIT INTERVIEW GUIDE

1. Use your spreadsheet to apply a linear model to the data for Chicago, Illinois. Discuss your results.

   a) What happens to the graph when you change *b*? *m*?

   b) What role does the graph play in finding the optimal linear model?

2. Explain the linear modeling process and its purpose.

3. Apply a linear model to the following data* using your spreadsheet template. Discuss your results.

If a person dives into cold water, a neural reflex response automatically shuts off blood circulation to the skin and muscles and reduces the pulse rate. A medical research team conducted an experiment using a group of ten 2-year-olds. A child's face was placed momentarily in cold water, and the corresponding reduction in pulse rate was recorded. The data for the average reduction in heart rate for each temperature are summarized in the table.

Water Temperature (°F)	Pulse Rate Reduction
50	15
55	13
60	10
65	6
70	2

*Problem 23, p. 1020. In Barnett, R. A., & Ziegler, M. R. (1996). College mathematics for business, economics, life sciences, and social sciences (7th ed.). Upper Saddle River: Prentice Hall.

4. Apply a linear model to the above data without using your spreadsheet template. Discuss your results.

5. How would you apply a linear model to data without using the spreadsheet template?

6. How would your results and interpretations be different?

7. How does the template aid the linear modeling process?

8. How does the template hinder the linear modeling process?

9. When is a linear model a “good” representation of data and a “good” predictor of future data?

10. For which other types of problems could the linear modeling process be applied?

APPENDIX E

LINEAR MODELING: HOMEWORK PROBLEMS FROM SECTION 2.1

11. City A had a population of 35,500 on January 1, 1985 and it was growing at the rate of 1700 people per year. Assuming that this annual rate of change in the population of City A continues, find (a) its population on January 1, 2000 and (b) the month of the calendar year in which its population will hit 85 thousand.

12. City B had a population of 375 thousand on January 1, 1992 and it was growing at the rate of 9250 people per year. Assuming that this annual rate of change in the population of City B continues, find (a) its population on January 1, 2000 and (b) the month of the calendar year in which its population will hit 600 thousand.

13. City C had a population of 45,325 on January 1, 1985 and a population of 50,785 on January 1, 1990. Assuming that this annual rate of change in the population of City C continues, find (a) its population on January 1, 2000 and (b) the month of the calendar year in which its population will hit 75 thousand.

14. City D had a population of 428 thousand on January 1, 1992 and a population of 455 thousand on January 1, 1997. Assuming that this annual rate of change in the population of City B continues, find (a) its population on January 1, 2000 and (b) the month of the calendar year in which its population will hit 600 thousand.

15. Find the month of the calendar year during which Cities A and C (of Problems 11 and 13) have the same population.

16. Find the month of the calendar year during which Cities B and D (of Problems 12 and 14) have the same population.

17. On January 1, 1995 City E had a population of 100 thousand and it was increasing linearly at the rate of 3,000 persons per year. At the same time City F had a population of 75 thousand and it was growing naturally with an annual growth rate of 4%. Assuming that the linear growth of City E and the natural growth of City F continue, on what calendar date will they have the same population?

18. On January 1, 1995 City G had a population of 100 thousand and it was increasing linearly at the rate of 5,000 persons per year. On that date City H had the same population but was growing naturally with an annual growth rate of 5%. Assuming that the linear growth of City G and the natural growth of City H continue, on what calendar date will the population of City H be double that of City G?