The Pennsylvania State University

The Graduate School

College of Education

MEASURING KNOWLEDGE OF MATHEMATICAL FUNCTIONS:

VALIDITY OF SCORES AND PROFILES OF PARTICIPANTS

A Dissertation in

Educational Psychology

by

Kelli Higley

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2009

The dissertation of Kelli Higley was reviewed and approved* by the following:

Jonna M. Kulikowich
Professor of Education
Dissertation Advisor
Chair of Committee

P. Karen Murphy
Professor of Education

Peggy Van Meter
Associate Professor of Educational Psychology

Dana L. Mitra
Assistant Professor of Educational Theory and Policy

Rayne Sperling
Associate Professor of Educational Psychology
Program Chair - Educational Psychology

*Signatures are on file in the Graduate School

ABSTRACT

Knowledge of mathematical functions is expected of university students. This knowledge was measured in this dissertation using two previously untested instruments. The aim of this dissertation was to validate the scores from these measures. The first instrument measured function knowledge in its declarative, procedural, or conditional form. The second instrument measured student ability to translate among representations of linear and quadratic functions. These instruments were intended to measure knowledge of functions in an efficient and effective manner such that deficits of knowledge could be clearly determined, along with clear directions for improving that knowledge.

As an essential correlate with knowledge, beliefs about mathematics were also measured, using the Conceptions of Mathematics Scale (Crawford et al., 1998b) and a modified Teachers Epistemic Beliefs Instrument (Hennessey, 2007).

The instruments were administered to 640 undergraduate students. The scores on the instruments were submitted to exploratory and confirmatory factor analyses in order to ascertain the structure of the scores; then profiles were established through latent class analysis. The instruments did not measure knowledge as expected. Instead the results indicate that mathematical knowledge of functions must be understood through the lens of representations, specifically their overall format of visual or symbolic representations. The beliefs measures did not substantially add to understanding of student proficiency in mathematics. Four profiles were identified, based primarily on total scores. The profiles provided a possible theory for acquisition of knowledge of functions. Limitations and directions for future research are considered.

TABLE OF CONTENTS

List of Tables vii
List of Figures viii
Acknowledgements ix

INTRODUCTION 1
Background for Instruments 3
Statement of Problem 7
Purpose of Study 8
Limitations 10
Definition of Key Terms 10

REVIEW OF THE LITERATURE 13
Measuring Student Knowledge of Function 15
Knowledge of Function Definition 15
Knowledge of Use of Functions 19
Summary: Measurement of Student Knowledge of Function 22
Types of Knowledge 23
Knowledge Definitions 24
Conceptual knowledge 24
Declarative knowledge 25
Procedural knowledge 25
Conditional knowledge 26
Knowledge Sets in Research 27
Mathematics education researchers 27
Anderson's Adaptive Control of Thought 30
Alexander et al 31
Summary of Types of Knowledge 32
Representations 34
Representations in Mathematics 34
Using Multiple Representations 37
Translation of Functions 40
Summary: Representations 42
Beliefs about the Nature of Mathematics 43
Statistical Analysis 47
Classification of Items 47
Exploratory factor analysis 48
Confirmatory factor analysis 50
Classification of Participants 52
Cluster analysis 52
Latent class analysis. 54
Classification in the Literature 57
Achievement. 57
Beliefs. 59

Summary of Statistical Analysis 61
Summary 62

METHODS 65
Participants 66
Measures 69
Basic Knowledge of Functions Test 69
Representational Transfer Test 71
Conception of Mathematics Scale 72
Pilot 73
Procedures 75
Recruitment and Data Collection 75
Data Entry 76
Splitting the sample 77
Data Analysis 78
Research Questions 1 and 2 78
Content validity 78
Factor analysis 79
Reliability 81
Research Question 3 81
Research Question 4 82
Research Question 5 82
Research Question 6 83

RESULTS AND DISCUSSION 84
Structure of the Instrument Scores 85
Research Questions 1 and 2 85
Content Experts 86
Parallel Items in the RTT 88
Exploratory Factor Analyses 90
Basic Knowledge of Functions Test (BKFT) 90
Representational Transfer Test (RTT) 95
BKFT revision 97
Discussion of Exploratory Factor Analyses 98
Confirmatory Factor Analyses 98
Basic Knowledge of Function Test (BKFT) 99
Representational Transfer Test (RTT) 102
Discussion of Latent Variables from BKFT and RTT 105
Basic Knowledge of Functions Test (BKFT) 106
Representational Transfer Test (RTT) 108
Summary of BKFT and RTT Measures 108
Research Question 3 109
Factor Analyses 110
Discussion of CMS 112
Research Question 4 113

Student Profiles....114
Research Question 5 and 6....114
Latent Class Analysis....114
Description of Classes....115
Relations with Other Variables....118
Discussion of Classes....120

CONCLUSIONS, IMPLICATIONS AND LIMITATIONS....122
Instrumentation and Procedure....123
Factors and Classes....126
Latent Factor Results....126
Latent Factor Discussion....128
Belief in Mathematics....130
Latent Class Results....133
Latent Class Discussion....134
Limitations....135
Future Directions....137
Conclusion....139

References....142
Appendix A: Basic Knowledge of Functions Test....156
Appendix B: Representational Transfer Test....165
Appendix C: Demographic Sheet....178
Appendix D: Demographics of Exploratory and Confirmatory Samples....179
Appendix E: Mean Scores (*SD*) on Measures Taken and External Criteria....180
Appendix G: Demographics by Latent Classes....181

List of Tables

Table 1. Majors by College 67

Table 2. Mathematics Courses Taken 68

Table 3. Comparison of Means after Replacing Missing Data 77

Table 4. Scale for Expert Analysis 87

Table 5. Content Expert Ratings by Parallel Items of RTT 89

Table 6. Fit Statistics for EFA Factor Solution Models for BKFT 92

Table 7. Item Descriptions for BKFT by Factor 93

Table 8. Fit Statistics for EFA Factor Solution Models for RTT 95

Table 9. Item Descriptions for RTT by Factor 97

Table 10. BKFT Models Run for Verification of Scores 100

Table 11. BKFT Components 101

Table 12. Factor and Component Scores for BKFT 102

Table 13. RTT Models Run for Verification of Scores 103

Table 14. RTT Components 104

Table 15. Factor and Component Scores for RTT 105

Table 16. Unstandardized Structure Coefficients for CFA of CMS 111

Table 17. Correlations among Latent Variables 113

Table 18. Fit of Solutions with 1 – 4 Classes 116

Table 19. Factor Scores by Class 116

Table 20. Univariate Results for Class Differences across Factor Scores 117

Table 21. Means and Standard Deviation Scores for External Measures by Class 119

List of Figures

Figure 1. Examples and Items by Knowledge Type from the BKFT 70
Figure 2. Example from the Representational Transfer Test 72
Figure 3. Examples from the Conceptions of Mathematics Scale 73
Figure 4. Pilot Items Missed by Over 50% of Participants.. 75
Figure 5. Items with Three Point Content Expert Disagreement 88
Figure 6. Scree plot of BKFT eigenvalues greater than 1... 91
Figure 7. Scree Plot of RTT Eigenvalues Greater Than 1... 95
Figure 8. BKFT Model From CFA Using Composite Scores 102
Figure 9. RTT Model from CFA Using Composite Scores.. 104
Figure 10. Percentage Scores by Class on Four Factors from BKFT and RTT 118

Acknowledgements

There are people I call the “angels of my dissertation,” people without whom this document would have never happened. First, my advisor Jonna Kulikowich. She captured my attention from my first days at UConn, and has continued taking my life down paths I never expected, whether that be earning a PhD, moving to central Pennsylvania, or any of the amazing things that have happened to me since we met. Always expecting excellence from me, I hope to not disappoint.

Then of course, my committee members, three strong, intelligent scholars who have molded me into the researcher I am today. Karen Murphy, who gave friendship a chance though not deserved; Peggy Van Meter, who always thinks I know more than I do; and Dana Mitra who exposed me to the qualitative side of educational research - I think about *everything* differently because of your voices that live in my head. Thank you.

My parents, family, and friends who believed in me every step of the way and have listened and celebrated with every milestone, no matter how many milestones there were. Specifically I am deeply grateful for my mom, my mother-in-law, and two dear friends who cared for Eleanor so I could write.

My husband Michael who never doubted I could do this, and do it well. When I wanted to quit, and found good reasons, he reminded me not only that I could do this, but that I should. His steadiness and joy in me ground me. And although she was nothing but a hindrance to my progress, I’m grateful for my dissertation baby, Eleanor, whose personality and temperament were exactly right for this accomplishment. It is her sweet self that got me to finish; she is exactly right for me.

CHAPTER 1

INTRODUCTION

Functions serve “a central and unifying role in mathematics” (O’Callaghan,1998, p. 23) and as such, a solid and appropriate understanding of the mathematical idea of function is key in the trajectory towards mathematical competence for a college graduate (Eisenberg, 1992; Miller, 2005). Although many students are accepted to college and get a good education without a solid knowledge of mathematics, the process will be easier, and more doors will be open to them with this competency (National Council of Teachers of Mathematics [NCTM], 2000; National Research Council [NRC], 1989, 2001). Perhaps more importantly, mathematics is interesting and enriching, and can be bettered valued with a moderate knowledge of the field (Wu, 1996).

No matter what field of study a student pursues, the liberal arts tradition in college ensures that there will be at least some mathematics which includes functions in one’s college career. Knowledge of functions is particularly important for those students pursuing degrees in science, technology, engineering, or mathematics fields (the STEM fields) so they understand the advanced mathematics they will be expected to use on a regular basis. In addition, many students who have academic ambitions beyond the undergraduate degree also need this understanding as they likely will be involved in collecting and modeling data of phenomena. The basic concept of finding how one variable can be mathematically defined as a function of another is a common theme of much research.

No absolute definition of function exists, but all definitions contain the basic idea of a correspondence between two sets. A member of a first set is related by a specific rule to a unique member of the second set. An excellent general definition is articulated on the online encyclopedia Wikipedia: “The mathematical concept of a function expresses dependence between two quantities, one of which is known and the other produced. A function associates a single output to each input element drawn from a fixed set” (Wikipedia, 2009). This can be compared with the typical definition of a college textbook: “Basically, a *function f* relates each element *x* of a set … with exactly one element *y* of another set” (Gill, n.d., p. 2). The central element of a fuction is the specific relationship or association. As functions define one quantity in terms of another, they can be used for description, prediction, and understanding in mathematics. Algebra, geometry, and calculus, as well as many other subfields in mathematics all utilize the function concept.

Although national organizations and researchers in mathematics promote the importance of student knowledge of function, study of the topic has limitations. Primarily this research has been in the context of measuring student definition of function (e.g. Dubinksy & Harel, 1992; Vinner, 1983), comparing teaching methods (e.g. O’Callaghan, 1998), or in the context of representation (e.g. Even, 1998). While these provide important insights to student knowledge, alone they are insufficient to describe whether or not students know functions, and if they do not, what deficits exist. A framework of what that knowledge should look like is essential to appropriately measure it. That framework is not evident. Good measurement of student knowledge is necessary to identify and remedy possible deficiencies.

In addition to lacking a theoretical framework, measurement of knowledge of function is often conducted using interviews, or open-ended questionnaires or instruments. While this is an excellent way to deeply understand the nature and nuances of this knowledge, it is time consuming to evaluate. Given what is already known about function knowledge, a measurement with efficient administration and evaluation can and should be used. This would allow the measure to serve not only in descriptive purposes for better understanding knowledge of functions, but it could also be used as a diagnostic measure to aid the students whose knowledge it describes.

Measuring and describing student knowledge of functions with instruments designed to fill these gaps is one of the main goals of this dissertation. This measurement will be done with two as yet untested instruments, one which uses the theoretical framework of declarative, procedural, and conditional knowledge of functions, and a second that measures ability to translate from one representation to another. The primary purpose of this dissertation is to apply these instruments and validate their design.

Background for Instruments

The term “to know” is broadly and generally defined; actual measurement of knowledge requires definitional precision (Alexander, Shallert, & Hare, 1991). Thus in this study knowledge is classified into two major categories, declarative and procedural (e.g. Anderson, 1996; Ryle, 1958; Ullman, 2004). Declarative knowledge refers to facts and ideas about a particular concept. It is knowledge that can be stated. It can range from simple facts about a subject (knowing that; Ryle, 1958) to the complex webs of information and connections that are the hallmark of an expert in a field (Hiebert & Lefevre, 1986). It is often considered one’s propositional knowledge that can be stated. For this study, declarative knowledge measures the facts, theorems, and definitions of mathematics.

Procedural knowledge is knowing how to do something (Ryle, 1958), such as mail a letter or solve a familiar form of a math problem. Procedural knowledge has been defined as "the ability to execute action sequences to solve problems" (Rittle-Johnson, Siegler, & Alibali, 2001, p. 346). This type of knowledge in mathematics is usually associated with knowledge of particular algorithms and connected to specific classes of problems. This knowledge is sometimes described as rote learning and can imply a lack of depth of understanding (Hiebert & Leferve, 1986) although many researchers recognize the existence and need for the deeper procedural knowledge that is one of the hallmarks of mathematical competence (Baroody, Feil, & Johnson, 2007; Star, 2005). Procedural knowledge in this study is the execution of algorithms and calculations.

An additional category of knowledge that is also included in this study is conditional knowledge. In Alexander et al.'s (1991) review of ways of knowing, conditional knowledge is described in association with declarative and procedural knowledge. This is knowing when to apply the declarative and procedural knowledge one has. The product of conditional knowledge appears to be either declarative or procedural knowledge, because it is the knowledge of when to use those types of knowledge. This way of knowing is represented in this study through application and modeling problems, in which the type and topic of knowledge needed for to solve the problem are not explicit in the question.

Although this framework of declarative and procedural knowledge is well-accepted in educational psychology research on knowledge, this framework was not found in studies of mathematical knowledge. The probable reason for this is that mathematics education research typically defines knowledge as being either *conceptual* or procedural. A representative definition of conceptual knowledge is “an integrated and functional grasp of mathematical ideas” (NRC, 2001, p. 118). From this definition, conceptual knowledge is the connected whole of declarative, procedural, and conditional knowledge. Yet studies of knowledge in mathematics do not consider the building blocks that form conceptual knowledge, just whether or not the completed structure exists.

The analysis of students’ declarative, procedural, and conditional knowledge of functions is important, but must be augmented by a consideration of student facility in using representations. Knowledge of functions cannot be considered without an investigation of representations (Duval, 2006). In the context of this study, representations are the varied external symbol systems that can be used to express information (Janvier, 1987). While the most common representation used in education is verbal (a description of the idea), students are soon introduced to charts, manipulatives, diagrams, graphs, pictures, and a myriad of other ways to represent information. Students are assumed to form connections among these representations, but this is often not true.

Initially, additional representations tend to be informationally similar, such that the intended information can be gleaned from either representation (Ainsworth, 1999). Because of this duplicity of information, students are not required to understand both representations and can often operate within a preferred format. Yet the value of having multiple representations is that each type allows for clarity in communicating certain aspects of the represented idea. That clarity comes at a cost though, because other aspects will be inaccessible through that form of representation (Leinhardt, Zaslavsky, & Stein, 1990; Tabachneck-Schijf, Leonardo, & Simon, 1997). Thus, students who favor one representation without understanding others will often not have the best tools available to learn from the multiple representations that are presented.

Specifically in mathematics, functions are one of the first situations within mathematics in which multiple representations are necessary, not just helpful, for understanding the concept. Functions can and must be represented in a variety of systems, such as algorithmic, graphical, tabular, and verbal (sometimes referred to as the 'rule of four'; e.g., Hughes-Hallet et al., 2002, Janvier, 1987). These various representations can be used for different purposes. Thus, depending on which representation the function is presented in, a student would be find some aspects of the function easier to recognize than others. Research shows that some knowledge about functions is representation specific (Dufour-Janvier, Bednarz, & Belanger, 1986; Gagatsis & Shiakalli, 2004). For example, while a student may be able to locate the zeros of a graphically presented function, he or she may not be able to do the same with an algorithmically based function. Provided with this same function in a table, or verbally, that information may not be available at all, unless the student is able to shift the function into an alternate representation.

In the context of functions, it is particularly important that students be able to translate from one representation to another. This translation ability is a hallmark of two important understandings in mathematics. First, it is evidence that students can access multiple representations, and the additional computation and solving tools that provides them. Second, translation ability is a necessary, but not sufficient, indicator that students are beginning to comprehend functions as the relationships that are depicted in the representations, as opposed to the representations themselves. "Transfer skills are prerequisite to the integration of information about functions into a single, unified conceptual image" (O'Callaghan, 1998, p. 23).

Thus, not only does a student need to know *about* functions (declarative knowledge), how to *use* functions (procedural knowledge), and *when* to use this knowledge (conditional), he or she must also know how to move from one type of representation to another. These ways of knowing are not unique to the study of functions, but could also be used to describe knowledge of many fields in business, sciences, engineering, as well as fields that participate in empirical research, in short, fields that use mathematics.

An additional factor that may influence student mathematical behavior is their beliefs on the nature of mathematics (Schoenfeld, 1989). Beliefs influence actions (Mandler, 1989) and a sense of student beliefs can provide insight as to how a student approaches their study of mathematics and connected understanding of the field.

Statement of Problem

Functions are described as a foundation to advanced mathematical studies, but measurement of student knowledge of functions has various inadequacies. First, it is not studied directly. Studies that focus on this knowledge in students measure their knowledge of the definition, measure learning within a comparison of teaching methods, or as a vehicle to study representations. Student ability to use functions, and know about functions (as opposed to defining functions) was not found as a specific research topic. Another concern is the way it is measured, which has been primarily for descriptive purposes. Description is an important purpose of research, but given the consequence of this topic and the previous research conducted, diagnostic tools are necessary. This would involve measurement with an appropriate instrument measuring essential knowledge that can be easily administered, evaluated, and interpreted. Finally, while functions should be studied as a knowledge base on their own, the connection with representations should not be overlooked. Consequently, functions should be studied in parallel with representations.

Purpose of Study

The first purpose of this research is to validate the design of two instruments purported to measure student knowledge of function, but have not yet been tested empirically. These instruments measure overall declarative, procedural, and conditional knowledge of functions, as well as specific knowledge of translation between function representations. These measures were intended to serve as diagnostic tools for understanding student knowledge, and should thus be efficient for administration, evaluation, and interpretation. While the instruments are intended for use at the high school or university level, they are tested here using university students.

The first instrument was designed to measure knowledge declaratively, procedurally, and conditionally; ways of knowing that have been shown to produce positive learning outcomes (Alexander et al., 1991). By studying knowledge as component pieces, strengths and weaknesses in student knowledge can be directly identified. The second measure was designed to measure students' abilities to translate between various representations of functions, a key skill in learning to understand and apply functions (Elia & Gagatsis, 2008). In addition, a measurement on beliefs about the nature of mathematics was administered. Factor analysis was used to identify latent variables in the measurements in order to verify that the design objectives were met.

The second purpose of this research was accomplished when the factors resulting from the instruments were utilized to characterize the nature of current function knowledge in a sample of university students. Once the structure of the instruments was determined, profiles describing student knowledge and beliefs were created using latent class analysis. Knowing how these variables worked together enhances the understanding of what it means to know in fields that are so dependent on mathematics and representations in their communication and demonstration of ideas.

The purpose of this study was to test instruments designed to provide information about university students' knowledge of functions, which would serve as diagnostic tools for measuring this important information. Evidence for the validity and reliability of the scores resulting from the tests was gathered. Finally, student learning about functions was described using the information resulting from the scales. With such purposes in mind, the research questions for this study are as follows:

1) To what extent does the factor structure of the Basic Function Knowledge Test (BFKT) support the validity and reliability of scores?
2) To what extent does the factor structure of the Representational Transfer Test (RTT) support the validity and reliability of scores?
3) Does the factor structure of the Conceptions of Mathematics Scale (CMS) match that of the literature? If not, what structure does seem to exist? Are these scores reliable?
4) What relationships exist between the scales from the BFKT, RTT, and CMS?
5) Given the scales of the three tests, can scale scores be used to form profiles of student knowledge of functions?

6) To what degree are profiles related to important academic performance criteria such as SAT mathematics scores and grades in mathematics courses?

Limitations

Although it is expected that the information resulting from this study will inform understanding about student knowledge of functions in many ways, there are also limitations to this study. The first is that the sample comes from a large-northeastern land grant university that is highly ranked for academic achievement. Whether the results from this population are generalizable to other populations remains an empirical question, and efforts should be made in future research to access other populations. In addition, the concept of functions is vast and varied. Providing a test that measures all aspects of knowing about functions would be prohibitively long for research or classroom purposes. While selection of topics to be measured was made with careful consideration, some mathematicians may feel that other information is more central to the concept, so this may not be the knowledge that all mathematicians would measure.

Definition of Key Terms

Class – Substantively meaningful groups of people that are similar in their responses to measured variables or growth trajectories (Nylund, Asparouhov, & Muthén, 2007, p.536).

Conceptual knowledge – “An integrated and functional grasp of mathematical ideas” (NRC, 2001, p. 118).

Conditional knowledge – Knowledge of when to apply other declarative and procedural knowledge one may possess (Alexander et al., 1991)

Declarative knowledge – "Knowledge that can be declared, usually in words" (Farnham-Diggory, 1994, p. 468).

Factor analysis – Factor analysis is a "process of discovering and defining latent variables and a measurement model that can provide the basis for a causal analysis or relations among the latent variables" (Loehlin, 2004, p. 152). It can be either exploratory to predict factors or confirmatory to corroborate relations.

Function – "Basically, a *function f* relates each element *x* of a set … with exactly one element *y* of another set" (Gill, n.d., p. 2).

Latent class analysis – An analysis in which objects are assumed to belong to one of a set of *K* latent classes, with the number of classes and their sizes not known a priori. Objects belonging to the same class are similar with respect to the observed variables (Vermunt & Magidson, 2002).

Latent variable – Unobserved variables hypothesized to predict the observed variables in factor analysis; also referred to as a factor (Loehlin, 2004).

Procedural knowledge – "The ability to execute action sequences to solve problems" (Rittle-Johnson et al., 2001, p. 346).

Representation – "A representation has two components: (1) a format for recording and presenting information and (2) processes for using and modifying information" (Tabachneck-Schijf, et al., 1997, p. 307). In mathematics, what is being represented is typically an abstract idea (Duval, 2006).

Symbolic system – A superordinate groups of mathematical representations, consisting of algebraic (and other) symbols as well as verbal descriptions, also called *semiotic system* (Duval, 2006).

Translation – “Translation ability refers to the psychological processes involved in going from one mode of representation to another” (Gagatsis & Shiakalli, 2004, p. 645.

Visual system – A superordinate group of mathematical representations, consisting of visual representations such as graphs and tables (Duval, 2006).

CHAPTER 2

REVIEW OF THE LITERATURE

The study of student knowledge about functions is important because of the central role of functions to further study of mathematics (NCTM, 2000) and the pervasive nature of misconceptions about functions (Leinhardt et al., 1990). In addition, functions are learned and understood using both algebraic and graphical systems, which need to be mastered both individually and in tandem in order to truly know functions (Leinhardt et al.). Functions are one of the first points in mathematical learning where mastery of multiple symbolic systems is necessary in order to know the topic (Janvier, 1987). Learning about the use of representations in functions enables understanding in other fields that use multiple representation systems. The study of knowledge of functions requires appropriate measurement tools.

The purpose of this research study is to validate the scores of theoretically-based instruments that would allow the study of knowledge of functions for students who should have a basic knowledge of functions. This includes students who have taken courses that would expose them to functions, specifically for this study, university students. The data from these instruments will be used to describe the current state of this knowledge. The purpose of this chapter is to review the literature that undergirds the central constructs for this study.

Various criteria were used in the location and inclusion of studies for this literature review. Within each major section (i.e., functions, knowledge, representations, and analysis) well-known studies, researchers, and acknowledged references were used as the foundation. In addition, directed searches were conducted with major search engines (e.g., PsychINFO, JSTOR) for studies of declarative, conceptual, procedural, and conditional knowledge. From the search results, studies pertaining to mathematics or science were reviewed. Most studies before the mid-1980s were not considered. Both theoretical and empirical studies were reviewed. Concerned about the lack of current studies on functions, a hand search of two major mathematics education journals, Journal for Research in Mathematics Education and Educational Studies in Mathematics, was undertaken. While these three methods provided an important base of articles for the review, another valuable source of articles was forward and backward reference searches, reading articles that were cited by and articles that cite the articles already included.

The review begins with a section outlining how knowledge of functions and other advanced mathematics has been previously measured. This is followed by three sections on the nature of the instruments. The three measures employed in this study addressed knowledge of topic measured declaratively and procedurally, knowledge of representation transfer, and beliefs about the nature of mathematics. Specifically, the first of these sections presents knowledge as defined by those who research it, with specific attention paid to knowing in mathematics. This leads to a section on knowing about representations that focuses specifically on translating between multiple representations. There is some overlap amidst the sections because there are not clear distinctions between functional knowledge and representational knowledge. Yet these two sections are distinguished because there are such different literature bases that led to these measures, even if many of the findings overlap. In addition there is a section on beliefs in mathematics, and the value of measuring not only what students know but also their beliefs on the nature of mathematics.

This chapter concludes with a section on statistical techniques for profiling knowledge, with attention given to factor analysis and latent class analysis. Included in this section are results from prior studies that have classified students according to their mathematical knowledge.

Measuring Student Knowledge of Function

This section presents a review of studies on knowledge of functions in mathematics. Both how this knowledge has been measured and the findings on function knowledge are provided. These findings are organized around two major aspects of function knowledge. The first focus of research was student concept of function and how this related to both their individual and the actual definition of function. The second focus was learning as related to various teaching methods. Participants in all these studies are primarily students in high school, university, or their teachers.

Knowledge of Function Definition

Vinner (1983; Vinner & Dreyfus, 1989) researched student "sense of functions" to explore how students defined functions, both in practice and in words. Participants in his studies were presented with a series of graphs and descriptions and were instructed to identify whether or not the object was a function, and why. The final question was to provide a definition of function in their own words.

Participants in the first study (Vinner, 1983) were students in the 10th and 11th grades (n = 146) who had taken a formal unit on functions during their 10th grade year. Almost 40% of students provided a textbook definition for the final question, and many of the other students provided definitions that had elements of the textbook definition in them. Despite the strong tendency to report the textbook definition, only 34% of those students classified functions according to the textbook definition they had provided. Vinner claimed that classifications were made based on an internal belief of what a function is, which Vinner referred to as a concept image. He suggested that a concept image was not closely associated with the formal definition.

A concept image is defined as the depiction of a concept that is held in one's mind (Vinner, 1983). It contains a combination of images, formulas, or ideas, and is developed through personal experiences with the concept. While some elementary concepts tend to be formed solely by image (such as the color “orange”) and defy a precise verbal definition, other concepts, especially those taught in school, are often associated with a formal definition as well. Vinner referred to the formal definition as the concept definition. The concept definition can be acquired from direct instruction, or it can be the result of an individual making explicit a previously implicit understanding of a concept. From the results of his study, Vinner postulated that the concept image and definition seemed to be stored separately in the brain, and these two ideas may or may not be connected. He assumed that participants in his study who used the textbook definition to determine functions (thus determined them correctly) had made connections between their image and definition.

Although students are provided with a definition when a concept is first taught at school, Vinner (1983) suggested that concepts are not applied using the definition, but instead are applied through the concept image. Concept definitions, if learned, initially remain inactive, except in such cases in which a student is asked to provide the definition. Vinner maintained that teachers either dismiss students’ prior experiences with a concept, or expect that students, once they have been exposed to a definition, will connect that to their image, and revise the image as needed. They expect that future learning with the concept should be done through the lens of the definition, with consultation to the image as needed. Vinner (1983) theorized that what actually happens is that work is done through the image, and the definition is consulted when students must provide it.

These assumptions were supported when Vinner and Dreyfus (1989) conducted a similar study using college students (n = 271) and some middle school teachers (n = 36) as their participants. The study provided six mathematical entities, which participants labeled as functions or not. This was intended to measure their concept image. The final question asked the participants to define a function in their own opinion, which was considered their function definition. They found similar results to Vinner's original study (1983). The definitions provided by participants did not match their concept image for 56% of their participants. They found though, that the students in fields that required more mathematics knowledge tended to have a more connected definition and image. The authors concluded that increased experience with functions in their varied forms was needed to have a good correspondence between the image and definition.

Studies after that of Vinner (1983) often followed similar methods, in which students were expected to appropriately classify examples and non-examples of functions. Researchers used these classifications to describe students' knowledge of functions. For example, Dubinsky and Harel (1992) looked at function knowledge in students taking a discrete mathematics course that was intended to increase the students' knowledge of functions. They interviewed half the students (n = 13) after the course was completed. Students were provided 24 descriptions of situations and asked if they could be described by functions. From this data, the authors found that even after a taking constructivist course intended to promote accurate recognition of functions, students were overly restrictive to what could be identified as a function.

To be classified as a function, most students required the situation to be an expression that could be evaluated, that used numbers, and was continuous. To students, these restrictions were part of the definition of a function. Not only did Dubinsky and Harel (1992) find many misconceptions about functions, they also found that students had a hard time explaining their beliefs about functions or using their beliefs in a consistent manner. As found by Vinner and Dreyfus (1989), the quality of their answers and justifications was related to their mathematical background.

Research on students' conception of functions was primarily conducted with small sample studies using open-ended questionnaires or interviews. From this research a set of robust findings describing student misconceptions about functions emerged. Students consistently showed that they intuitively understood that functions expressed a relation between two variables. But the formal definition encompasses many more types of relationships. Hence, functions that met these intuitive criteria were readily recognized, but the many functions that did not match the intuitive definition were dismissed by students as not being functions (Leinhardt et al., 1990). For example, students believed that functions must be continuous and defined at all points (Sierpenska, 1992). They expected functions to be orderly, continuous, and familiar (Ferrini-Mundy & Graham, 1994). Students tended to reject functions with split domains (Sfard, 1992). Students saw functions as processes, not objects; they expected functions to be something that required them to solve or calculate (Sfard).

Research on functions focused on the coherence of the students' conception of function and the actual definition. The results indicated that participants' sense of what constituted a function was characterized by misconceptions. In particular, participants were not able to correctly distinguish between examples and non-examples of functions. The discrepancy between concept image and concept definition (Vinner, 1983) could be seen as an example of the knowledge and belief learning espoused by Murphy (2007). She postulated that although students know what is explicitly taught in the classroom, the learning that they will apply and use is the learning that they believe, which comes from their own experiences. According to her theories, teachers not only need to teach the definition, but also must engage students in seeing how their personally constructed definitions are not sufficiently inclusive. Only with teaching methods that encourage engagement and interaction of the newly presented information with their prior held beliefs will students truly connect the ideas into examined understanding that will enable an appropriate application of the function definition. Otherwise, students will use their personal conceptions, which may not be entirely correct, in their practice of mathematics.

Students' sense of functions is an important part of describing their knowledge and point to possible misconceptions in the application of functions in mathematical situations. The next section presents an overview of studies that investigate a more comprehensive picture of knowledge of functions, measuring functions used in calculations, interpretations, and application.

Knowledge of Use of Functions

Few studies were located that study knowledge of functions as encountered in the classroom, as opposed to definitions. The two studies located were designed as comparison studies of teaching methods. Huntley, Rasmussen, Villarubi, Sangtong, and Fey (2000) compared traditional and reform teaching in high school students. The two teaching conditions in the study were a *Standards*-based curriculum and a traditional curriculum. The *Standards*-based (NCTM, 2000) teaching emphasized conceptual learning and was characterized by only requiring the students to learn how to plan and interpret the math involved in algebra, because the tools of calculators and computers could do the calculations faster and with less chance of error. The classes in this condition were from six schools in different regions of the U.S. that were piloting a new three year curriculum. Between 90 – 180 students participated at each site, for a total of almost 600 students in either the control or experimental group. Given the diversity of location, there was little control over the fidelity of the implementation of the curriculum, and it was put into practice in various ways (e.g., curriculum used for more advanced students, struggling students, or heterogeneous groupings). The traditional classrooms tended to follow the typical curriculum for high school math, which were assumed to focus more on procedural skills.

After completing three years of high school math in either a reform or traditional teaching situation, students completed assessments measuring students' understanding, skill, and problem solving abilities in algebra and functions. Instruments had four "super problems" which were sets of questions contextualized to a problem situation (aligned with the aims of the reform curriculum). Students also solved 28 non-contextualized problems without a calculator (presumed to be aligned with the skills learned in the traditional classes). The *Standards*-based classes did better on the first, conceptual portion of the assessments, and the traditional classes did better on the second, procedural portion of the assessments. Despite these relative strengths though, students in both reform and traditional classrooms, in reality, did poorly on the measures (mean overall scores on measures ranged from 27 - 43%). Their knowledge was much weaker than would be hoped. As had been shown throughout the earlier qualitative research, students did not know this set of mathematics as measured, even after three years of instruction.

Another study that compared teaching methods was conducted by O'Callaghan (1998). He compared a teaching method that focused on concepts, real-world problems and technology with a more traditional procedures-based course. Given the importance of functions in a students' mathematics trajectory, he studied the impact of this teaching method on student understanding of functions. O'Callaghan defined knowledge of functions as being made up of four parts: modeling, interpreting, translating, and reifying. Modeling was defined as the ability to analyze a problem or situation and produce the appropriate function in a usable format. Interpreting was being able to use a function and extract the requisite information from the representation. Translation was moving from one representation to another. Reification was the mental process that takes functions from procedures and instructions to objects that can be manipulated and acted upon in their own right – such as compositions of functions, adding functions, etc.

O'Callaghan (1998) studied college students taking algebra. Students from three classrooms were involved in the study, for a total of 109 students. Two courses used traditional instruction, while the third was taught with the reform curriculum. All students took researcher-designed pre- and posttests intended to measure their knowledge over the four elements of function knowledge, balanced over representation type. The students in the reform curriculum were better at modeling, interpreting, and translating functions in comparison to the students in the traditional courses. Both groups were equally poor at reifying functions as measured by the researcher designed instrument. Reform students performed somewhat poorer on standard operations and transformations of algebraic formulas as measured by the course final, but this could have been a result of their initial lower mathematical ability. Again, despite differences in groups, all averages were fairly low (less than 50%). Although students taught reform curriculum showed evidence of having greater knowledge of how to understand functions, students taught with traditional methods did better in solving algorithmic problems with functions, yet neither showed exceptional knowledge.

These studies show consistent evidence of a lack of knowledge of function. Although the studies by Huntley et al. (2000) and O'Callaghan (1998) show that teaching methods affect what students learn, they also seem to indicate that the understanding of function knowledge is incomplete. Functions are also studied in the context of representations. These studies will be reported in a later section of this review.

Summary: Measurement of Student Knowledge of Function

The participants in these studies were high school or college students, and all had exposure to functions through mathematics classes. The studies about general knowledge of functions demonstrated that most students, despite instruction, seemed to have a similar conception of function, one that did not match the mathematical definition (Vinner, 1983). Yet increased mathematics knowledge did relate to a better interpretation of functions (Vinner & Dreyfus, 1989). Because of these findings, students overall were thought to have a poor knowledge of functions (Dubinsky & Harel, 1992; Sfard, 1992). In addition, studies specifically measuring students' abilities to apply function knowledge in solving both contextual and non-contextual problems showed again a lack of knowledge (Huntley et al., 2000).

The majority of these studies were conducted with relatively small sample sizes, and all measured knowledge with interviews or open-ended questionnaires. This required a depth of analysis and allowed a rich description of some aspects of knowledge of function to be developed. What is lacking from this body of research is the specific study of function knowledge, not as an embedded topic of another study (e.g. teaching methods). Students' ability to work with functions has not been measured directly. In addition, the instruments used in these studies were not conducive to large-scale administration and measurement. This restricts how and when this knowledge can be measured, and what can be done with those measurements. Using instruments that cannot be scored and interpreted quickly do not allow for diagnostic measurements that could be used to help students with insufficient knowledge. A large scale measurement requires a framework in the development of the instrument. The next section provides a review of various frameworks that researchers in knowledge and mathematics have developed.

Types of Knowledge

The expected outcome of any educational experience is knowledge. Students who have graduated from high school and entered college are expected to have certain knowledge, and functions are part of that knowledge. But appropriate measurement of knowledge is only possible if the term is precisely defined and operationalized. Its definition influences research questions, measurement tools, results, and the conclusions drawn by both the researchers and those who read the results (Alexander et al., 1991;

Rittle-Johnson et al., 2001; Star, 2005). This section contains a review of the various types of knowing and learning outcomes that have been applied to mathematics.

Humans seem to instinctively recognize that people can know about things or how to do things. This self-evident division is enacted in different ways depending on the research base though. Research on knowing in mathematics focuses on two types of knowledge, conceptual and procedural (e.g., Hiebert, 1986). In the field of studying knowledge more generally, conceptual knowledge is defined as encompassing procedural knowledge, as well as declarative and conditional knowledge (e.g., Alexander et al., 1991). In this section these four major classifications of knowledge are defined. Conceptual knowledge is presented first, followed by declarative knowledge, then procedural, then conditional. Evidence of their use and application is provided. Research examples of these knowledge types, primarily within mathematics, are then considered.

Knowledge Definitions

Conceptual knowledge. Within research on mathematics learning, conceptual knowledge is broadly considered to be knowledge about the overarching concepts in math. It has been defined as a "rich network of relationships between pieces of information" (Carpenter, 1986, p. 113). It is seen as the backbone of true understanding in mathematics. It serves as a check for procedural knowledge, and represents a web of knowledge in which relations between facts are as important as the facts themselves (Hiebert & Lefevre, 1986). Conceptual knowledge is either an implicit or explicit understanding of principles and relationships between units of knowledge. The possessor of such knowledge may not be able to verbalize this understanding depending on its nature (Rittle-Johnson et al., 2001).

Given the encompassing nature of conceptual knowledge, there is no one way to measure it. Conceptual knowledge has been measured by having students attempt novel tasks, under the assumption that unfamiliar tasks require conceptual, not procedural knowledge (Rittle-Johnson et al., 2001). It has been measured through interviews of students (Harel & Dubinsky, 1992) or open-ended questionnaires (Vinner, 1983). Sometimes multiple-choice conceptual questions are asked, hoping to tap into implicit conceptual knowledge (LeFevre et al., 2006). In other studies students evaluate solution methods (Rittle-Johnson & Star, 2007) with the assumption that recognizing errors is a conceptual, not procedural, task.

Declarative knowledge. Ullman (2004, see also Gabrieli, 1998) described declarative knowledge as being associated with the learning, representation, and use of knowledge about facts and events. Declarative knowledge is based on associative learning (the pairing of two pieces of information) and this knowledge often can be acquired in single stimulus presentation. Declarative knowledge is explicit (or can be made explicit when cued) and is accessible in all learning situations, both declarative and procedural. Anderson (1996) defined declarative knowledge as a fact-based semantic chunk, such as a definition, theorem, or other piece of verbally-coded information. No research was identified in this review that directly measured declarative knowledge in mathematics. This may be because little or no attention is paid to declarative knowledge in mathematics education research.

Procedural knowledge. Anderson (1996) defines procedural knowledge as the knowledge required to take an action, whether it be cognitive or motor, which is executed when appropriate conditions have been met. Procedural knowledge is involved in learning new, or controlling already acquired, sensori-motor and cognitive skills (Ullman, 2004). Thus procedural knowledge is used in acquiring and using habits, skills, or other procedures. How procedures are enacted is generally not available for conscious access. Procedural learning is gradual, with slow improvement of skill with practice (Gabrielli, 1998).

Specifically within mathematics education research, procedural knowledge is defined as “step-by-step procedures executed in a specific sequence” (Carpenter, 1986, p. 113). Procedural knowledge is often characterized by the mechanical nature of the process, and rote learning. It is the ability to execute action sequences in order to solve a problem (Rittle-Johnson et al., 2001). The measurement of procedural knowledge is more straightforward than that of conceptual knowledge. It is the assessment of whether a participant can successfully accomplish the specified task.

Conditional knowledge. Conditional knowledge was identified in a review of how the term knowledge is used in learning research conducted by Alexander et al. (1991). It is defined as recognizing the appropriate conditions in which to apply other knowledge, or the “when” of knowing. This type of knowledge is challenging to measure though, because it is enacted as knowledge that one can state (declarative) or the knowledge one can do (procedural). This type of knowledge also has not been measured in research on mathematics learning.

Conceptual, declarative, procedural, and conditional knowledge have been defined in the preceding paragraphs. Individual understanding of these types of knowledge is valuable, but these categories of knowledge tend to appear in research in groups. The next section identifies the typical groups of knowledge and provides some findings that result from the particular framework of knowledge.

Knowledge Sets in Research

Types of knowledge are paired in various ways in research. In this section, divisions are presented, along with how that interpretation informs measurement of mathematics knowledge. Mathematics education researchers (e.g. Hiebert, 1986; Rittle-Johnson & Siegler, 1998) consider primarily conceptual and procedural knowledge. This is the richest source of information on measuring knowledge of functions so will be presented first. Then, targeted findings from Anderson's (1996) research, which focuses on declarative and procedural knowledge, is presented. This is followed by knowledge structure as reported by Alexander and colleagues (1991).

Mathematics education researchers. Conceptual knowledge is most often associated with depth of knowledge while procedural knowledge is identified with a shallower understanding (e.g., Hiebert, 1986). Recently though, some researchers have insisted that this characterization restricts the understanding of these two forms of knowledge, and the learning they represent (Star, 2005). As the depth of knowledge increases about either conceptual or procedural knowledge about a certain topic, the two forms of knowledge become intertwined. Conceptual knowledge without an understanding of the accompanying procedures cannot exist (Baroody et al., 2007). Rittle-Johnson and collegues (2001) conceptualize the two types of knowledge as ends of a continuum. At the ends they can be distinguished, but there is middle ground where the two overlap.

Rittle-Johnson and Siegler (1998) conducted a review of the literature to better understand knowledge development in mathematics and the relationship between these conceptual and procedural knowledge. Advanced mathematics is studied less than elementary mathematics (Leinhardt et al., 1990), and this is evidenced in this review. All the articles corresponding to conceptual and procedural knowledge reviewed by Rittle-Johnson and Siegler were about preschool and elementary mathematics. But some important conclusions emerged about the acquisition of these types of knowledge that may generalize. The first was that conceptual and procedural knowledge almost always worked in tandem. Students who knew procedures tended to know concepts and vice-versa. This contradicts the general conviction that students can know one without the other (Hiebert, 1986; Star, 2005). The types of knowledge were disjoint only when the concept was first taught, but as expertise grew there was an increase in both conceptual knowledge and procedural skill. The authors raised the possibility that these two types of knowledge cannot be measured independently.

Another important finding of Rittle-Johnson and Siegler (1998) was about knowledge acquisition. They found that the first type of knowledge to develop was the one that the student had the most experience with, either in daily use or in school settings. This meant that if students were exposed first to the procedures, and were well-practiced there, then procedural knowledge developed first with conceptual knowledge following. If the concepts were experienced first, then conceptual knowledge lead.

This finding was supported by the research by both Huntley et al. (2000) and O'Callaghan (1998). Some of the findings from these studies were reported in an earlier section. In both studies the experimental group were classes taught conceptually and the control groups were taught procedurally (or traditionally). On post-intervention measures of knowledge students performed better for the tests measuring the more familiar type of knowledge. Accordingly, the conceptually taught students were more able to answer conceptual questions while the procedurally taught students were better at procedural questions. Neither focus of teaching resulted in independent mastery of the other type of knowledge.

While the difference between conceptual and procedural knowledge provides a valuable framework for understanding the knowledge required to do mathematics, studying mathematics using these dimensions has complications. These classifications of knowledge were not intended to describe all mathematical knowledge (Baroody et al., 2007; Hiebert, 1986), but research in student knowledge of mathematics rarely moves outside of these two forms of knowledge. In addition, by only measuring conceptual (not declarative) knowledge a complete representation of knowledge cannot be formed. Defined this way, there is no mechanism for measuring the theorems, definitions, and facts that are the foundation of conceptual knowledge. Excellence in mathematics happens when these distinct pieces of knowledge are combined in novel ways, thus the dichotomy championed by the math educational researchers ignores some important mathematical knowledge. By looking at the big ideas, the wide lens misses some of the more detail oriented nature of mathematics. In addition, if conceptual knowledge is found to be lacking, more studies are necessary to determine if it is the knowledge pieces or connections that are missing

Conceptual knowledge is one measurement of information that can be stated. Declarative is another. At first glance conceptual knowledge and declarative knowledge appear similar in that they are contrasted with procedural knowledge and both are measures of knowledge that can be stated. The most notable difference though is in the scope of the knowledge. Conceptual knowledge is that of big ideas and overarching themes. Specific facts are only measured as part of a larger conceptual understanding. Some researchers define conceptual knowledge as including that, but do not include it for measurement purposes (NRC, 2001; Rittle-Johnson et al., 2001). Declarative knowledge is facts and chunks of verbal information. The next two organizations of knowledge use declarative knowledge as the companion to procedural knowledge.

Anderson's Adaptive Control of Thought. Through his Adaptive Control of Thought-Rational theory (ACT-R; Anderson, 1982; 1996; Anderson, Bothell, Byrne, Douglass, Lebiere, & Qin, 2004), Anderson attempts to model human cognition, specifically through computer programming. It has been used to successfully predict thinking in solving proofs (Anderson, 1983), writing algebraic expressions, tutoring high school algebra (Koedinger & Anderson, 1997) and many other mental processes. In order to model human cognition with computers, all mental processes must be reduced to one of two categories, declarative or procedural. Declarative knowledge, according to ACT-R, is a fact-based semantic chunk, such as a definition, theorem, or other piece of verbally-coded information. Declarative knowledge chunks are formed either through direct encodings from environment, or as an outcome of productions. Procedural knowledge is the encoding of actions, whether they be cognitive or motor, which are taken when appropriate conditions have been met. These two types of knowledge work together in the form of production rules.

Productions are if-then statements that inform the mind on the appropriate action. For example, a production may be “if equation is in form *x* + *number1* = *number2*, and goal is to know *x*, then find difference between numbers.” Thus, “production rules embody procedural knowledge, and their conditions and actions are defined in terms of declarative structures” (Anderson, 1996, p. 356).

While Anderson has been successful at modeling knowledge using these categories, his interpretation of them is overly fine-grained for the purpose of evaluating (not modeling) the state of student knowledge. A classification scheme that uses declarative and procedural knowledge at a more general level is that presented in the review of knowledge terms by Alexander and colleagues (1991).

Alexander et al. The review of knowledge terms by Alexander et al. (1991) describes all four knowledge types highlighted in this review of the literature. Conceptual knowledge is defined as an individual’s knowledge of ideas or concepts. It is an overarching type of knowledge that subsumes specific knowledge of a topic. Declarative, procedural, and conditional are not defined in connection with conceptual knowledge, but are included as a triad of knowledge types that could be used to identify any knowledge about a subject. Using this framework, knowledge can be organized specifically into knowing factual information about something, knowing how to execute various procedures, and knowing when those facts and procedures are applicable.

Anderson does not provide a distinct type of knowledge for knowing when. According to the definition of productions provided by Ritter, Anderson, Koedinger, and Corbett (2007), they already contain the conditions that are necessary to carry out the action, thus for Anderson, it seems that conditional knowledge is the interaction between declarative and procedural knowledge, but does not qualify as a separable type of knowledge. It may be that identifying conditional knowledge as distinct from declarative procedural knowledge is not empirically possible.

The divided ways of knowing, declarative and procedural, are common in most literature involved with the general description and understanding of learning (e.g., Anderson, 1983; 1996; Gabrieli, 1998; Ryle, 1958, Ullman, 2004). With minor variations, the basic framework is consistent. Declarative knowledge is knowledge that can be stated, while procedural knowledge is knowledge that is enacted. To repeat the common distinction, declarative knowledge is “knowing that” and procedural knowledge is “knowing how.” Clearly knowledge about mathematical functions comes in both varieties. But in addition to problems that can be nicely categorized (e.g. recognizing a definition for declarative knowledge, or solving for the variable as a procedural question), there are many mathematics activities that draw from both types of knowledge.

Summary of Types of Knowledge

In this section, definitions of knowledge types and their interactions as defined through research were reported. The consensus among various fields of cognitive research is that there is more than one way to know (e.g., Alexander et al., 1991; Anderson 1996; Hiebert, 1986; Rittle-Johnson & Siegler, 1998). While there is a consensus that we seem to 'know that,' and 'know how,' the nature of what that is, and how the two interact, and whether there is more, is not settled. Math education researchers differentiate between the big ideas of mathematical concepts and carrying out procedures. Knowledge researchers draw the distinction between things that can be stated and things that can be done. Distinctions such as this help describe the phenomena of learning and help researchers define future research.

There are problems with measuring mathematics knowledge according to either of these divisions though. First, no matter how the knowledge is split, scant attention is spent on the relation between conceptual and declarative knowledge, and the fundamentally intertwined nature of knowledge that exists at the higher level of learning that is expected in college. Declarative knowledge does not measure the connections between pieces of knowledge, and conceptual knowledge does not measure the pieces of knowledge.

Studies that actually measured mathematical knowledge in a declarative-procedural framework could not be found. Using only the conceptual-procedural division creates a measurement situation in which misunderstanding in connections or in building blocks of the conceptual framework cannot be differentiated. Measuring only declarative and procedural knowledge misses the connections that students have and need in order to demonstrate competence. Thus any measurement working under this framework would also require some sort of conditional knowledge measure as well, or items that exhibit a sense of when this information is useful and necessary. It is be possible that while these various “ways of knowing” are interesting theoretical distinctions that allow us to better conceptualize knowledge, in actual measurement of advanced math practices these divisions have no practical application.

The previous section reviewed the different types of knowledge and the frameworks that are used in studying how students know. There is a consensus that students should know *about* ideas, and they should know *how* to do things, but the nuances of those definitions vary. In studying knowledge of functions, one area that is essential to include is knowledge of representations, which is reviewed in the next section. While there is some overlap with previous sections, research on knowledge of representations draws from a different literature base than that for general knowledge of functions. This next section focuses on representations.

Representations

Representations are used to record or present information, to communicate, or to manipulate in order to demonstrate an idea or solve a problem (Larkin & Simon, 1987). While representations can be internal or external, only research pertaining to external representations is presented here. The specific focus is the ability of students to use and transfer among various external representations used with mathematical functions. First, the nature and use of representations in mathematics is overviewed. Attention is then paid to the importance and interactions between multiple representations, which is the foundation of transfer between representations. Finally specific studies about students' abilities to transfer between representations are reviewed.

Representations in Mathematics

Representations are essential in mathematics as they provide the system for communicating otherwise abstract ideas (Duval, 2006). Various representations are used to communicate mathematics, such as expressions, equations, coordinate graphs, data tables, hybrid natural language and more (Kaput, 1989). Within mathematics, the most commonly considered representations are those of the graph, equation, and table. Verbal descriptions are also representations. Specifically with functions, understanding the various representations is an essential part of knowing functions (Gagatsis & Shiakalli, 2004). It is important for students and teachers to understand the various representational forms, as well as their inherent strengths and weaknesses (Dufour-Janvier et al., 1987).

Most mathematical ideas can be displayed in more than one representation. Often representations of the same idea in different formats are intended to be informationally equivalent (Larkin & Simon, 1987), such that that the same information can be extracted from either representation. Yet within any particular format, some information is more easily accessed while other information is only implied. The value of knowing different representations is that the most effective representation can be chosen for the task (e.g. solving a problem, describing a situation, etc.). The effectiveness of a representation is based on how easy it is to extract the needed information. Effectiveness of a representation is not an objective label can be assigned to a representation; instead it is dependent on the familiarity of the user with both the type of representation and the task that is to be done (Tabachnik-Schijf et al., 1997).

Meltzer (2005) studied representational facility of students in college physics courses and found that representation type influences knowledge measurement. In his study (n = 408) he found that often students performed differently depending on problem representation. For example, some students who correctly answered the questions presented verbally would then respond incorrectly for the pictorial representation. The findings from his study suggest that expertise develops by representation, which influences students' abilities to exhibit knowledge. The possibility that knowledge can be demonstrated in one representation but not in another has serious ramifications for any measurement of student understanding.

This was also demonstrated in a study by Monk (1992). He interviewed 20 college students, who were either currently taking calculus or were senior university mathematics students. He provided each student with a physical model of a situation, then he observed his or her ability to translate it into graphical model. Questions were posed to the students about their thought process as they modeled. In this process, Monk realized that students' understanding of the concepts necessary for graphing was deficient in certain situations. He recognized that the students had only partial knowledge of concepts. They were understood well enough to accomplish some mathematical tasks, but not others. The partial knowledge became more evident when students were asked similar questions in different representations. He found that an idea that was well formed in one representation often had deficiencies in another.

Lesh, Behr, and Post (1987) distinguished representations by how easy they were to use, suggesting a continuum of transparent to opaque, referring to how directly the representation illustrates the meaning of what it represents. They used a diagram as an example of a transparent representation, and an equation as a more opaque one. Zazkis and Lijedahl (2004) pointed out that in fact most mathematical representations are fairly opaque, although each representation has transparent elements. Opaque and transparent seem to be classifications similar to "effective" – the user's familiarity with the representation influences how those terms can be applied. Until students gain familiarity with representations of functions, they would all appear opaque to them.

A representation used in mathematics gains transparency with use and familiarity. As more representations can be used effectively, students gain additional tools for doing mathematics. This principle was illustrated in a study of tenth-grade classrooms in preuniversity math courses in the Netherlands (van Streun, 2000). Three classes (n = 74) were taught the mathematics with the aid of a graphing calculator. Four classes (n = 103) were taught the concepts in traditional analytic fashion (equations on paper). The posttest measured knowledge of functions and their applications. Students taught to use a graphing calculator did moderately better than the control group. A further look at the results shows that the two groups did equally well in using analytic solutions, but those who were able to solve in two ways (analytic and graphical) were able to use choose among methods, thus had more items they could complete. Given any set of systems of equations, some will be easier to solve through graphing, and some will be more conducive to analytical methods, but students only used the solution representations that were transparent to them. More experience provided more transparency.

Although students will use the most familiar method when they are given the choice, they often do not think to move to another representation. Even (1998) found this in a study of students' understanding of functions. An open-ended questionnaire was administered to 162 college math students (mostly seniors) and pre-service math teachers. She found that when a problem was presented algebraically, students did not think to solve it using any representation but the problem type, even if another representation would have simplified the solution process. This suggests that they do not understand the value and utility of multiple representations, the focus of the next segment of the review.

Using Multiple Representations

Expertise in a subject such as mathematics is characterized not only by the ability to manipulate ideas within a single representation, but also to translate from one system to another, and recognize ideas across different representations (Lesh et al., 1987).

Multiple representations aid learners of mathematics in the comprehension of function principles. Thus, it becomes important for students to know and understand each way to explore a function, through each representation, which increases their ability to apply this knowledge. With sufficient practice, students become familiar with some representations in isolation. Then they must learn to work with multiple representations.

The purpose of multiple external representations seems to be threefold, to complement other representations, to constrain understanding, or to construct deeper understanding (Ainsworth, 1999). Representations can be used to complement by providing different views on the same concept through various representations. Since different representations provide different views of a topic, multiple representations may be the best way to illustrate the varied features of an object. Multiple representations can also be used to constrain interpretations when a well known representation is used to help understand a less familiar or newer representation. Drawing connections between the familiar idea and the new representation enables the user to see how the other representation operates. Finally, representations enable the construction of deeper knowledge when they are used to promote abstraction, support extension, or teach relations between representations. This use almost always requires translation between representations. The process of translation is not automatic, and is rather complicated. Function knowledge uses representations for each of these purposes.

Translation from one representation to another can be considered an analogy, in that the inherent structure of two disparate representations actually tell the same story between an independent and dependent variable. An analogy is formed through a process of selecting similarities to make salient, and repressing other features that are less important to connect the two objects or phenomena (Holyoak & Thagard, 1997). Few, if any, surface features are similar between the various mathematical representations, thus it takes analogical reasoning to see these relationships. According to a study by Gentner and Markman (1997), deep conceptual knowledge is required to see analogical similarities that are not transparent “[P]ossessing a systematic higher order structure can permit transfer even under adverse transparency conditions” (p. 53).

While multiple representations are needed and used by those who work with functions, research does not support the idea that students have the requisite conceptual understanding to readily integrate these various representations. In fact, the more different two representations appear, the less likely that they will be integrated. This means that for the various representations available for functions, there is little chance of integration without specific instruction on how the representations map one to another.

Seufert (2003) studied helping students learn through multiple representations in science. College students (n = 86) were learning about human metabolism through text and diagrams. They were either provided with directed, non-directed, or no help in integrating the representations. Recall and comprehension measures were tested after the lesson. The participants with directed help had significantly higher recall and comprehension scores than other methods of help. Connections needed to be explicitly stated and demonstrated. Gentner and Markham (1997) suggested that especially when representations are physically different, seeing the underlying similarity is less likely to occur, thus assistance in identifying those similarities is useful to teach connections between representations.

The value of multiple representations is lost if the hard work of connecting the representations is not accomplished. Mayer and Sims (1994) conducted an experiment that illustrated just how much work it was. In this study, 86 college students were to connect two representations in science, a verbal description and a visual diagram. In order to connect the representations, students had to form a mental understanding of both representations individually, but also see and code the relations between the representations as well. "Meaningful learning involves more than building either a verbal or visual representation; the additional component is building referential connections between the two kinds of mental representations" (Mayer & Sims, 1994, p. 400).

Building and using these referential connections between representations is part of a sufficient knowledge of functions. Specific studies on the nature of translation with mathematical functions are treated in the following section.

Translation of Functions

It seems that in general, students are not fluent in translating between common representations in mathematics. In the study by Huntley and colleagues (2000) comparing the two teaching methods, one specific content area studied was translation between various representations. In considering the three representations of graph, table and symbol, students were best at moving from equation to graph. After that was translation from table to graph or equation, then graph to equation, then table to equation. Yet the average score for all of the translation measures was 25% (range: 14 – 33%).

Gagatsis and Shiakalli (2004) also studied students' translation skills. Students from a university in Cyprus (n = 195) participated in a study of their ability to translate among functions. Students took three tests. One provided a verbal description of a set of functions, and then students translated them into a graphical and algebraic format. The other form began with graphical, and participants translated them to verbal and algebraic. The same functions were used on both tests. Students performed differentially depending on the source and target representations. Some combinations, such as verbal to algebraic, were straightforward, and students performed very well on them (average success: 86.6%). Other combinations, such as graphic to verbal, were much harder (average success: 55.2%). Again, representation differences changed the nature of the problem. This supports the task analysis of the transformation process that shows that different analytical processes are needed to perform the various transformations (Janvier, 1987).

Gagatsis and colleagues conducted many other studies of this nature (i.e., Elia & Gagatsis, 2008; Elia, Panaoura, Eracleous, & Gagatsis, 2007; Elia & Spyrou, 2006; Evangelidou, Spyrou, Elia, & Gagatsis, 2004). These studies have been conducted on high school and university students alike. Students are asked to translate from one representation to another, using the same functions but having different starting representations. Findings consistently indicated that performance on items is based on source representation. The results of the studies on representation indicated that students may be proficient in moving from one representation to another (Gagatsis & Shiakalli, 2004), but have not integrated these representations. Each representation continues to be a stand-alone mathematical object, not a set of representations that illustrate an object. In addition, much knowledge appears to be representation dependent, thus a students' ability to demonstrate that knowledge may only be accessible in a specific format.

It is hard to determine why the performance of the students in Gagastis and Shiakalli's (2004) study was more accurate than those students in the study by Huntley et al. (2000). Students were known to be at different schooling levels (university or high school) and geographic locations (Cyprus or USA). Unfortunately not enough information on the instrument was provided in the study by Huntley et al. to investigate this difference any further.

Summary: Representations

Representations are used both individually and multiply in the teaching and doing of mathematics. Representations are intended to be assorted formats of the same idea. Experience is required to see them that way though, and research indicates that many high school and college students do not have sufficient experience to view the ideas behind the representation. In fact, some research indicates that students perceive changing the representation as completely changing the nature of the problem.

To complicate matters further, Even's (1998) study found that college math students' understanding of functions was not only significantly related to representation, but also to the type of function used within the problem. Thus a concept about functions that was understood for one type of function may be missed in another type. These patterns raise questions about the source of a misunderstanding. For any question about knowledge of advanced mathematics or other science, possible sources of misunderstanding could be the concept, the representation, or the situation where it is being applied.

This review has presented information about research findings about student knowledge of function, what it means to know, and external representations in mathematics. In addition to measuring cognitive components of student understanding of function, it is also important to consider beliefs. While knowledge is the primary focus of this research, cognition does not exist independently of beliefs (Pintrich, Marx, & Boyle, 1993), thus the next section briefly overviews beliefs in relation to mathematics.

Beliefs about the Nature of Mathematics

This section reviews findings about beliefs of students on the nature of mathematics. Beliefs are the impetus to action (Hart, 1989), thus are an important consideration in understanding student achievement in mathematics.

Schoenfeld (1989) studied beliefs about mathematics in high-achieving high school students (10^{th} – 12^{th} grade). His study of 230 students was intended to fill the research gap between the fine-grained qualitative studies and the coarse-grained national studies in finding explanations for students professing conflicting beliefs. He found that students believed that the field of mathematics is creative and useful for learning to think logically about complex problems. In contrast, he also found that students believed that most mathematics problems could be solved in two minutes, and could be determined to be impossible after twelve. They also felt that mathematics has rules for all situations and is best learned through memorization.

These contradictory beliefs had also been found in other research (Carpenter, Lindquist, Matthews & Silver, 1983, as cited in Schoenfeld, 1989). Schoenfeld postulated that the apparent discrepancy of beliefs was generated by students perceiving two types of mathematics: school mathematics and real-world mathematics. If two types of mathematics existed then students no longer hold contradictory beliefs, but instead were describing two applications. This adherence to two seemingly conflicting sets of beliefs was also found in research on college students conducted by Crawford, Gordon, Nicholas, and Prosser. (1998b).

Crawford et al. (1994, 19998a, 1998b) studied students' conceptions about mathematics, or students' beliefs about the nature of mathematics. In the initial study students were asked the open-ended question, "Think about the maths you've done so far. What do you think mathematics is?" (Crawford et al., 1994). Five conceptions were identified from student responses. These five categories are hierarchical, such that the first conception is the lowest, simplest idea of mathematics, which expands and develops into the next. Within the five categories, there is a structural shift in how the field of mathematics is conceived, such that the first two categories represent a fragmented sense of mathematics and the other three categories represent a cohesive view of mathematics. The fragmented categories are described as "Maths is numbers, rules, and formulae," and "Maths is numbers, rules and formulae which can be applied to solve problems." The cohesive categories are "Maths is a complex logical system and way of thinking," "Maths is a complex logical system which can be used to solve complex problems," and "Maths is a complex logical system which can be used to solve complex problems and provides insights used for understanding the world." (Crawford et al., 1998b, p. 88).

These categories of fragmented and cohesive were validated in a later study through principal components analysis (Crawford et al., 1998b). Student statements were used to write Likert-style items. Through repeated iterations of the measure, 19 items that described the two scales as anticipated were developed. Although the two factors appear to be opposite ways to perceive mathematics, the factors were not correlated (r = -.13). This was used as evidence that fragmented and cohesive factors are in fact separate ideas. The non-correlated nature of the factors also implied that many students scored high on both fragmented and cohesive scales, manifesting again the conflicting beliefs found by Schoenfeld (1998).

In the original identification of these conceptions (Crawford et al., 1994), approximately 300 first-year math students were asked to provide their conceptions of mathematics and study habits. It was found that fragmented views of mathematics corresponded with surface, or reproductive, strategies for learning, and cohesive views matched with deep, or understanding, strategies for learning.

This relationship was confirmed and extended in a later study (Crawford et al., 1998a). For this study the sample was composed of another set of 300 university students taking first-year math (all had taken two years of high school calculus). They responded to two measures at the beginning of the year, one measuring their conceptions of mathematics (fragmented or cohesive) and one concerning their approaches to study (surface or deep) customized to the study of mathematics. At the end of the semester, participants retook these two measures, as well as taking a questionnaire on their experience in their math course. This measured students' perception of teaching and learning in five categories: good teaching, clear goals, inappropriate workload, inappropriate assessment, and independence in learning.

First the subscales (nine in total) were subjected to correlational analysis, then a factor analysis was conducted on the subscales. Finally a cluster analysis of the students on subscales was conducted (Crawford et al., 1998a). Each of the three analyses indicated the same overall relationship; cohesive views of mathematics matched with positive learning habits and outcomes. Students who indicated that mathematics was a cohesive body of knowledge were more likely to a) use study techniques building towards deeper, more connected knowledge, and b) indicate that their class had good teaching, clear goals, and that independence in learning was promoted. In addition, higher mathematics grades also related to the cohesive view. A fragmented perception of mathematics was associated with studying at the surface level (a goal to pass, not to learn), and perceptions of their math class having an inappropriate workload and assessments.

Students for this study (Crawford et al., 1998a) were first year math students. The report gave no indication of whether students came from one course or many, but given the number of participants, it can be assumed that both participants with fragmented and cohesive views of mathematics were in the same classrooms. Yet they had different perceptions of the math course they took that semester. Thus, it is plausible that perception of learning and teaching are based, at least partially, on incoming perceptions. The authors made no claims to causal relationships, and in the earlier study warned against creating causal links from their correlational analysis (Crawford et al., 1994).

Studies of knowledge are enhanced with an additional look at beliefs because beliefs influence behavior (Mandler, 1989). Knowing how students perceive the field of mathematics enables a more complete picture of their current mathematical status. Awareness of students' perceptions on the nature of mathematics is informative to understanding their knowledge of functions. If mathematics is seen as a fragmented field, students will have no need to connect representations or functions and see the big picture. In contrast, a student possessing a cohesive view will look for those connections which enable a stronger knowledge of functions, and thus future mathematics, as well as a set of positive learning and teaching indicators that associate with cohesive conceptions of mathematics.

To this point, this literature review has highlighted constructs and studies that undergird the measurements used in this study. The next section pertains to statistical procedures that will be used to analyze and make sense of the data, specifically factor analysis and latent class analysis.

Statistical Analysis

This final section in the review of the literature highlights the statistical analyses that are used in synthesizing the data. Factor analysis is a statistical method used to classify variables in large data sets. There are also methods for classifying cases or participants. Two examples are cluster analysis and latent class analysis. These methods are intended to help find the balance between the loss of information when all cases are combined into aggregate scores, and the noise involved in considering all scores individually (Ward, 1963). They are used for several reasons. At their most basic form, they identify groups which can then be labeled, such that large sets of data can be simplified into smaller, more describable sets (Everitt, Landau, & Leese, 2001). This section presents research on classification by item and by case.

Classification of Items

Item classification is commonly done through factor analysis, which is a statistical process used to find the structure of scores within a data set. It is a way to reduce the number of variables to increase ease of interpretation. Factor analysis is a mathematical procedure that allows a larger number scores on observed variables to be organized into smaller groups that seem to have be influenced by the same unmeasured, or latent, construct (Der & Everitt, 2002). The process of conducting a factor analysis evaluates the shared correlations between various observed variables to find the commonalities among them, and identifies those items that group together. Factor analysis methods were developed for use with variables measured on a continuous scale. Procedures have been developed that allow factor analysis to be executed with binary and categorical scores as well. For binary scores, tetrachoric correlations are calculated as opposed to Pearson's correlation (Muthén & Muthén, 1998 – 2007).

Factor analysis is a "process of discovering and defining latent variables and a measurement model that can provide the basis for a causal analysis or relations among the latent variables" (Loehlin, 2004, p. 152). Given a set of measures, factor analyses can be exploratory (EFA) in order to identify the latent variables, or confirmatory (CFA), in which the paths are set and the model is tested for goodness of fit (Der & Everitt, 2002).

Exploratory factor analysis. There are two basic steps in an EFA. The first step is to find the relations between the observed variables and the latent variables (also called factors), using the smallest number of latent variables as will appropriately describe the sample. Factor analysis partitions the variance, attributing the variance of each item among the latent variables. The second step is to rotate the solution, which enables more straightforward interpretation of the scores and their underlying factors (Darlington, 1997). After rotation, an item is said to "load" on the latent variable that is responsible for the greatest amount of variance, but other latent variables are also identified as providing some of the variance of that item (Borgen & Burnett, 1987). This simple explanation glosses over two important choices that a researcher must make, specifically how many factors to extract and what method of rotation to use. These are discussed in the following paragraphs.

In an exploratory setting, the number of factors is not known before the analyses are run. There are a number of heuristics used to determine the best model. The most common is known as Kaiser's criterion (Kaiser, 1960), where the number of eigenvalues greater than 1 is used as the number of factors to extract. This rule has been shown to extract more factors than meaningful though, and Cattell (1966) recommended interpreting a scree plot, which is a plot of the eigenvalues. His research showed that an appropriate number of factors could be identified by taking the number of factors from the “elbow” of the plot, or the point where the graph of eigenvalues levels off. This tends to be imprecise in practice though, as often the “elbow” is not at one point, but over a series of points. Another suggested method is a parallel analysis (Zwick & Velicer, 1984), where random score data is generated, and factor analyzed. The eigenvalues from the actual data and the random data are compared, and the number of factors is determined by how many actual eigenvalues are greater than those generated by random data. Most researchers use a combination of methods.

Not all of these methods for determining the number of factors are applicable to binary scores. Consideration of eigenvalues and scree plots are still used in determining the number of factors (Muthén, 2002), but parallel analysis is not effective with binary data (Tran & Formann, 2009). In addition, some programs provide fit statistics for binary exploratory factor analyses, similar to those found with confirmatory analyses. These are interpreted along the same criterion values (which will be presented in the confirmatory section). While numerical values help the decision process, interpretability of factors is essential in determining the best solution (Muthén, 2002).

Once the number of factors is determined, the second step in an EFA is rotate the solution to simplify the interpretation of the scores (Loehlin, 2004). The type of rotation used is based on the relation between the factors. If the factors are assumed to be unrelated, then orthogonal rotational methods are used, such as the varimax rotation. Correlated factors are rotated with oblique methods, such as the promax rotation (Browne, 2001). A rotated solution provides loadings of an item onto each factor. Loadings range from -1 to 1, and can be interpreted as correlation coefficients between the item score and the factor. There is no absolute standard for when an item loads onto a factor. Various researchers require minimum loadings ranging from .3 to .5 (Costello & Osborne, 2005; Cattell & Vogelmann, 1977). Promax and varimax rotations provide solutions where the majority of items only load onto one factor. Other rotation methods are also available but are less commonly used.

Once the items which load on a factor are identified, the sets of items are analyzed for similarities that may identify the latent factor. Researchers often label factors with descriptive names to provide a sense for their nature and meaning. An exploratory factor analysis only provides a possible structure of item scores and variables based on the results of one sample, and should be interpreted as such. When possible, an exploratory analysis should be confirmed.

Confirmatory factor analysis. In confirmatory factor analysis, the model is specified by the researcher. The research provides a theoretical model of which item scores are predicted by which factors, which is verified through various measures of fit. Individually, each item should have a statistically significant pattern coefficient. In addition, the model should have sufficient model fit statistics. Some major goodness-of-fit indices are chi-square goodness of fit index, Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Root-Mean-Square Error of Approximation (RMSEA) and Standardized Root-Mean-Square Residual (SRMR).

The model chi-square statistic provided in factor analysis tests the null hypothesis that the number of factors identified is sufficient to describe the data. Thus, a non-significant chi-square value ($p > .05$) is desired for this test. This score is always used in conjunction with other measures though, because with large sample sizes, this test has been shown to be overly restrictive (Der & Everitt, 2002). The Comparative Fit Index (CFI) and the Tucker-Lewis Index (TLI) are two sample-based incremental fit statistics. The value of the index compares the fit of the model with the null model. The CFI ranges from 0 – 1, while the TLI can have values greater than 1 (Kline, 2006). Higher values indicate better fit. Yu (2002) recommends a minimum threshold of 0.95 for the CFI and 0.96 for the TLI for good fit.

The root-mean-square error of approximation (RMSEA) is in fact a badness of fit index, thus lower scores are desirable. This measure is population based, and less dependent on sample size. This index measures the error that comes from estimation of the model (Kline, 2006). Yu (2002) recommended that for binary or continuous data, a value lower than .05 indicates good fit. When used with continuous data, this statistic also returns a confidence interval, allowing interpretation of how likely good fit truly is. The standardized root-mean-square residual (SRMR) is a standardized measure of the difference between predicted correlations and observed correlations. This value should be below .10 (Kline, 2006). The SRMR is not an effective measure for binary data (Yu, 2002).

Factor analysis is a powerful analytic tool because it allows large data sets to be reduced to fewer more interpretable factors. It only helps to better understand the nature of the items though. In order to also understand the nature of the participants who responded to the measures, the participants should also be grouped. This can be done with some variety of cluster analysis.

Classification of Participants

There are multiple methods available to create groups of participants. All types of clustering operate under the assumption that populations have homogeneous subgroups with individual profiles along the measured variables (Magidson & Vermunt, 2002). Meaningful groups can be described by the group's unique set of strengths and weaknesses across these specific variables. Once such descriptions have been established, information can be gathered to learn more about the causes and influences of such groups (Zeruth, Kulikowich, & Murphy, 2006). This section will explore two of methods of grouping, sample cluster analysis and latent class analysis. In addition, studies in which these techniques are used in mathematics will be reviewed.

Cluster analysis. Cluster analysis is a system of analyses that partitions sample data to form sets of objects. In it most common application, it divides the cases (often people) from a data set into homogeneous subgroups that minimize differences within groups, while maximizing differences between groups (Borgen & Barnett, 1987). Clustering is a statistical method that groups participants with relevant similarities so that aggregating their data provides meaningful and interpretable sets. Hierarchical cluster analysis allows for an exploration of an initial data set, which can then be confirmed through non-hierarchical, or *K*-means methods. Both use similar grouping techniques and are considered here collectively.

Exploratory clustering initially considers each case as an individual cluster, then clusters the two most similar objects. Similarity is determined by distance measures, such as squared Euclidean Distance or Average Linkage method (Borgen & Barnett, 1987). These two objects receive a combined score and are then considered one object. The process is repeated with the next two most similar objects until all objects are clustered into one group (Ward, 1963). Once an object is categorized, it remains in that cluster even if another cluster emerges that may provide a better fit. The results can be produced as a tree diagram showing how the groups are combined. The number of clusters is determined by considering the theoretical support for the clusters and differences in quality of various partitions.

Confirmatory clustering, or *K*-means clustering, is used when the number of clusters is known, or postulated. *K*-means can be used to verify a solution previously identified through hierarchical methods or when the number of clusters can be predicted theoretically. *K*-means begins by specifying the number (*K*) of clusters that are to be identified. Seeds, or starting values of the centroids, are provided for the clusters. All objects are classified with the nearest seed in order to form clusters. Similar distance measures as those used in hierarchical clustering are utilized. Once all members are classified, the new value of the centroid is calculated. Then the clustering process is repeated. This may result in the reclassification of objects into a more appropriate cluster. This process of recalculating the centroids and reclustering the objects is repeated until a pre-specified number of iterations are run or until cluster membership stabilizes (Steinley, 2003). This becomes the solution.

Cluster solutions are interpreted through comparing the means and distributions of the groups, and describing the particular profile that a cluster appears to represent. The quality of the solution is also determined by considering the relationships of the clusters to external criteria. These groups are then described, and often labeled for ease of interpretation. When speaking of cluster analysis, most researchers are referring to either hierarchal or *K*-means methods (Hartigan, 1996). But, there are additional analyses that classify participants. One of these is latent class analysis.

Latent class analysis. An additional method of finding groups within the data set. LCA is a probabilistic method that determines the likely membership of patterns of scores into groups. It is often applied to variables, thus returning a conclusion similar to a factor analysis, but when applied to people it is sometimes known as Latent Profile Analysis (Muthén & Muthén, 1998-2007). It serves the same goal as hierarchical cluster analysis, in that the scores of participants are analyzed in order to find meaningful partitions. Both work under the assumption that the groups within the sample have unique means, variances, and covariances among the measured variables. Originally, LCA was developed for use with categorical variables (Ubersax, 2008), usually those representing the presence or absence of some condition. Thus hierarchal and *K*-means cluster analysis formed clusters with continuous data while LCA used categorical data. LCA has since been expanded to work with both categorical and continuous data (Vermunt & Magidson, 2003), allowing for more varied data sets to be analyzed.

LCA is a model-based approach. Conceptually, LCA is designed to detect classes and predict membership given the frequencies of the various response patterns. For each participant, the probability is calculated that they would have membership in each class given his or her specific response pattern. They are then classified into the group with the highest probability (Vermunt & Magidson, 2003). Probabilities are used to estimate a model of the population. The fit of the model is tested and various fit statistics are provided that enable the solution to be statistically evaluated. This is contrasted with sample cluster analysis methods which are intended to describe the groups that currently occur in the existing sample and are not mathematically intended to represent a population model, nor do they provide fit statistics. Thus the class membership as determined in LCA is less subjective than cluster analysis, and fit statistics are available (Magidson & Vermunt, 2002). With the inclusion of continuous variables, LCA is a powerful tool for determining population models (Vermunt & Magidson, 2003).

LCA uses the method of maximum likelihood on log-likelihood functions (Magidson & Vermunt, 2002). This enables the assignment of person to cluster to be optimized given certain criteria. Similar to hierarchical and non-hierarchical cluster analysis, the attempt is to minimize within class variance and maximize between class variance. One significant advantage to LCA is that since it provides model data, parameters can be constrained or freed, and fit statistics are provided to help evaluate models and the appropriate number of classes.

Fit of a model is determined by comparing each model to the previous model with one less class (thus the 4-class is compared to the 3-class). Some measures of fit criteria are entropy, information criteria, and the Lo-Mendell-Rubin Likelihood Test. Entropy measures the stability and reliability of the classes, and should be greater than .70 (Schwartz & Zamboanga, 2008). In addition, the appropriate enumeration should have minimum information criteria (IC). IC are predictive fit indices, in that they “assess model fit in *hypothetical* replication samples” (emphasis in original, Kline, 2006, p. 142) The three that are most commonly used are the Akaike's Information Criterion (AIC), the Bayesian Information Criterion (BIC), and the sample size adjusted BIC (adjusted-BIC). If these three values do not minimize at the same number of classes, then the adjusted-BIC should be used because it is the most stable (Nylund et al., 2007). One concern with using the minimum IC score is that oftentimes the values will level off, so that two or three possible solutions have basically equivalent IC scores. To address this issue, an additional fit criterion was developed, the Lo-Mendall-Rubin Likelihood Ratio Test (LMR) to help determine the appropriate number of classes. This test compares the goodness of fit of a solution of *k* classes against the solution of *k-1* classes. This test has been shown to be effective in identifying the “true” number of classes (Nylund et al.). The test statistics of the LMR have a known distribution, so this test returns a *p* value indicating a measure of improvement in the solution that came from the addition of a class.

LCA is not sample dependent and the number of classes is determined more objectively than other clustering methods. When compared with other clustering methods, LCA was more accurate in identifying the correct number of classes and sorting participants into the classes (Magidson & Vermunt, 2002). Both cluster analysis and LCA methods have been used in research on mathematics knowledge. The final section reviews examples of both methods applied to mathematical knowledge and beliefs.

Classification in the Literature

Neither cluster analysis or LCA has frequently been applied to studies of students' ability or achievement in mathematics. No studies were found that considered the nature of clusters in mathematics achievement at a secondary or tertiary level. Classification in mathematics has been applied to special education situations (Yang, Shaftel, Glasnapp, & Poggio, 2005), elementary students (Kyriakides, 2002), and beliefs (Buehl & Alexander, 2005; Zeruth, et al., 2006). An example of clustering on achievement in engineering is also included (Higley, Litzinger, Van Meter, Masters, & Kulikowich, 2007). The results of the studies will be considered here, along with conclusions that are germane to this project.

Achievement. Yang and colleagues (2005) looked for significant groups within students with learning disabilities in the 4th grade. Using LCA, the responses and scores of 2681 learning disabled students on the state's annual assessment were grouped. They determined that there were two classes. Yang et al. expected that the classes could be defined by mastery in some content areas, and weaknesses in others, and that this would correspond with subsets of LD categorizations (e.g., mentally retarded, visual impairment, etc.), allowing for more targeted prescriptions of assistance. Instead they found that the groupings were not according to skill sets, but instead related to overall mastery of content. One group did better in all content areas than the other. Students of all LD classifications were included in both groups. The authors recognized the groups were described by quantitative differences, as opposed to the qualitative differences they were expecting to find. Although unanticipated, this idea of quantitative differences provides a useful frame for interpreting other clustering research.

Kyriakides (2002) investigated the mathematical abilities of students (n = 1670) in Cyprus entering their first year of primary school and then again at the end of their second year. The students were given a developmentally appropriate test of performance over the six sets of math knowledge that would be expected of students in their first year, and then again at the end of their second year (one test for each year). Two hierarchical cluster analyses were run, one for each year. Five clusters were established from the data of the first year. Three clusters matched the quantitative conclusions of Yang et al. (2005). There was an overall high performing cluster (students who had mastered most of the 20 mathematical objectives), an overall low performing cluster (students who had mastered only a few of the objectives), and an average cluster. They did find two qualitatively different clusters, though, showing strengths in some areas and weaknesses in others. When the second cluster analysis was run on the performance measure from the end of the second year, only three clusters still existed: the high, average and low. This caused Kyriakides to reconsider the first data set, and it was found that the three cluster solution had equivalent explanatory power as the five. Thus, the more parsimonious model was chosen, exhibiting the three clusters representing three levels of overall proficiency.

Higley et al. (2007) used cluster analysis techniques to study background knowledge in second year college engineering students. Using Ward's hierarchical clustering methods, they clustered students on concept knowledge, SAT scores, and spatial reasoning scores. Six clusters emerged, allowing for the variation in measures that one would expect, where a particular group has, for example, high SAT scores, but average concept knowledge, or typically low scores except a high spatial reasoning score. As it was hierarchical clustering though, solutions with fewer clusters could also be investigated. The three cluster solution led to the quantitative conclusion, overall levels of performance emerged: all high, all average, and all low. While the chosen six-cluster model had better explanatory power than the three-cluster, it is not clear that the additional clusters were statistically or practically significant.

While it is no surprise that students who do well in one area measured would also do well in another, cluster analysis is often thought of as a way to find unknown profiles, the combination of skills and qualities that would not be anticipated, thus the reason for these exploratory analyses. As a result, the steadiness of the high and low responses gives pause to the value of applying cluster analysis to achievement data. It appears that if qualitative differences are desired, then solutions must have more than three groups. However, solutions with larger numbers of clusters should be compared to more parsimonious solutions to ensure that the additional groups are meaningful.

Beliefs. When cluster analysis is used on beliefs, this quantitative grouping effect is tempered, but still exists. Buehl & Alexander (2005) studied epistemic beliefs of college students in history and mathematics. Four clusters were found for the three types of epistemic beliefs in history (certainty of knowledge, authority of sources, isolation from other subjects). The four clusters were basically indicators of amount of endorsement of beliefs. The first group endorsed the beliefs the most; the last group endorsed them the least. The relation between the three sets of beliefs stayed basically constant across the groups. The same three sets of epistemic beliefs were also measured in mathematics. Again four clusters were found. A slightly different set of profiles emerged. Although two clusters showed evidence of being only distinguishable through their level of endorsement of the beliefs (not the relation of the beliefs to each other), the other two groups were distinct, showing different relationships between the three types of beliefs. The quantitative difference of Yang et al. (2005) did not apply to all clusters.

A study investigating beliefs about mathematics by Zeruth and colleagues (2006) also returned profiles that were not solely based on level of endorsement. They identified three types of beliefs, importance or value of mathematics, difficulty of mathematics, and certainty of mathematics. As defined by the researchers, a high score in importance and low scores in difficulty and certainty would be the profile one would expect from a mathematician. If the scores for difficulty and certainty were reversed (such that they become measures of non-difficulty in learning and non-certainty, respectively), one cluster scored high across all beliefs. The next one has moderate importance and non-certainty beliefs with low non-difficulty beliefs, while the final one has low importance and non-certainty beliefs, with moderate non-difficulty beliefs. Thus there is evidence that, as measured, these beliefs do not cluster solely on the basis of levels of endorsement.

Three knowledge measures were included in the further investigations of the clusters (Zeruth et al., 2006). These measures were not used to establish the clusters, but to learn about the nature of the clusters. The knowledge measures also did not group along quantitative divisions. A measure of recall corresponded with importance and non-certainty beliefs. The multiple-choice test on declarative and procedural knowledge of probability, and four problem-solving tasks corresponded with non-difficulty beliefs.

Clusters are intended to represent subpopulations within a larger group that can be differentiated by the characteristics of their scores. For the articles reviewed for this dissertation, clusters done on mathematical abilities formed clusters based on level of knowledge, not types of knowledge, such that only one real profile emerged, with varying degrees of intensity. When mathematical beliefs were used as the clustering variables, then profiles of varying levels between the variables did emerge. There are many differences between the knowledge and belief studies though. Primarily, there is the major difference of cognitive versus belief measures, but there is also a difference in age groups. Children tended to group in fewer clusters with less variability, adults (college students) into more groups with greater variability. The study by Higley et al. (2007) suggests that perhaps much of the profile effect of clusters comes with a larger number of clusters (see also Kyriakides, 2002). Everitt et al. (2001) propose that more clusters are more informative.

Summary of Statistical Analysis

Creating profiles of variables or cases reduces the multivariate space of the data, thus making data easier to understand. Identifying and naming groups is a powerful way to increase understanding of a phenomena (Hartigan, 1996). There are various way to find these groups. Factor analysis is the most common method of grouping variables. It allows researchers to identify and test models of relationships of between scores from observed and latent variables. Analogous to factor analysis are various methods to classify participant. Both cluster analysis and LCA create homogeneous subgroups from the sample. Both factor analysis and latent class analysis provide models of the population allowing generalizability to other samples. All methods require research judgment and heuristics to interpret the solutions.

Summary

The focus of this review of the literature was to describe the constructs undergirding the measures and that will inform the analysis of the data. Past studies of student knowledge of function, defining knowledge, representations in mathematics, and statistical analyses used in the study were reviewed here.

The first section reviewed what past research has reported about students' knowledge of functions. The primary conclusions of the studies were that students did not know functions sufficiently well. Students did not seem to know precisely what a function is, although they had a general image of it (e.g., Monk, 1992; Vinner, 1983). Studies of general knowledge of applied functions evidenced similar deficits in knowledge. Studies that focused only on student knowledge of functions were hard to find.

Most of the studies of function knowledge were conducted many years ago with small samples and through qualitative means. While this research provided rich descriptions and deep understandings of the phenomena at the time, generalizability comes from larger scale research. This current study attempts to apply the findings of earlier studies and see if they replicate in larger samples. Those studies that were conducted on larger samples were not actually studies of function knowledge, but studies of teaching methods. These provided comparative measures of function knowledge, but direct measures are also needed.

The study of knowledge, especially in mathematics, requires that the various types of knowledge be considered. Thus, a significant portion of the review was dedicated to dissecting conceptual, declarative, procedural, and conditional knowledge. The basic distinctions are between knowledge that can be stated and knowledge that can be enacted. These take on different forms depending on the field of research. What studies have found is that at basic levels, these divisions make sense. No matter the complexity of the task, if a fine enough grained analysis is done, the task can be split into declarative and procedural knowledge (Anderson et al., 2004).When studying children's beginning of knowledge, conceptual and procedural knowledge are separable (Rittle-Johnson & Siegler, 1998). But at more complex levels, such as advanced mathematics, there is much more interplay and these distinctions are harder to separate (Baroody et al, 2007; Byrnes & Wasik, 1991; Star 2005).

Yet utilizing a test designed with knowledge distinction is worthwhile, because to truly master a subject, one must know it both procedural and declaratively. A test utilizing both declarative and procedural knowledge and application items allows for a more balanced test that provides a broader test than is often measured in a research study, and a more focused test than that which takes place in a classroom. It is possible that at complex levels these types of knowledge are inseparable, but that is an empirical question.

The third section of the review demonstrated the central importance of representations in the measurement of knowledge of functions. In many studies not intending to study representations, this became the subject of research because of how the students responded and used representations (Even, 1998; Meltzer, 2005). Although the goal of teaching about functions is to have students look beyond the representations and see the objects that are depicted by them (Lesh et al., 1987), research shows that changing the representation changes the problem (e.g., Elia et al., 2007). Thus, any research on functions must either measure knowledge of representations, or carefully control for their influence. Truly, given of the fundamental link between function knowledge and representations, it essential to study them in tandem. Although much research studied representations of functions (e.g., Even, 1998; Gagatsis & Shiakelli, 2004), no research studied both applied knowledge and representations with equal focus.

The review also included research on beliefs in mathematics, specifically a measure that was expected to augment the findings of the two more cognitive instruments. The Conceptions of Mathematics instrument (Crawford et al., 1998b) measures how students perceive the field of mathematics, but also relates to students' study habits and perceived experiences in mathematics classrooms (Crawford et al., 1998a). This provides an important perspective on the groups of students that are expected to be found through this research.

The final section reviewed methods and relevant research on grouping items and people. The basic framework of factor analysis and latent class analysis were presented, as well as research that used these methods in studying mathematics achievement. Using these methods in this research will provide valuable insights to the empirical nature of functional knowledge and how that knowledge describes groups within the population. The research on grouping indicates that at least two groups will be identified, those who perform well and those who do not. Additional groups may allow for qualitative differences to be identified and described.

CHAPTER 3

METHODS

The purpose of this dissertation was to measure and interpret the results of college students' knowledge of functions. This knowledge was theorized to be in four forms. Three were declarative, procedural, and conditional knowledge. These were measured with one instrument, the Basic Knowledge of Functions Test. The fourth type of knowledge was the ability to translate between representations; which was measured with the Representational Transfer Test. The first focus of this dissertation was to establish the validity and reliability of test scores and to investigate the connections among the various subscales of the instruments. The second focus was to form profiles of student knowledge in order to better understand the nature of this knowledge. In order to accomplish these tasks, a number of research questions were asked:

1) To what extent does the factor structure of the Basic Function Knowledge Test (BFKT) support the validity and reliability of scores?
2) To what extent does the factor structure of the Representational Transfer Test (RTT) support the validity and reliability of scores?
3) Does the factor structure of the Conceptions of Mathematics Scale (CMS) match that of the literature? If not, what structure does seem to exist? Are these scores reliable?
4) What relationships exist between the scales from the BKFT, RTT, and CMS?
5) Given the scales of the three tests, can scale scores be used to form profiles of student knowledge of functions?

6) To what degree are profiles related to important academic performance criteria such as SAT mathematics scores and grades in mathematics courses?

In order to answer these questions, the instruments were administered to a sample of university students. The descriptions of the sample, instruments, and procedures used to analyze the results follow.

Participants

Participants for this study were undergraduate students at a large northeastern research university. They were recruited from courses offered from the College of Education and from the College of Science, department of Biology. The education courses were two introductory educational psychology classes; the science courses were a human anatomy class and a physiology course. In addition, students taking the Physiology Lab (not necessarily the students currently taking the physiology course) were also recruited. The education course is required of all Elementary and Secondary Education students. The biology courses are required for many science majors and fulfill general education requirements, thus these courses have high enrollment and a wide range of students. All students received course extra credit for their participation. Differences were expected in groups of students from the two types of courses (education or science). This matter was explored when the sample was split.

The instruments were completed by 642 students. Two cases were deleted because the students were not undergraduates. Approximately 38% of participants were from the education courses, and 62% were from the biology courses. Four hundred ninety-nine (78.0 %) of the respondents were female. As would be expected from the demographics of the institution, the students who responded were primarily white (88.4 %). The other ethnicities reported, African-American, Asian-American, Hispanic, or foreign, each represented 2 – 3% of the respondents. A little more than half of the participants (58.1%) were sophomores, with the rest of the students basically split across the remaining three years of university attendance. Seventy-two distinct majors or minors were listed by students in the sample. See Table 1 for a list of majors identified by at least ten students.

Table 1. Majors by College

College Major	n	Percent
Education		
Rehabilitation and Human Services	13	2.0
Education	13	2.0
Elementary Education	98	15.3
Secondary Education	43	6.7
Special Education	10	1.6
Engineering		
Bioengineering	27	4.2
Health and Human Development		
Biobehavioral Health	63	9.8
Communication Sciences and Disorders	37	5.8
Kinesiology	119	18.6
Nursing	69	10.8
Nutrition	16	2.5
Liberal Arts		
Psychology	12	1.9
Science		
Biology	20	3.1
Life Sciences	11	1.7
None		
undecided	10	1.6

Note. Only majors indicated by at least 10 students listed here.

The average GPA was 3.36 (*SD* = 0.39). The average math SAT score was 593.9 (*SD* = 82.7), and the mean verbal SAT scores was slightly lower, 575.6 (*SD* = 74.0) (total = 1169.5). Forty-one percent of the students (260) reported that they did not take either high school or college calculus, while 22.2% (142) students reported taking it both in high school and college. Of the 352 students who reported taking at least high school calculus, their average grade was a B+/A- (*M* = 3.52; *SD* = 0.63, where A = 4, B = 3, etc.). Of the 170 students who reported taking the 1st semester college calculus, the average grade was a B (*M* = 3.03, *SD* = 0.84). Most of the students (85.5 %) reported taking four years of high school math (range 2 – 4). The range of college courses was larger, 0 – 6, but 80% took between one and three math courses (*M* = 1.75). See Table 2 for a list of the courses taken by at least 10 students and the number of students who indicated taking that course.

Table 2. Mathematics Courses Taken

Course Title and Number	n
Finite Mathematics	35
College Algebra I	167
College Algebra II & Analytic Geometry	104
Plane Trigonometry	56
General View of Mathematics	18
Trigonometry and Analytic Geometry	18
Techniques of Calculus I	26
Calculus with Analytic Geometry I	129
Calculus with Analytic Geometry II	99
Problem solving in Mathematics (El Ed only)	97
Matrices	31
Calculus and Vector Analysis	22
Calculus of Several Variables	17
Ordinary and Partial Differential Equations	23
Analysis and Interpretation of Statistical Data in Education[a]	38
Statistical Concepts and Reasoning[b]	58
Elementary Statistics[b]	103

Note. Only courses indicated by at least 10 students listed here.
[a]This course is offered through the educational psychology program.
[b]These courses are offered through the statistics department.

Measures

This section describes the four measures taken by the participants in this study. There were two cognitive measures and a belief scale. This was the first research application of the two cognitive measures; the belief scale had been previously investigated.

Basic Knowledge of Functions Test

The Basic Knowledge of Functions Test (Higley, Higley, & Kulikowich, 2005; see Appendix A) was the first measure investigated. The content of this measure (basic knowledge) was determined through an investigation of what is initially taught in college algebra texts, the NCTM standards, and a review of literature on student knowledge of functions (e.g., Dugopolski, 2003; Even, 1998; Hughes-Hallett et al., 2002; Miller, 2005; Mueller & Brent, 2006; NCTM, 2000). This instrument was designed to measure procedural, declarative and conditional knowledge (Alexander et al., 1991) about functions. See Figure 1 for a list of questions by knowledge type and an example of each type of question.

There were 31 multiple choice items on this measure. The questions were about definitions, range and domain, evaluating functions, and identifying families of functions, as well as questions that require students to interpret function representations. Declarative knowledge was measured with 11 items. These questions focused on definitions and rules presented in both words and diagrams. Another 12 questions measured procedural knowledge, which was operationalized as calculations or following algorithms to do mathematics. Conditional knowledge was measured with 8 application items requiring interpretation of functions in their various representations, both in concrete and abstract examples. The 31 multiple-choice items were scored as correct or incorrect, thus the highest possible score was 31. The Kuder-Richardson-20 (KR-20) for the scores of this measure was 0.75.

Figure 1. Examples and Items by Knowledge Type from the BKFT

Declarative – Items 1 – 10, 12

4. What is the domain of a function?

*a. the set of possible values of the independent variable

b. the set of values for the dependent variable

c. the values of x chosen for graphing purposes

d. the arithmetic process used to find the dependent values

Procedural – Items 11, 13 – 23

14. Find $f(3)$ if $f(x) = 2x^2 + 4$.

a. 13

b. 40

c. 10

*d. 22

Conditional – Items 24 – 31

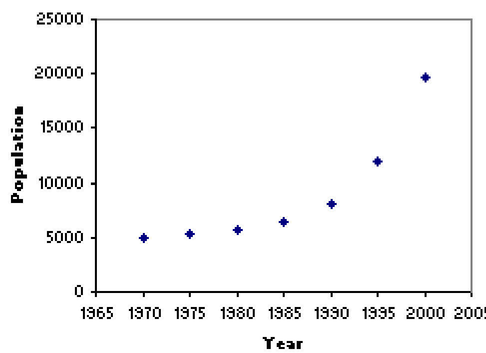
28. Given the following graph, what sort of function would best model the data?

a. linear

*b. exponential

c. quadratic

d. inverse



Free-Response Item – Item 32

32. Please graph the following, and state whether or not it is a function.

$$f(x) = \begin{cases} x^2; for\ x \leq 2 \\ 1; for\ x > 2 \end{cases}$$

The final item of the BKFT (#32; see Figure 1) was not considered part of the instrument for scoring and analysis purposes, but instead was used as a criterion measure for the latent class analysis. This question measured the participant's ability to construct a piece-wise graph and recognize it as a function. Two scores were assigned to this problem, one for the graph and one for identification of function. Both scores ranged from 0 – 2. Zero meant no answer was provided, 1 indicated that the provided answer was wrong, and a score of 2 was given for the correct answer.

Representational Transfer Test

The Representational Transfer Test (RTT; see Appendix B) was originally designed by Feinn and Kulikowich (1999). This instrument was designed to measure ability to recognize one representation of a function given another. For each question, participants were provided with a source function in one of the four modalities: verbal, graphical, algorithmic, or tabular. They then chose the appropriate target function from a set of four options. There were 12 questions for each function family included (i.e., linear, quadratic, cubic, logarithmic, exponential, trigonometric) in order to measure each combination of modalities. The original instrument contained a total of 72 items.

In evaluating the test for use in this study, there was a concern that using less familiar functions would not measure translation ability, but would instead measure knowledge of advanced functions, thus the RTT was edited and condensed. First, only questions about linear and quadratic functions were used. Then, since the scope of function type was limited, answer choices were altered so that they primarily only included linear and quadratic functions. The shorter, more familiar test was also thought to reduce possible testing fatigue. In either format, this was the first utilization of this test in a research study.

There were 24 multiple-choice items on this test. Items were scored as correct or incorrect. Because of a typo in the test, one item did not have a correct answer, thus the highest possible score for this test was 23. The KR-20 measure of reliability of the scores for this test was 0.81. See Figure 2 for an example of a quadratic question, moving from a verbal source to a table target response.

Figure 2. Example from the Representational Transfer Test

1. An upward opening parabola in which the vertex is one unit down from the origin.

a. *x*	*y*	b. *x*	*y*	c. *x*	*y*	*d. *x*	*y*
0	-1	0	1	0	-1	0	-1
1	-2	1	-1	1	0	1	0
2	-3	2	1	2	1	2	3
3	-4	3	-1	3	2	3	8

Conception of Mathematics Scale

The belief measure, Conception of Mathematics Scale (CMS), was taken from the literature (Crawford et al., 1998b). This 19 item instrument used Likert-scale responses (ranging from 1 – 5) to determine whether students view mathematics as a coherent body of information or as fragmented, unconnected facts. Ten items measure the fragmented scale and nine items measure the cohesive scale. This two-factor structure has been replicated in the literature (Alkhateeb, 2001; Arigbabu & Mji, 2005; Mji, 1999; Mji & Klaas, 2001). See Figure 3 for examples of an item from both scales. The reported post-test Cronbach's alpha for this scores of this test are .88 for the cohesive scale, and .85 for the fragmented scale. Cronbach's alpha for the scores of this sample was .89 for the cohesive scale and .81 for the fragmented scale.

Figure 3. Examples from the Conceptions of Mathematics Scale (Crawford et al., 1998b)

Fragmented 18. Mathematics is the study of the number system and solving numerical problems.
Cohesive 19. Mathematics is models which have been devised over years to help explain, answer and investigate matters in the world.

Epistemic Beliefs Scale

The second affect measure was a variation of a Teachers' Epistemic Beliefs Instrument (TEBI; Hennessey, 2007; see Appendix C). The instrument was originally written for teachers, referring to their beliefs of best teaching practices. It was designed to measure their belief frameworks of foundationalism, coherentism, and reliabilism, and was shown to be reliable and valid for those purposes (Hennessey). No factor analysis was performed on the scale when used with the teachers because of insufficient sample size. For this application, the TEBI was modified for students. While significant differences existed between each of the scale scores, the exploratory factor analysis returned a solution with two factors. When a three factor solution was fitted, the items did not correspond with their intended groups. Whether this is an indication of a problem with the measure, that the measure is not suitable for students, or a loss of psychometric properties when varied is an empirical question. Given the indeterminate results, the findings for this instrument were not included in this study.

Pilot

A small pilot of the instruments was conducted before its general use. Approximately 50 students were recruited from the course, Calculus and Vector Analysis

(Math 230), to take the instrument. A gift certificate for an ice cream cone was offered in return for their participation. Completed instruments were returned by 15 students. Most of these students were sophomores. All had taken at least four mathematics courses in college, and had taken 1st semester calculus. Three were math education majors, the remaining 12 were in science majors. They were compatible with the actual sample in terms of gender, ethnicity, and age, but on average had more math course work.

Given their mathematical knowledge, it was expected that most students would ceiling on most items. All but one item in the RTT were answered correctly by at least 80% of the students. Twenty out of 31 items had over 80% correct for the BKFT. Any item answered incorrectly by more than seven (approximately half) of the students was investigated further. One item from the RTT missed this criterion, and five from the BKFT. Of these six items, three of them had mistakes. These were corrected before the final instruments were distributed. One typo on the RTT was not found because it was the most reasonable answer.

This left three items from the BKFT in which less than 50% of the students answered the questions correctly (#s 3, 5, and 12; see Figure 4). The content of these items were directly based on the definitions of functions and their inverses, and were designed to mirror misconceptions often identified in previous studies. Given past research on functions, students do not have a solid knowledge of definitions, especially when their use of the concept does not require a precise definition (Vinner, 1983). These items were left intact because the low performance was expected, and the relation of these definitional (declarative) items with other procedural items was of interest for this investigation.

Figure 4. Items Missed by Over 50% of Pilot Participants

3. Which of these equations defines y as a function of x?
   - a. $(y+2)^2 = x$
   - b. $(x-2)^2 + (y-1)^2 = 9$
   - *c. $y^3 + 3 = x$
   - d. $x = 3$
5. Which of the following must be true for all functions?
   - a. There must be an equation that defines it.
   - b. It must be continuous.
   - c. An equation must exist defining one variable by the other.
   - *d. It is a mapping from one set to another set.
12. If f(x) and g(x) are inverse functions defined on the real numbers, which of the following is NOT always true?
   - a. f(g(x)) = x, for x in the domain of g.
   - b. g(x) is a one-to-one function
   - c. the graphs of f(x) and g(x) are symmetric about the line y = x.
   - *d. the domain of f(x) and g(x) are equivalent

Procedures

Recruitment and Data Collection

Students were recruited from classes offered in two colleges, the College of Education and the College of Science. The researcher visited the education and biology classes informing them of the nature of study and the opportunity for extra credit. A follow-up e-mail was sent after the visit with a sign-up for a specific time to take the measures. Because of the large number of labs (approximately 15), individual visits were not made, but a more detailed e-mail was sent to serve as the recruitment. Research sessions were scheduled over a one week period. An hour was provided for participants to respond to the measures, but no student took longer than 45 minutes.

Upon arrival to the research sessions, informed consent forms were distributed to students with some verbal explanation of the form provided. All measures were taken in one sitting. The measures were paper-and-pencil, which enabled the collection of not only answers, but also a trail of any written work that may have been required to arrive at the answers.

The measures were ordered into four forms. All students first filled out the demographics page (see Appendix D). Each form started with a cognitive measure, either the BKFT or the RTT. One of the beliefs measures followed, either the CMS or the TEBI (results of this scale not used in this analysis). Then the other cognitive measure was given, then the remaining belief scale. This counterbalancing effort was to protect against order effects in the taking of the tests. On the final page, students indicated what course they were recruited from and their student id number in order to receive extra credit. This information was removed from the rest of the measures, which were then identified simply with a random id number.

Data Entry

Data were entered into a spreadsheet by the researcher. Responses from the BKFT and RTT were entered as the student's response, then transposed into binary scores, 1 for correct and 0 for incorrect. Responses for the CMS were coded as the student's response. Unanswered items or items in which two selections were made were left blank in the original spreadsheet.

There were two types of missing data, individual items or connected sets. For the cognitive measures, individual items were replaced with a 0, under the assumption that no response indicated that the students did not have the threshold of knowledge for that item to earn a 1. Sets of unanswered items were left blank because it could not be determined whether the student missed a page of the measure or chose not to respond to those questions. Replacing the missing data made only small changes in the aggregate data, but recovered 60 cases for data analysis (see Table 3). Missing data for the belief scales were left blank. Thus, the final sample sizes were 636 for the BKFT, 630 for the RTT, and 630 for the CMS. There were 620 cases with complete data.

Table 3. Comparison of Means after Replacing Missing Data

			With all Missing Data			Missing Data Replaced		
Measure	min	max	*n*	*M*	*SD*	*n*	*M*	*SD*
BKFT	5	31	576	16.15	4.79	636	16.07	4.82
RTT	2	23	570	16.94	4.16	630	16.66	4.30

Splitting the sample

The sample size was sufficient to allow the sample to be split. This enabled both exploratory analyses and confirmatory analyses for the factor analyses, so more confidence could be placed in the findings. Gender and course effects were expected, thus stratified random sampling was used. The sample was first divided into groups according to course and gender, creating four groups, plus a fifth including students who were missing some of this demographic information, thus could not be classified according to this scheme. The groups were: females in the education courses, males in the education courses, females in biology courses, males in biology courses, and students missing gender or course information. Each group was randomly split, then combined to form two sample groups. One group was used for exploratory analyses (n = 301), and the second (n = 339) was used for confirmatory analyses. See Appendices E and F for the demographics and descriptives of both samples. This method of splitting the sample generated very similar groups, both demographically and by scores on measures. ANOVAs were run to ensure the comparability of those groups. The subsample was used as the independent variable and the study measures were used as dependent variables. The ANOVAs showed no significant differences (BKFT: $F(1, 628) = 1.183$; RTT: $F(1, 634) = 1.172$; CMS: $F(1, 628) = .042$). Chi-Square tests of demographic groups also showed the groups to be similar. The entire sample was used for the latent class analysis.

Data Analysis

The section overviews the data analysis techniques that were used to answer the research questions of this study. The first aim of this dissertation was to verify whether the scores from the instruments grouped as anticipated, and if not, then what latent variables were measured. This was done by conducting factor analyses on the instruments. Relationships between the factors were observed. The second aim of this dissertation was to find profiles of the students who participated. This was done by identifying profiles of factor scores using latent class analysis. This section is organized by research questions; questions 1 and 2 are treated together because the same processes were used.

Research Questions 1 and 2

The first two research questions addressed the validity and the reliability of the scores of the BKFT and the RTT. Since the same procedures were followed for both tests, the methodology is only presented once. Validity was assessed in two ways. The first was through a content validity check. Following the content validity, factor analyses were run. After the analyses were completed, reliabilities of the factor scores were calculated.

Content validity. The first step to ensuring content validity was in the test development, in that a mathematician was one of the authors of the each of the tests. Thus in the writing of the test, questions were checked for mathematical accuracy and appropriateness. Once the tests were completed, they were provided to two content experts who evaluated each item for its fit as knowledge that was procedural, declarative, or a combination. One content expert is the head of undergraduate mathematics teaching at a university; the other is a high school math teacher of 15 years who has taught calculus for five years. Both have a comprehensive knowledge of functions as taught to students.

Factor analysis. Factor analyses are intended to uncover the latent variables that contribute to the observed variables (Loehlin, 2004). In this situation, the observed variables were the individual items from the measures, and the latent variables represented types of knowledge that influence the responses to the items. Factor analysis requires the observed variables to be continuous, but methods have been developed to handle binary scoring of those variables. These methods operate under the assumption that a threshold exists along the continuum that marks the division of 0 and 1. This assumption was met with this data. Factor analysis of binary variables uses tetrachoric correlations (Muthén & Muthén, 1998-2007).

Exploratory factor analyses (EFAs) were conducted on the scores from the exploratory sample using Mplus 4.21 (Muthén and Muthén, 2007). Listwise deletion was used to handle missing scores. Mplus was chosen because of its ability to handle binary outcomes variables. Mplus uses WLSMV (weighted least squares with mean and variance adjusted) as the default estimator of categorical data (Muthén & Muthén, 1998 – 2007). WLS estimation is suited for use on categorical and binary data because these data do not meet the normality assumptions required for maximum likelihood estimation (Kline, 2006). Mplus produces a number of factor solutions (as specified by the researcher) and the appropriate solution is then identified. Solutions with one to four factors were evaluated. The identification of the best solution was done by observing eigenvalues patterns (Kaiser, 1960), scree plots (Catell, 1966), chi-square goodness-of-fit tests, the Root Mean Square Error of Approximation (RMSEA) (Yu, 2002), and the interpretability of the solution (Muthén, 2002). Parallel analysis has been shown to be an effective method of establishing the number of factors given continuous observed variables (Zwick & Velicier, 1984), but does not perform with similar efficacy with binary data (Tran & Formann, 2009), so was not used.

In Mplus (Muthén & Muthén, 2007), factor solutions are rotated using both Varimax and Promax rotations. Promax rotations were interpreted for the purposes of the BKFT and RTT because of the assumed correlational nature of the factors. The minimum factor loading was set at .30 (Costello & Osborne, 2005). The set of items that made up each factor was analyzed in order to interpret their similarity and to potentially identify the latent variable.

Factors were to be confirmed on the confirmatory sample using the CFAs specifically tailored for binary data available through Mplus. Weighted least-squares estimation was again used. Three models were run for each test. One was simply the items that loaded onto the factors in the EFA. One was a model of the strongest items, using only items that had high factor loadings (greater than .50) on the factor. The final model used all items from a test, and was constructed of conceptually fitting items despite their loading during the EFA. The three models were each considered for fit through appropriate goodness-of-fit indices (i.e., chi-square, CFI, TLI, and RMSEA).

Reliability. From the results of the factor analyses, component scales were created for both the BKFT and the RTT. These component scores were summed into scale scores representing specific latent variables. Reliability estimates (Cronbach's alpha) were calculated for these subscales to further support their validity.

Research Question 3

The third research question addressed the validity and reliability of the belief measure, the Conceptions of Mathematics Scale. Crawford et al (1998b) reported a two-factor structure of the CMS. They analyzed the instrument using Principal Components Analysis with maximum likelihood estimation and varimax rotation. These results were verified in this study by exploratory and confirmatory factor analyses using Mplus (Muthén & Muthén, 2007) to mirror the analysis used for the BKFT and RTT. Two exploratory factor analyses were run. The first treated the data as continuous, using a maximum likelihood estimator, as is commonly practiced for Likert-type data. In contrast, Muthén & Muthén (1998 – 2007) recommend treating Likert-type data as ordered categorical scores and suggest conducting a weighted-least-sqaures analysis. This analysis was also run. The two solutions were compared with each other and with the reported structure (Crawford et al.) to determine best fit.

The results of the EFA were replicated in a CFA, also using Mplus, to confirm the factors. Fit was determined by considering the appropriate fit statistics. Once the factors were confirmed, Cronbach's alpha was calculated for each scale to provide evidence of reliability.

Research Question 4

Once the scores from the tests and subscales were shown to be valid and reliable, the next research question addressed the relations that existed between and among the various constructs. Correlation coefficients between scales of the three measures were analyzed. Pearson's r (Glass & Hopkins, 1996) was calculated for relationships between the various subscales of the BKFT and RTT. In addition, these subscales were correlated with the factors underlying the CMS.

Research Question 5

The fifth research question addressed whether meaningful profiles could be created given the score patterns. After determining that informative factors had been established for the various measures taken by the participants, student profiles were determined. This was done using the scores on the subscales of the BKFT, RTT and CMS in a latent class analysis. Mplus 5.2 (Muthén and Muthén, 2007) was used to perform this analysis.

A one-factor model was fit first for comparison purposes. Subsequent models added an additional class until the appropriate solution was identified. The number of classes was determined by considering entropy values for the solution, finding minimum information criteria (AIC, BIC, adj-BIC), and a non-significant Lo-Mendell-Rubin (LMR) test. Entropy is a measure of reliability of classes, and the minimum value was set at .70 (Schwartz & Zamboanga, 2008). Three information criteria were observed, Akaike's IC, the Bayesian IC, and the sample-size adjusted Bayesian IC. Minimum IC values were desirable. Finally, the LMR test will be non-significant ($p > .05$) at the correct number of classes (Nylund et al., 2007).

Means and standard deviations on the grouping variables for the classes are provided in order to describe the different groups. A MANOVA was run to ensure significant differences among the classes. This analysis was run using the class as the independent variable and the latent variables from the measures as the dependent variables. Other demographic information by class is also set forth.

Research Question 6

The final research question addressed relations between class membership and other external criteria such as SAT scores, GPA, whether calculus was taken, grades in university calculus courses, and the score on the graphing item from the BKFT. This item provided the equation of a function, and students graphed the function and identified whether they believed it was a function or not. Although this item was initially assigned two scores, one for the graph and one for the identification, for the purposes of this analysis, it was scored as correct or incorrect.

Descriptives were produced for each class, and compared to the total sample. Differences in the classes were determined by looking for significant differences between the classes and their scores on these criteria. ANOVAs were run to see if there were significant differences in SAT, GPA scores, and grades. Chi-square tests were run to see if there were differences in how many students within a class had taken calculus and correctly responded to the graphing problem.

CHAPTER 4

RESULTS AND DISCUSSION

The research done for this dissertation had two primary purposes. The first purpose was to validate two sets of test scores intended to measure knowledge about functions. In addition, those measurements, and one concerning students' beliefs on the nature of mathematics, were used to define profiles of university students' knowledge. The first four research questions addressed the tests and the scales that composed the tests:

1) To what extent does the factor structure of the Basic Function Knowledge Test (BFKT) support the validity and reliability of scores?
2) To what extent does the factor structure of the Representational Transfer Test (RTT) support the validity and reliability of scores?
3) Does the factor structure of the Conceptions of Mathematics Scale (CMS) match that of the literature? If not, what structure does seem to exist? Are these scores reliable?
4) What relationships exist between the scales from the BKFT, RTT, and CMS?

The final two research questions addressed the relations among the scales and the profiles of student knowledge described by those scales:

5) Given the scales of the three tests, can scale scores be used to form profiles of student knowledge of functions?
6) To what degree are profiles related to important academic performance criteria such as SAT mathematics scores and grades in mathematics courses?

As such this chapter is presented in two parts, the first addressing the structure of the instruments' scores, and the second presenting the findings concerning the student profiles.

Structure of the Instrument Scores

This section reports the results to the first four research questions about the structure, validity, and reliability of the scores of three instruments used in this research and their correlations. For all three tests, factor analyses were used to determine the underlying structure of the test scores. Given that the sample size was sufficiently large, the sample was split, allowing both exploratory and confirmatory factor analyses to be conducted. The scores of BKFT and RTT were interpreted as stand-alone tests, but as the interpretation of one informed the other, they will be reported and discussed in tandem. The results for the belief scale (CMS, Crawford et al., 1998b) follow. At the end of the section, the correlations between the latent variables are presented.

Research Questions 1 and 2

The Basic Knowledge of Functions Test (BKFT) was intended to measure the procedural, declarative, and conditional knowledge of algebraic functions. While there are other ways to distinguish different ways of knowing, these three particular classifications of knowledge are thought to be comprehensive (Alexander, 1991; or the two of declarative and procedural, Anderson, 1996). The BKFT was expected to produce factor scores that either grouped basically according to the three knowledge classifications or along declarative-procedural lines, given that conditional knowledge is exhibited through the application of the other two forms of knowledge (Alexander et al., 1991). The Representational Transfer Test (RTT) was written to measure students' ability to transfer between function representations. In past research transfer questions have grouped according to the source representation (Elia & Gagatsis, 2008), but the RTT measured this skill through multiple choice questions as opposed to free response, and included more representation types. Thus, there was no preconceived expectation of how the scores of these items would contribute to an understanding of the latent constructs predicting the scores. Given the encompassing nature of the declarative-procedural knowledge classification, it was anticipated that the RTT factor descriptions could be related to the same established knowledge classification as used in designing the BKFT.

Content Experts

Two content experts evaluated both the BKFT and RTT. The established knowledge classifications of declarative and procedural were used to interpret both tests. Although the categories of declarative and procedural knowledge are distinct and separable, solving multi-step problems may tap into both forms of knowledge. Some mathematics education researchers postulate that for purposes of classifying mathematics, these types of knowledge exist on a continuum (Rittle-Johnson & Star, 2007). As a consequence, content experts were instructed to classify all items along a five point scale, from “procedural” to “declarative,” with the middle anchor being “both” knowledge types. They were also given the opportunity to comment on the items, but neither made any comments on major issues in the tests or items. The first content expert (CE1) has a PhD in mathematics, is a professor and the head of undergraduate mathematics at a large university. The second content expert (CE2) has an M.S. in mathematics and has been a high school mathematics teacher for 15 years. For the past five years she has taught AP

Calculus. Both are in positions to have a good grasp of the expectations of knowledge of beginning college students' knowledge of functions.

Their ratings were transposed into a quantitative scale by the researcher. Procedural items were scored as -2, declarative items as 2, and items using both types of knowledge as 0. Table 4 provides the rating scale they used. Mean scores were calculated for each test by expert; the RTT received negative scores near zero, indicating the experts viewed the overall test as slightly more procedural than declarative (CE1: $M = -0.375$; CE2: $M = -0.125$). The BKFT received slightly positive scores, thus it was more declarative than procedural (CE1: $M = 0.258$; CE2: $M = 0.700$).

Table 4. Scale for Expert Analysis

Analysis environment	Codes
Expert Analysis	P ------ | ------ B ------ | ------ D
Quantitative Analysis	-2 ----- -1 ------ 0 ------ 1 ------ 2

Agreement between the two experts was not high. Perfect agreement occurred on 36% of the items (6 out of 24 for the RTT; 14 out of 31 for the BKFT). The experts were within one point of agreement on an additional 33% of the items (9 from RTT; 9 from BKFT). There was no consistent pattern of one expert considering items more procedural or more declarative. They were two points apart for all but two of the remaining items. The scores on the two items, one from each instrument, differed by three points. The item from the BKFT (#16, see Figure 5) required students to read the period of a function from the graph. CE1 classified it as procedural (-2), while CE2 sorted it as primarily declarative (1). The RTT item (#9, see Figure 5) required transposing a graph to an equation. CE1 indicated it was declarative (2), yet CE2 considered it primarily procedural

(-1). (All scores by content experts are included in Tables 7 and 9 describing the EFA analyses.)

Figure 5. Items with Three Point Content Expert Disagreement

BKFT item 16

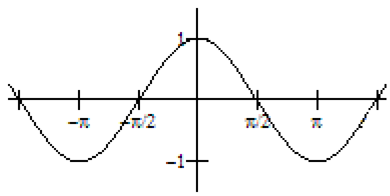
16. What is the period of this function?

a. 1

b. π/2, 3π/2,…

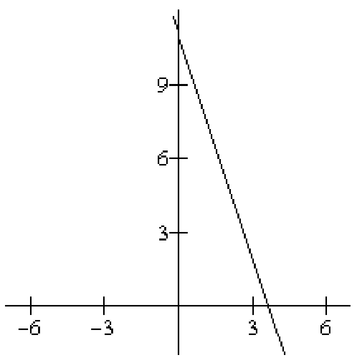
*c. 2π

d. 2



RTT Item 9

9.



*a. $y = 11 - 3x$

b. $y = e^x + 11$

c. $y = 3^x - 11$

d. $y = 11 + x^2$

Parallel Items in the RTT

As each translation type within the RTT is represented twice, once with a linear function and also with a quadratic function, the content of the RTT is designed such that each item has a parallel item within a different function family. As a consequence, there are two items transposing from a graph to equation, one linear and one quadratic. Thus more information about the difference in content expert scoring can be learned through observing the experts' classifications of the two items. See table 5 for a list of all sets of parallel items and their content expert scores.

Table 5. Content Expert Ratings by Parallel Items of RTT

		Linear			Quadratic		
Problem Representation	Solution Representation	Item	CE1	CE2	Item	CE1	CE2
Equation	Graph	13	P	B	18	P	B
Equation	Table	1	P	B	8[a]	P	Pd
Equation	Verbal	3	Dp	Dp	5	Dp	Dp
Graph	Equation	9	D	Pd	22	B	Dp
Graph	Table	11	Pd	Pd	21	P	B
Graph	Verbal	16	D	B	7	Dp	Dp
Table	Equation	20	P	Pd	4	P	Pd
Table	Graph	23	P	B	10	P	Pd
Table	Verbal	6	B	Pd	14	B	B
Verbal	Equation	24	Dp	D	19	Dp	Pd
Verbal	Graph	12	Dp	Dp	15	Dp	B
Verbal	Table	2	Pd	B	17	B	P

Note: CE1=content expert 1; CE2=content expert 2
[a]item not included in further analyses

Item 9 had a three point discrepancy. Its corresponding item (#22) had only a one point discrepancy. Both experts gave different ratings to the two items. This could be the result of the difference between linear or quadratic functions. Comparing other sets of parallel items indicates that a different explanation would be the subjective nature of these classifications. It appears that these types of knowledge are confounded, perhaps even more deeply than those who study declarative and procedural knowledge using these classifications assume. In addition, the classification could be more dependent on the test takers' experience and background knowledge, as opposed to being an objective classification.

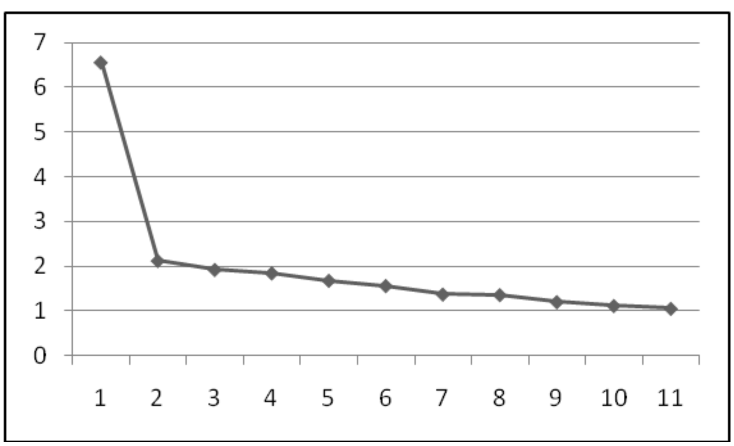
The analysis of the content expert classifications was the first indicator that despite the purpose of the tests, the tests would not break cleanly into the expected factors. The experts were not in high agreement with each other. But when their responses for the parallel items of the RTT were considered, there were many sets of items for which there was low within rater agreement as well. The between rater differences could be credited to their different sources of expertise, but it also provided evidence that classification of knowledge in categories of declarative and procedural is at best an imprecise process. Consequently, the factor analyses were run as planned, but without preconceived notions of the number or nature of factors that might emerge.

Exploratory Factor Analyses

For both the BKFT and RTT, EFAs were conducted on the scores from the exploratory sample using Mplus 4.21 (Muthén & Muthén, 2007) because of its capacity to run analyses with binary data (correct or incorrect). In an EFA analysis, Mplus is designed to extract a range of factor solutions, as specified by the user. Because the tests were intended to measure three types of knowledge, three was the expected number of factors. The results of the content experts resulted in some doubt as to how well this framework would work, so four solutions (with one to four factors respectively) were extracted. This range allowed consideration that a) since all items were about functions they could be described as one factor and b) the presence of more factors than were expected. The results for the BKFT were considered first, then the RTT, and are reported in that order. But the factors from the BKFT were difficult to interpret, and the interpretation of the RTT factors suggested an alternate interpretation of the BKFT, which is presented after the RTT factors. The modified BKFT factors related to the RTT factors were used for CFA and LCA purposes.

Basic Knowledge of Functions Test (BKFT). The 31 multiple choice items were intended to measure the declarative, procedural, and conditional knowledge of functions. The first decision was the appropriate number of factors to extract. Eleven eigenvalues were greater than 1. This value commonly predicts many more factors than are interpretable, particularly in binary data (Costello & Osborne, 2005; Loehlin, 2004). Thus, other criteria were used, such as the scree plot, and the various fit statistics available through Mplus such as the chi-square goodness-of-fit test and the root-mean-square error of approximation (RMSEA). In addition, interpretability of the factors was essential (Muthén, 2002). A scree plot is a created by plotting the eigenvalues (Cattell, 1966). The number of factors can be determined by observing the location of the “elbow” or where the graph abruptly changes direction. The chi-square goodness-of-fit test measures whether the number of factors is sufficient to describe the data, and a non-significant value at an alpha of .05 is desirable (Yu, 2002). The RMSEA is a measure of badness of fit, and measures the amount of error stemming from the approximations. The RMSEA for binary data should be less than .05 (Yu, 2002). The four factor solution was rejected because it returned factor loadings greater than 1. This is an indicator of a Heywood case, which indicates that the solution is not appropriate for the data (Loehlin, 2004; Muthén, 2004; Tomarken & Waller, 2003). Thus, only the one, two, and three factor solutions were considered.

Figure 6. Scree Plot of BKFT Eigenvalues Greater than 1



The scree plot displayed a definite elbow at two factors, indicating that would be the correct number of factors (see Figure 6). The chi-square test indicated that three factors was the correct solution (see Table 6). The RMSEA values indicated that each of the three solutions had sufficient fit. Given the various indices that supported either the two or three factor solution, both were considered for interpretability.

Table 6. Fit Statistics for EFA Factor Solution Models for BKFT

Model	$\chi^2(df)$	*p*-value	RMSEA
1- Factor	198.17 (147)	.003	.034
2- Factor	173.78 (143)	.041	.027
3 - Factor	156.05 (137)	.127	.021

Both solutions were challenging to interpret. But given that chi-square goodness-of-fit tests tend to be overly restrictive and the clarity of the scree plot, the more parsimonious two factor solution was chosen for interpretation. This solution explains 28.0% of the variance. This suggests that although these factors tell some of the story, many details are left untold. Promax rotation was used to interpret the scores. The correlation between the factors was 0.351.

To help in the interpretation, or naming (Everitt et al., 2001), of the factors, each item was classified by a) knowledge type (e.g., declarative, procedural, or both) as assigned by the content experts, b) general topic of the question, c) what representations were used in the problem, d) question type (e.g. recognition, application, etc.), and e) what type of function was used. These classifications were used to determine the meaning of the factors. The following chart contains information about the two factors and those items that did not load onto a specific factor. The cut-off value for a factor loading was

Table 7. Item Descriptions for BKFT by Factor

Item	CE1[a]	CE2[a]	Topic	Representation	Question Type	Function	Loading
				Factor 1 - Visual			
10	B	B	Growth	Graph	Comprehension	Quadratic	0.587
17	Pd	Dp	Inverse	Graph	Comprehension	Cubic	0.585
7	D	D	Periodicity	Verbal	Recognition	Periodic	0.538
9	P	B	Growth	Equation	Recognition	Standard	0.534
16	P	Dp	Periodicity	Graph	Comprehension	Trigonometric	0.532
8	Dp	B	Domain	Graph	Recognition	Standard[b]	0.529
6	D	D	Range	Verbal	Recognition	All[c]	0.515
21	P	P	Inverse	Equation	Calculate	Linear	0.469
4	D	D	Domain	Verbal	Recognition	All[c]	0.453
18	Pd	B	Transposition	Eq-Graph	Calculate	Quadratic	0.408
28	Dp	Dp	Fn types	Graph	Modeling	Exponential	0.347
11	Dp	Dp	Inverse	Table	Comprehension	Standard	0.339
				Factor 2 - Symbolic			
22	P	P	Composition	Equation	Calculate	Lin-Quad	0.686
31	B	Dp	Growth	Eq-Verbal	Modeling	Standard[b]	0.508
15	Dp	Dp	Function Def.	Table	Recognition	Standard[b]	0.504
26	P	M	Graphing	Graph	Modeling	Standard[b]	0.474
14	P	B	Evaluate	Equation	Calculate	Quadratic	0.467
12	Dp	Pd	Inverse	Verbal	Recognition	All[c]	0.459
30	Dp	D	Coefficients	Verbal-Eq	Analysis	Quadratic	0.454
19	Dp	Dp	Symmetry	Graph	Comprehension	Quartic	0.448
13	Pd	Pd	Evaluate	Equation	Calculate	Linear	0.446
29	D	D	Coefficients	Verbal-Eq	Analysis	Quadratic	0.434
5	D	Dp	Function Def.	Verbal	Recognition	All[c]	0.408
				No Factor			
1	D	D	Function Def.	Verbal	Recognition	All[c]	
24	B	Pd	Coefficients	Verbal-Eq	Analysis	Quadratic	
2	D	Dp	Function Def.	Graph	Recognition	Standard[b]	
23	Dp	D	Range	Table	Comprehension	Standard[b]	
3	Dp	Dp	Function Def.	Equation	Recognition	Standard[b]	
27	B	Dp	Graphing	Graph	Modeling	Standard[b]	
25	P	B	Graphing	Graph	Modeling	Standard[b]	
20	D	D	Inverse	Graph	Comprehension	Standard[b]	

Note. CE1 = content expert 1, CE2 = content expert 2.
[a]D = declarative, Dp = declarative with some procedural, B = both, Pd = procedural with some declarative, P = procedural; [b]multiple examples are given that would be commonly encountered in a classroom; [c]no specific function is indicated, true for all functions

set at .30 (Costello & Osborne, 2005). Items that met this criterion for both tests were put with the factor with the highest loading.

As expected from the results of the content expert analysis, the factors were messy. The first thing that became clear was that knowledge type (as expected), function type, and question type were not specifically related to the factors. In addition, the first factor was “cleaner” than the second. The overarching topic of the first factor seemed to be *graphing*. The questions and concepts were those that were needed for and associated with graphs, such as domain and range, growth curves, periodicity, and inverses. These concepts are not exclusively graphing concepts, but graphing, and the tools needed to do it, seemed to be the theme of the first factor. Not all items from the measure that seemed to fit in this factor (given the current description) were here. This basic factor (with some item variations) existed in the three-factor solution as well.

The second factor was less straightforward in its meaning. Factor 2 contained an even mix of the three types of knowledge. Items also varied significantly in their difficulty, from questions based on student misconceptions to the most straightforward calculation questions. It also had questions that dealt with simple mathematical modeling, but also theoretical knowledge of equation coefficients. The most fitting description seemed to be *application* because to succeed on these items, one must apply knowledge of definitions, theory, using equations, and models.

With a rough description of what these factors may represent, the RTT factors were interpreted. That interpretation was more straightforward, and also provided clues to better understand the factors of the BKFT. Therefore, the BKFT factors were reinterpreted. This description is provided after the report of the results of the EFA on the RTT data.

Representational Transfer Test (RTT). The 24 multiple choice items were intended to measure the ability to recognize translations between various functional representations. Seven eigenvalues were greater than 1, but the scree plot again indicated two factors (see Figure 7). The two-factor solution also qualified with appropriate chi-square significance values and RMSEA values, although the fit statistics indicated that the one-factor model could also be correct (see Table 8). Because the primary concern of choosing a solution for an exploratory factor analysis should be interpretability, both the one and two solutions were considered.

Figure 7. Scree Plot of RTT Eigenvalues Greater Than 1

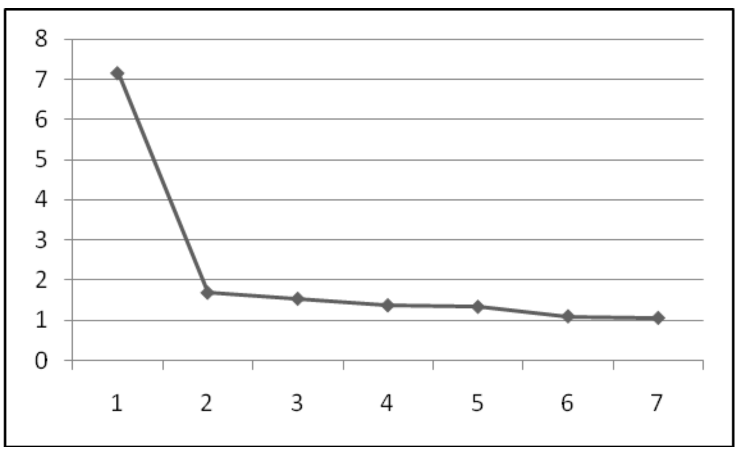


Table 8. Fit Statistics for EFA Factor Solution Models for RTT

Model	$\chi^2(df)$	*p*-value	RMSEA
1- Factor	94.38 (90)	.355	.013
2- Factor	82.20 (85)	.560	.000
3 - Factor	75.77 (88)	.821	.000

The two factor solution was the most descriptive solution, so it is reported here. This factor solution explained 38.4% of the variance. Promax rotation was used to interpret the scores. The correlation between the factors was 0.621.

Table 9 contains information about the items loading onto the two factors as well as the items that did not load onto a specific factor. The cut-off value for a factor loading was set at .30. Each item was classified by a) procedural or declarative classifications, b) problem representation, c) solution representation, and d) type of function used. These classifications were used to determine the meaning of the factors.

There was no influence from the function type. This is reasonable given the assumed familiarity of undergraduate students with linear and quadratic functions. It may have had a stronger effect if less familiar or more complex types of functions were included. In addition, there did not appear to be a pattern of procedural or declarative knowledge types that described either of the factors.

The two factors were best interpreted together. The factors showed evidence of a difference in types of representational transfer, specifically *far transfer* and *near transfer*. This conclusion came from considering the four representations in two basic systems, a visual system for tables and graphs and a semiotic system of words and equations. Items with translation across system (e.g., graph to equation) loaded on the first factor, indicating *far transfer*, while items with translations within systems (e.g., graph to table) loaded on the second factor, indicating *near transfer*. Again, many items did not match the theoretical descriptions of the factors, but these labels seemed to indicate trends of the data. A confirmatory factor analysis was necessary to give support to these possible factors. The analysis and results pertaining the CFA are considered in the next section.

Table 9. Item Descriptions for RTT by Factor

Item	CE1[a]	CE2[a]	Problem Representation	Solution Representation	Function Type	Loading
			Factor 1			
20	P	Pd	Table	Equation	Linear	0.849
4	P	Pd	Table	Equation	Quadratic	0.679
14	B	B	Table	Verbal	Quadratic	0.628
7	Dp	Dp	Graph	Verbal	Quadratic	0.610
1	P	B	Equation	Table	Linear	0.596
6	B	Pd	Table	Verbal	Linear	0.493
22	B	Dp	Graph	Equation	Quadratic	0.443
2	Pd	B	Verbal	Table	Linear	0.437
13	P	B	Equation	Graph	Linear	0.320
			Factor 2			
19	Dp	Pd	Verbal	Equation	Quadratic	0.846
24	Dp	D	Verbal	Equation	Linear	0.608
16	D	B	Graph	Verbal	Linear	0.533
18	P	B	Equation	Graph	Quadratic	0.528
17	B	P	Verbal	Table	Quadratic	0.451
5	Dp	Dp	Equation	Verbal	Quadratic	0.442
9	D	Pd	Graph	Equation	Linear	0.413
21	P	B	Graph	Table	Quadratic	0.410
11	Pd	Pd	Graph	Table	Linear	0.321
			No Factor			
10	P	Pd	Table	Graph	Quadratic	
23	P	B	Table	Graph	Linear	
12	Dp	Dp	Verbal	Graph	Linear	
3	Dp	Dp	Equation	Verbal	Linear	
15	Dp	B	Verbal	Graph	Quadratic	
8[b]	P	Pd	Equation	Table	Quadratic	

Note. CE1 = content expert 1, CE2 = content expert 2.
[a]D = declarative, Dp = declarative with some procedural, B = both, Pd = procedural with some declarative, P = procedural.
[b] item 8 was not included in analysis

BKFT revision. After interpreting the factors of the RTT, it made sense to reconsider the BKFT, using the lens of representations as opposed to type of knowledge. The two factor solution for these items were interpreted as manifestations of the two representation systems, *visual* and *symbolic*. The majority of the items that loaded on the first factor were about graphing and creating representations that are images. The majority of the items loading on the second factors were based more on solving equations, knowledge of theorems, and interpretation of equations.

Discussion of Exploratory Factor Analyses

All factors resulting from the scores of the two instruments were based on representation generally, and specifically the conceptual divide of visual and symbolic representations. The RTT factors were relatively straightforward to interpret while the BKFT factors were not. This is most likely the result of the instrument design. The RTT had a more narrow scope of inquiry than the BKFT. The RTT intended only to look at transformations, so type of function was controlled for (only linear and quadratic were used) to avoid introducing effects based on depth of knowledge of functions. It was also carefully designed to use every representation an equal number of times in the same relative positions. This attention to narrowing the scope of inquiry was not given to the BKFT specifically because of the purpose of the test, which was to measure student knowledge of the very broad subject of functions within the framework of declarative and procedural knowledge. Thus, while care was spent to balance use and type of function and representation, the more important goal was to include items that could measure mastery of basic information across ways of knowing. The instrument was not expected to be based on latent factors of representations. Given the unanticipated results, CFA analyses were required give support to these findings that items were organizing by major categories of mathematical representations.

Confirmatory Factor Analyses

Again, the CFAs for the BKFT and RTT were analyzed following the same procedures. The confirmatory subset of the sample was used for these analyses. Three models were run for each test to ensure the chosen model had the best fit and was maximally descriptive. Results indicated that fit statistics were an artifact of the number of items used in the analysis, as opposed to a measure of goodness-of-fit. Thus, similar items were bundled into component scores, allowing CFAs to be run with scale data, not binary. The results of the initial CFA models, the component formation, and final CFAs are reported below. The results of the BKFT are presented first, and thus the procedures are described in more detail.

Basic Knowledge of Function Test (BKFT). The two-factor model postulated from the EFA was replicated using confirmatory factor analysis. As explained in the methods section, three variations of the model were run (see Table 10). The first model (replication) was a direct replication of the EFA model, using all items that loaded onto the factors in the model, whether they theoretically matched the factor description. The second model (high loading) was postulated to be a cleaner version of the first. For inclusion in the second model, items had to have high loadings (greater than .50), thus ensuring a strong indicator of fit with the factor. The final model (all items) included all items from the instrument. Each item from the BKFT was classified as visual or symbolic, and matched with the appropriate factor. Each model was tested in a CFA using Mplus (Muthén & Muthén, 1998-2007) with the default robust weighted least squares analysis for binary data. Table 10 lists the fit statistics of the models.

For binary data, an appropriate model has a non-significant chi-square goodness-of-fit statistic, a CFI greater than 0.95, and a TLI greater than 0.96. In addition, the RMSEA value would be less than 0.05 (Yu, 2002). The results of these analyses indicated that the first two models exhibited poor fit by all indicators (replication: $\chi^2(153) = 410.4$,

p = 0.000; CFI = .525, TLI = 0.584, RMSEA = .071; high loading: $\chi^2(143)$ = 510.4, p = 0.000; CFI = .322, TLI = 0.365, RMSEA = .088), while the third model had acceptable RMSEA fit (all items: $\chi^2(159)$ = 216.1, p = 0.002; CFI = .895, TLI = 0.911, RMSEA = .033).

Table 10. BKFT Models Run for Verification of Scores

Model	$\chi^2(df)$	p-value	CFI	TLI	RMSEA
Replication[a]	410.37 (153)	0.000	.525	.584	.071
High Loading[b]	510.29 (143)	0.000	.322	.365	.088
All Items[c]	216.08 (159)	0.002	.895	.911	.033

Note. CFI = Comparative Fit Index, TLI = Tucker Lewis Fit Index, RMSEA= Root-Mean-Square Error of Approximation.
[a]Factor 1 items: 4, 6, 7, 8, 9, 10, 11, 16, 17, 18, 21, 28; Factor 2 items: 5, 12, 13, 14, 15, 19, 22, 26, 29, 30, 31; [b]Factor 1 items: 4, 6, 7, 8, 9, 10, 16, 17, 28; Factor 2 items: 5, 13, 14, 22, 29, 30, 31; [c]Factor 1 items: 2, 4, 6, 7, 8, 9, 10, 11 12, 15, 16, 17, 19, 20, 25, 26, 27, 28; Factor 2 items: 1, 3, 5, 13, 14, 18, 21, 22, 23, 24, 29, 30, 31.

The results seemed to indicate that model fit was based on number of items, rather than actual structure of factors. A greater number of observed variables will improve the fit of a model (Tomarken & Waller, 2003). To test this on the specific data of the BKFT, the items were divided randomly into two groups (of same size as the all items model). The fit statistics were basically identical. Given this artifact of number of items, and the difficulty of estimating factors with binary items, factors were confirmed in another manner.

For the second round of CFAs, composite scores were used. For each factor, three components were identified. These components were determined by considering items that loaded onto each factor. There appeared to be topical patterns within the visual and symbolic categories; these were used to form the components. For the visual factor, the components were domain and range, trends of functions, and inverse. For the symbol factor, the components were definitions, evaluation, and modeling. Items that supported the components and had high factor loadings were chosen to form composite scores that were indicated by the latent variables. See Table 11 for the items that were used to form component scores.

Table 11. BKFT Components

Factor	Component	Items
Visual	Domain and Range	4, 6, 8
	Trends	7, 10, 16
	Inverse	11, 17, 21
Symbol	Definitions	3[a], 5, 15
	Evaluation	13, 14, 22
	Modeling	29, 30, 31

Note: [a]Did not load on any factor in EFA.

A CFA was then run. The components were considered categorical data with four categories, thus the weighted least squares estimator was again used (Muthén & Muthén, 1998-2007). The fit statistics indicated excellent fit ($\chi^2(8) = 4.695$, $p = .790$; CFI = 1.000; TLI = 1.019; RMSEA = .000). To ensure that this was not again an artifact of the data, the same items were reorganized into random groups, and that model was run. The random model produced a non-positive definite latent variable matrix, which in this situation was caused by the two factors being statistically indistinguishable (Muthén, 2006). Thus, the random model was not appropriate to describe the data. The non-positive definite matrix indicates that when the items were grouped randomly, no meaningful differences existed in what was being measured. See Figure 8 for the accepted model with non-standardized loadings.

Table 12 contains means and standard deviations for the factors in the BKFT. The reliability of the scores from the *visual* factor was calculated to be .518, and .537 for the scores for the *symbolic* factor. Although these a low reliabilities, only three scores are being used in their calculation, which lowers reliability coefficients (Cortina, 1993).

Figure 8. BKFT Model from CFA Using Composite Scores

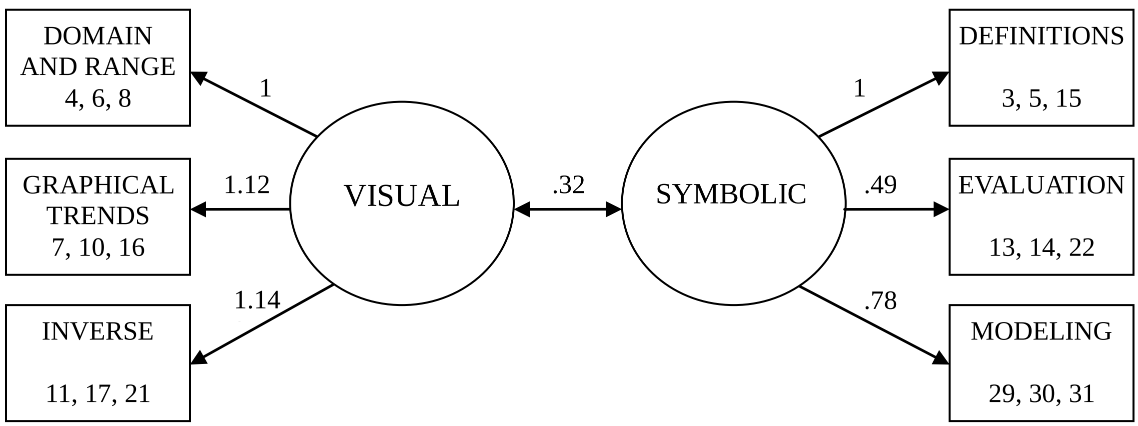


Table 12. Factor and Component Scores for BKFT

Factor Component	Mean	*SD*
Visual[a]	4.02	2.076
Domain and Range[b]	1.46	1.043
Graphical Trends[b]	1.48	0.927
Inverse[b]	1.07	0.931
Symbolic[a]	4.89	1.890
Definitions[b]	0.73	0.748
Evaluation[b]	2.14	0.904
Modeling[b]	2.02	0.957

Note. [a]Maximum score is 9. [b]Maximum score is 3

Representational Transfer Test (RTT). The same process of analyzing three possible models was followed for the RTT. The same models, replicate, high loading, and all items, were postulated. Again, the fit statistics showed, a) only good fit for the all items model, and b) evidence that the fit statistics were an artifact of the number of items as opposed to model fit, see Table 13.

Table 13. RTT Models Run for Verification of Scores

Model	$\chi^2(df)$	p-value	CFI	TLI	RMSEA
Replication[a]	706. 593 (85)	0.000	.358	.426	.148
High Loading[b]	852.857 (83)	0.000	.205	.272	.167
All Items[c]	136.768 (102)	0.012	.964	.973	.032

Note. CFI = Comparative Fit Index, TLI = Tucker Lewis Fit Index, RMSEA= Root-Mean-Square Error of Approximation.
[a]Factor 1 items: 2, 4, 6, 7, 14, 20, 22; Factor 2 items: 5, 9, 11, 17, 18, 19, 21, 23, 24.
[b]Factor 1 items: 2, 4, 6, 7, 14, 20, 22; Factor 2 items: 5, 11, 19, 21 ,23, 24.
[c]Factor 1 items: 1, 2, 4, 6, 7, 9, 12, 13, 14, 15, 16, 17, 18 20, 22; Factor 2 items: 3, 5, 10, 11, 19, 21, 23, 24.

Again, components were then established. Three components were formed for each factor of the RTT. Components were based on the source representation. Four source representations were possible: graph, table, equation, or verbal description. Only three of these were used; the equation source representation was not given a component. There were two reasons for this decision. First, the dropped item had equation as its source representation, making fewer items available for the component. In addition, when the three factor EFA solution was considered, the factors showed evidence of being related to source representation, but only for table, graph and verbal. As a consequence, equation source representations were not included in the component CFA. Thus, the components for both factors were verbal source, graph source, and table source. Each component was balanced to include equal number of linear and quadratic items. Because there were so few near transfer items, only two items were used in each component. See Table 14 for component items.

Fit for this model was good ($\chi^2(7) = 17.594$, $p = 0.014$; CFI = 0.966; TLI = 0.952; RMSEA = .067). In order to check the fit, this model was also contrasted with the same items grouped randomly (similar to the process done with the BKFT). The matrix again was non-positive definite, indicating the two test factors were indistinguishable. See Figure 9 for the model with non-standardized loadings.

Table 14. RTT Components

Factor	Component	Items
Far	Verbal	2, 15[a]
	Graph	16 [a], 22
	Table	14, 20
Near	Verbal	19, 24
	Graph	11 [a], 21
	Table	10 [a], 23

Note: [a]Did not load on any factor in EFA.

Figure 9. RTT Model from CFA Using Composite Scores

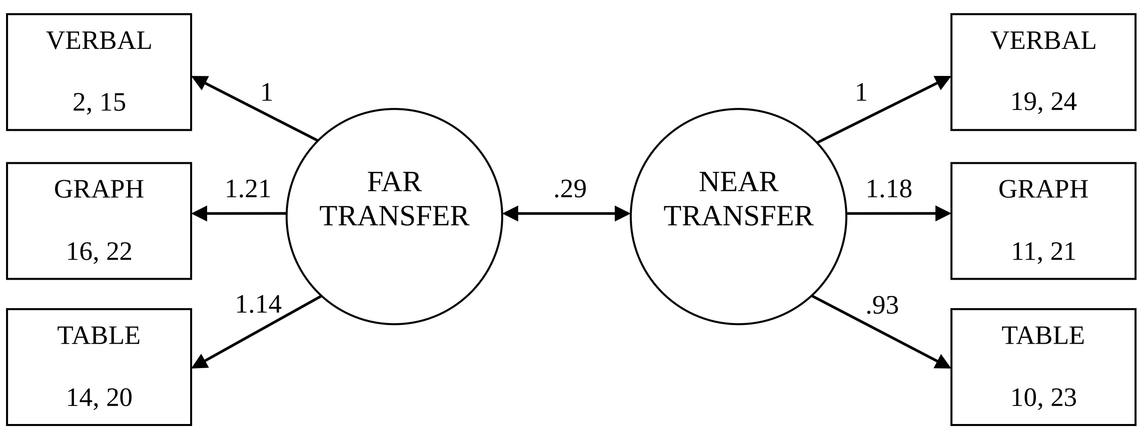


Table 15 contains means and standard deviations for the components and factors of the RTT. Cronbach's alpha for the scores on the Far Transfer factor was .577, and .483

for the scores from the Near Transfer factor. These are low reliabilities, but acceptable given the criterion nature of the test and the small number of items (Cortina, 1993).

Table 15. Factor and Component Scores for RTT

Factor Component	Mean	*SD*
Far Transfer[a]	4.59	1.393
Verbal[b]	1.33	0.657
Graph[b]	1.64	0.606
Table[b]	1.63	0.628
Near Transfer[a]	4.37	1.434
Verbal[b]	1.48	0.667
Graph[b]	1.48	0.683
Table[b]	1.41	0.693

Note. [a]Maximum score is 6. [b]Maximum score is 2

Discussion of Latent Variables from BKFT and RTT

The BKFT and RTT were theorized to measure declarative, procedural, and conditional knowledge of functions and ability to translate between representations. While certainly each of these ways of knowing was represented in the instruments, factor analyses indicated that different latent variables were driving the performance on the tests. Both tests split into two factors each. These latent variables were based not on ways of knowing but on types of representations.

The four representations used in these tests were equations, words, graphs, and tables. These four types divide into two major categories, visual representations (graphs and tables) and symbolic representations (equations and words) (Duval, 2006). Solid mathematical knowledge requires a knowledge of the use and interpretation of the various representations, as well as the ability to move from one representation to another. Some of these connections are well practiced and easy for students to utilize. Others are either less practiced, more difficult, or both, so are done with increased difficulty. Mathematical expertise comes from the understanding that each format is a different expression of the same underlying object. Research shows though, that most students do not arrive at the point of surpassing the representation, but instead continue to see the representation as the mathematical object, not as a depiction of it (O'Callaghan, 1998; Rittle-Johnson & Siegler, 1998; Sfard, 1992). The evidence from these factor analyses are that college students continue to consider functions within representations.

Basic Knowledge of Functions Test (BKFT). The *visual* latent variable is about creating and using visual representations of functions. Some of the concepts held within this variable are those that are used to graph, such as domain, range, trends of graphs, and cycles of graphs. In addition, the concept of inverse is part of this factor. This could indicate that by and large, students understand inverses in terms of their visual appearance, as opposed to any other salient feature. The concepts that match with this factor are those that enable students to properly use and manipulate the visual elements of functions.

The second latent variable, *symbolic*, within the BKFT was less about the visual aspects of functions and more about representations that use algebraic symbol systems, such as equations and descriptions. Calculation, definitions, knowledge of algebraic notation, and modeling with equations were all part of this factor. As opposed to the first factor that is composed of various concepts used when working with the visual system, this latent variable is composed ways of understanding and using the mathematical symbol system.

Although not anticipated from this instrument design, the central role of representation is similar to what researchers have learned about representations in mathematics. Students see various representations as completely different problems, not as variations of the same (Sfard, 1992). Elia and Gagatsis (2008) found a similar representational divide in their study of high school students' knowledge of function. They had 587 high school students take two measures, one transposing graphical representations into verbal and equation representations, and one changing verbal representations into graphical and equation formats. The same relations (three inequalities and three functions) were used on both instruments. They analyzed this data using multiple types of analysis that identify underlying structure of scores, including a hierarchical item clustering method and a CFA. Each analysis showed that the most salient distinction of the problems was the source representation, i.e., graph or verbal, then target representation. The results of the analyses suggested that "the graphic form and the verbal form of the same mathematical content stimulated the use of distinct conversion processes by the students" (Elia & Gagatsis, p. 145).

The study by Elia & Gagatsis (2008) was specifically designed to investigate these distinct representation forms. In contrast, the BKFT was theorized to be a measurement that was representation neutral. This means that many items did not include specific representations, and representations that were used were balanced across knowledge types. Thus it was not projected that representations would still be the salient descriptor of scores. This is further evidence of the compartmentalized nature of functions in mathematics (Sfard, 1992). Visual and symbolic representations may be the primary influence in knowledge of algebraic functions.

Representational Transfer Test (RTT). The factors of the RTT provide additional information about how students conceptualize the various representations. Given the factors from this test, student scores seem to differentiate among distance of representation transfer. The distinction drawn from the factors of the RTT are those of *far transfer* and *near transfer*. The systems of representations are visual, graphs and tables, and symbols, equations and verbal descriptions. Translations within a representation set are classified as *near transfer*. For example, both the graph and the table are depictions of sets of ordered pairs, so that information is easily accessible from both representations without changing format, after initial familiarity is established. A similar near relation exists between equations and words. Words are often the writing out of the symbols, so as familiarity with representations grows, symbols become more of shorthand for the words.

Far transfer measured transpositions across the visual-symbolic divide. These shifts require the movement from ordered pairs to the algorithmic language. As experience with visual and symbolic representations grows this leap is lessened, but it is always a different type of translation. There are fewer points of similarity to match between the representations. Elia and Gagatsis (2008) reported that source representation was the most salient feature in grouping items, paying little attention to the target representation. The results of the RTT factor analysis indicate that another important consideration is the similarity or dissimilarity of the source and target.

Summary of BKFT and RTT Measures

These four factors from the two tests underscore two important points about mathematics learning. The first is that mathematics learning is deeply coupled with the type of representation being used (e.g., Duval, 2006; Elia & Gagastsis, 2008; Even, 1998;

Gagatsis & Shiakalli, 2004; Janvier, 1987). While the foundational ideas of mathematics are representation free, as soon as those ideas need to be communicated, whether amongst mathematicians or in order to teach, representations become essential. Results indicate that these representations become the most salient aspect of the mathematics, overshadowing the underlying abstract ideas.

The other main point from these results is that while ideas of procedural and declarative knowledge are useful for descriptions and understanding, they may not be measurable in the more advanced mathematics addressed in this study. These divisions may in fact be better thought of as descriptive terms for understanding the nature of knowledge, with the realization that rarely is one type of knowledge used in isolation of the other. If one were determined to measure these types of knowledge, much effort would be required to make sure that the items were equivalent across representation and function so that the nature of the types of knowledge would not be overshadowed by other characteristics of the questions.

Research Question 3

The Conceptions of Mathematics Scale (CMS; Crawford et al., 1998b) measures students' ontological beliefs about the field of mathematics. It is intended to measure whether students see mathematics as a fragmented field of unconnected knowledge or a cohesive body of knowledge. This was done by providing students with statements about mathematics (e.g. Mathematics is a subject where you manipulate numbers to solve problems) and the student indicating their agreement on a scale from 1 (strongly disagree) to 5 (strongly agree). The scale was developed and replicated with samples of college students entering mathematics programs of study, so it was important to verify that the measurements still held for a mixed sample of students. This was done by running an exploratory then confirmatory factor analysis.

Factor Analyses

EFAs were run using Mplus 4.21 (Muthén & Muthén, 1998 – 2007), requesting from one to three factors. This allowed consideration of the score structure as reported in the literature, or a collapsing or splitting of a factor. In addition, two methods of estimation were analyzed, both maximum likelihood (ML) estimation and weighted least squares (WLS) estimation. The ML estimation was used to match the analyses found in the literature, but Muthén and Muthén, (1998 – 2007) suggest that ML is inappropriate for Likert-type data, and WLS will better model the data. Outcomes were similar for both analyses, so the weighted least squares results were interpreted to match the analyses of the other two instruments.

The scree-plot indicated that there were two or three factors (eigenvalues greater than 1: 5.809, 4.962, 1.064). The chi-square goodness-of-fit tests were significant for both solutions, while the RMSEA showed sufficient fit (two factor: $\chi^2(62) = 213.64$, $p = .000$, RMSEA = .091; three factor: $\chi^2(57) = 147.51$, $p = .000$, RMSEA = .073). Given these results, both the two and three factor solutions were considered. The two-factor solution matched that from Crawford et al. (1998b). The three factor solution was in fact the two factor solution. No items loaded on the third factor. Thus, the two subscales of *fragmented* and *cohesive* conceptions of mathematics exist in this data as well.

The results of the EFA were confirmed with a CFA using weighted least squares estimation. The fit statistics were not quite in the desired range ($\chi^2(41) = 270.92$, $p = .000$; CFI = .890; TLI = .925; RMSEA = .130). Although not all fit statistics were within the appropriate range, given the results of the EFA and other replications of the structure found in the literature (Alkhateeb, 2001; Arigbabu & Mji, 2005; Mji, 1999) it is clear that these two factors are a good description of the scores resulting from the measure, and apply to this sample in particular (see Table 16).

Table 16. Unstandardized Structure Coefficients for CFA of CMS

Items	1	2
Fragmented Scale		
12. The subject of mathematics deals with numbers, figures and formulae.	1.394	
9. Mathematics is figuring out problems involving numbers.	1.328	
13. Mathematics is about playing around with numbers and working out numerical problems.	1.210	
18. Mathematics is the study of the number system and solving numerical problems.	1.203	
2. Mathematics is a lot of rules and equations.	1.155	
5. Mathematics is about calculations.	1.048	
1. For me, mathematics is the study of numbers.	1	
7. What mathematics is about is finding answers through the use of numbers and formulae.	0.944	
16. Mathematics is a subject where you manipulate numbers to solve problems.	0.942	
4. Mathematics is simply an overcomplication of addition and subtraction.	0.663	
Cohesive Scale		
17. Mathematics is a logical system which helps explain the things around us.		1.381
8. I think mathematics provides an insight into the complexities of our reality.		1.321
10. Mathematics is a theoretical framework describing reality with the aim of helping us understand the world.		1.318
19. Mathematics is models which have been devised over years to help explain, answer and investigate matters in the world.		1.231
14. Mathematics uses logical structures to solve and explain real life problems.		1.212
11. Mathematics is like a universal language which allows people to communicate and understand the universe.		1.209
6. Mathematics is a set of logical systems which have been developed to explain the world and relationships in it.		1.041
3. By using mathematics we can generate new knowledge.		1
15. What mathematics is about is formulae and applying them to everyday life and situations.		0.875

For the complete sample, the reliability of the scores from the fragmented scale was .81, and that of the scores on the coherent scale was .89. The mean of the fragmented scale was 37.88 (*SD* = 5.10). The cohesive scale mean was 32.39 (*SD* = 6.52). As originally reported by Crawford et al. (1998b), the subscales are independent of each other. The correlation between the variables was –.068 (p = .087). This signifies that the subscales measure different concepts, not just opposite ends of one concept. The ways of conceiving mathematics were organized by Crawford et al. as a hierarchy, meaning that ideas that belong to the fragmented scale could conceivably still be held by someone who also conceptualizes mathematics as a cohesive unit.

Discussion of CMS

This belief scale returned two independent factors, as reported in the literature (Crawford et al., 1998a, 1998b). The scales measured whether students see mathematics as a fragmented set of unconnected facts or a cohesive body of knowledge used to solve problems and understand the world. It is expected that those who have more experience with mathematics and better understanding of it will see the links among the various concepts, and recognize the connected nature of the discipline. This should correspond with higher scores on the BKFT and RTT.

The lack of association between the two subscales is worth noting. Although there are participants who score high on one measure and low on the other, as would be expected, there are also many who indicated high levels of endorsement on both, or neither, measures. It may be that while some students have advanced sufficiently in their perceptions of mathematics that they recognize its utility, they have not released their earlier conceptions. High scores on the cohesive scale have been shown to associate with positive learning outcomes (e.g., study habits, grades; Crawford et al., 1998a). It would be expected that these factors relate to higher scores on the BKFT and RTT as well.

Research Question 4

This section answers the research question of how the scales of the instruments are related. In order to investigate this, factor scores were totaled for the entire sample, and then Pearson Correlations were calculated. See Table 17 for these correlations. All correlations were significant at the .05 level except the one between the fragmented and cohesive factors of the CMS.

Table 17. Correlations among Latent Variables

	1	2	3	4	5
1. Visual (BKFT)	--				
2. Symbolic (BKFT)	.433**	--			
3. Far Transfer (RTT)	.351**	.493**	--		
4. Near Transfer (RTT)	.316**	.435**	.453**	--	
5. Fragmented (CMS)	-.220**	-.165**	-.162**	-.091*	--
6. Cohesive (CMS)	.273**	.271**	.213**	.165**	-.068

**p < .01 * p< .05

The correlations between the cognitive scales are medium (Cohen, 1988), ranging from .316 - .493. This is reasonable given that the underlying content of all items are mathematical functions as initially formally encountered. It also indicates that these factors are not measures of the same thing, but represent distinguishable areas of knowledge. The cohesive scale is positively related to all the cognitive measures with small correlations, while the fragmented scale is negatively related. This provides additional evidence to the claim that better mathematics knowledge corresponds with more recognition of the connected nature of the discipline.

Student Profiles

This section responds to the final two research questions which investigate the existence and nature of latent profiles within the data. Latent class analysis (LCA) can be applied to variables or participants; in this research it was used to establish profiles of student groups. The profiles were then described using both the instrument scores and other collected demographic data. These analyses were conducted using the full sample.

Research Question 5 and 6

The underlying premise of LCA is that within a population, there are homogeneous subgroups that differ from other groups in their scores across the measurements. The analysis is conducted with a specific number of classes to be found within the data, then maximum likelihood estimation is used to predict the likelihood that any member of a class has a particular score for some variable (Garson, 2008). This allows for calculations of probabilities that an individual's score trajectory belongs to each class. A participant is then classified into the most likely group. LCA is a population based technique, which gives more confidence in the results than a traditional cluster analysis (Magidson & Vermunt, 2002). Initially, the one class model is tested to verify the assumption that multiple groups exist. Then additional groups are iteratively added until the best-fitting model is found. Once the appropriate number of classes is established, they can be described using both the factors employed in the analysis and other external criteria.

Latent Class Analysis

Mplus 5.2 (Muthén & Muthén, 1998 – 2009) was used for this analysis. The four factors from the BKFT and RTT along with the two subscales from the CMS were used as the defining variables. Twenty participants were missing scores on one or more clustering variables and were not classified, thus the sample size was 620. The one-class model was tested first. Each subsequent model (containing additional classes) was judged against the fit of the previous models. Fit is determined by entropy, information criteria, and the Lo-Mendell-Rubin test (LMR). Entropy is a measure of the reliability of the factors, and this value should be above .7 for a stable solution (Schwartz & Zamboanga, 2009). Three information criteria, Akaike's (AIC), Baysien (BIC) and sample size adjusted Baysien (adj-BIC), help determine the fit of the model. The solution with the best number of classes will have minimum IC values (Yang et al., 2005). Some changes in IC values are small, and the question arises as to whether the change is significant. To respond to this concern, the Lo-Mendell-Rubin test was developed. This statistical test measures whether the addition of a class better describes the data. The distribution of the test statistic is known, allowing for a significance (p) value with the test. This test was found be the most accurate in correctly identifying the accurate number of classes in simulated data (Nylund et al., 2007)

For this research, models with up to five classes were run. All models had sufficiently high entropy scores. The five class model had the lowest BIC scores, while the four-class model had the lowest AIC score. Quality of fit as measured by the LMR test indicated that the increase in classes past four was non-significant, thus the four class model was chosen (see table 18).

Description of Classes

Means by class and overall are presented in Table 19. As an overall set (total), students earned the lowest scores on the visual factor, then the symbolic factor. Students performed better on both far and near transfer items than on the BKFT. Both the transfer scores had similar mean difficulties.

Table 18. Fit of Solutions with 1 – 4 Classes

# classes	AIC	BIC	adj-BIC	Entropy	LMR
1	16958.052	17037.786	16980.639	n/a	n/a
2	16798.697	16909.440	16830.069	.851	169.587, *p* = .000
3	16666.835	16808.586	16706.991	.788	142.692, *p* = .000
4	16604.225	16776.984	16653.166	.833	74.945, *p* = .000
5	16572.260	16776.027	16629.985	.783	44.966, *p* = .295

Note: AIC = Akaike's Information Criterion, BIC = Baysiean Information Criterion, adj-BIC = sample size adjusted Baysiean Information Criterion, LMR=Lo-Mendell Rubin Likelihood Measure.

Table 19. Factor Scores by Class

		M (*SD*)				
		1	2	3	4	Total
Factor	Max	(n = 56)	(n = 179)	(n = 372)	(n = 113)	(n = 640)
Visual	9	2.79 (1.44)	3.34 (1.55)	3.36 (1.32)	7.27 (1.21)	4.01 (2.08)
Symbol	9	3.00 (1.50)	4.16 (1.84)	5.00 (1.56)	6.73 (1.13)	4.89 (1.89)
Far	6	1.39 (0.65)	3.65 (0.48)	5.43 (0.50)	5.65 (0.48)	4.59 (1.40)
Near	6	3.11 (1.34)	3.77 (1.46)	4.62 (1.27)	5.33 (0.87)	4.36 (1.43)
Fragmented	50	39.09 (4.40)	38.91 (4.43)	38.41 (4.77)	34.49 (5.78)	37.90 (5.10)
Cohesive	45	30.02 (6.49)	30.84 (7.35)	32.25 (5.83)	36.50 (4.56)	32.41 (6.50)

Class 4 (n = 113), *high performing*, was defined by their high scores across all measures. They found the transfer items to be easier than the BKFT scales, with their best items being far transfer. Class 3 (n = 272) was the *high average* group. Among the four measures, they were most successful on far transfer, and least successful on the visual items. Interestingly, class 2 (n = 179) scored almost identically to class 3 on the visual factor, but then scored lower on the other three indicators, thus they are *low average*. Class 2 still found transfer items to be easier than BKFT items, especially near transfer items. Class 1 (n = 56), *low performing*, clearly struggled with the concepts measured here. They scored approximately 30% on all scales except near transfer, in which they got a little over half correct.

The MANOVA conducted on the classes and factor scores yielded a significant multivariate effect, Wilks' $\lambda = .058$, $F(18, 1729) = 166.45$, $p < .001$, partial-$\eta^2 = .612$, indicating that the class differences were significant. Univariate effects are listed in Table 20. Tukey post-hoc tests were run to find the nature of the differences. All classes were differentiated on all measures except for the visual factor. For this factor, groups 2 and 3 had similar scores.

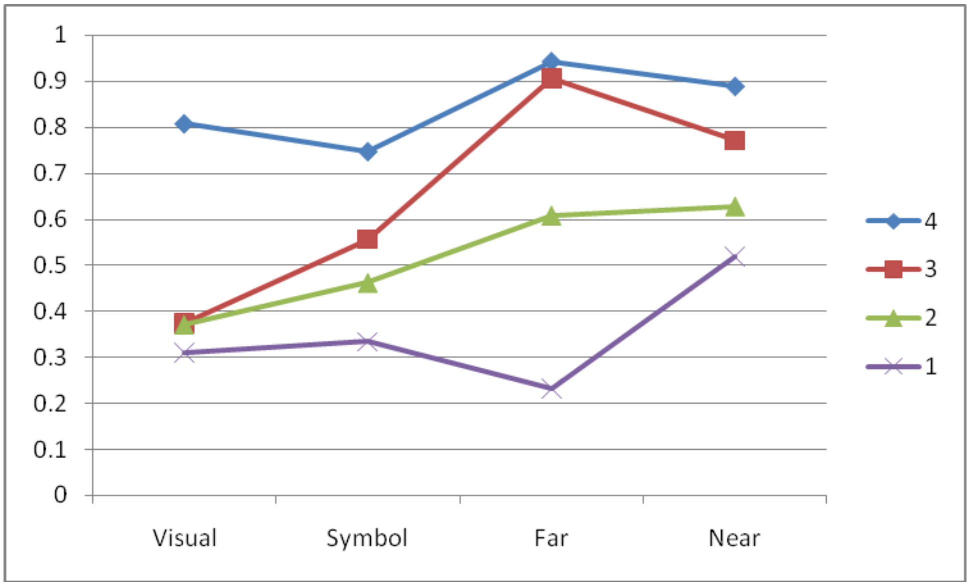
Table 20. Univariate Results for Class Differences across Factor Scores

Measure	*F*-ratio	*p*	partial-η^2
Visual	256.867	.000	.556
Symbolic	91.192	.000	.308
Far Transfer	1379.403	.000	.870
Near Transfer	56.544	.000	.216
Fragmented	23.387	.000	.102
Cohesive	23.238	.000	.102

Note. Degrees of Freedom for all tests: (3, 616)

The primary distinction between classes is total score. Class 1 had the lowest scores, then class 2, class 3, with class 4 performing best. But each class had a unique score trajectory, which can be seen in Figure 10. This figure shows percent correct on the factors of the BKFT and RTT by class. Clearly class membership indicated level of knowledge on these concepts, but each class had individual strengths and weaknesses within its ranking of high, average, or low.

Figure 10. Percentage Scores by Class on Four Factors from BKFT and RTT



On the belief measure, classes 1, 2, and 3 had similar profiles. Although their actual scores were significantly different, all had the overall pattern of higher fragmented scores and lower cohesive scores. As math expertise grew, so did scores on the cohesive scale. Class 4 (high performing) indicated a higher endorsement of cohesive views than fragmented views.

Relations with Other Variables

This section investigates class membership by their relationships with other variables of interest. Appendix G describes each class by demographic characteristics such as gender and course. Specifically to answer research question 6 though, class differences were considered across the variables SAT scores, GPA, whether calculus was taken, grades in university calculus courses, and the score on item 32 from the BKFT. This question presented a split-domain function in equation format, and asked students to a) identify if it was a function, and b) graph the function.

ANOVAs were run to determine if differences existed in SAT scores, GPA scores, and calculus grades. Significant differences were found in for SAT scores (Math: $F(3, 522) = 65.57$, $p < .0005$; Verbal: $F(3, 519) = 6.09$, $p < .0005$), GPA scores ($F(3, 568) = 9.73$, $p < .0005$) and college calculus grades ($F(3, 164) = 5.84$, $p = .001$). Tukey post-hoc tests were then run to find which groups had significantly different scores.

For SAT-math scores, all groups were significantly different, in the order expected. For SAT-verbal scores, class 1 scored significantly lower than class 3 and 4, class 2 scored significantly lower than class 4, as did class 1 in comparison to class 3. There were in effect two subsets, the high scoring classes 3 and 4 and the low scoring classes 1 and 2.

In GPA scores, Class 1 had significantly lower GPA scores than all other classes. There were no significant differences among the other classes. For calculus grades, class 4 had significantly higher scores than classes 2 and 3. The difference between class 4 and 1 was larger than any other group, although it was not significantly different. This is most likely an artifact of sample size. See table 21 for the class scores on these variables.

Table 21. Means and Standard Deviation Scores for External Measures by Class

Measure	1	2	3	4
SAT-M	520.22 (71.38)	560.43 (67.63)	600.73 (72.51)	669.48 (70.70)
SAT-V	547.61 (93.86)	565.17 (65.05)	579.83 (69.36)	596.04 (80.27)
GPA	3.12 (0.45)	3.34 (0.37)	3.40 (0.37)	3.42 (0.38)
Calculus Grade[a]	2.50 (1.29)	2.80 (0.76)	2.88 (0.81)	3.34 (0.72)

Note: SAT-M=SAT mathematics score, SAT-V = SAT verbal score, GPA= grade point average, Calculus Grade is for university calculus course
[a]2 = C, 3 = B, 4 = A. Class 1 n = 4, Class 2 n = 25, Class 3 n = 69, Class 4 n = 70

Chi-square tests of association were used to determine if there were significant differences in whether students took calculus and their success on the free-response item from the BKFT. Although this item was originally scored in two parts, recognition of the function, and a correctly drawn graph, for this analysis it was scored as correct or incorrect.

As expected, the chi-square was significant for both measures. For students taking calculus, the chi-square value was 99.59, $p < .0005$, with 3 *df*. Almost 94% of students in class 4 took calculus in high school or calculus. Approximately 62% of class 3, and 22% of class 2 took this class. Only 4% of class 1 took calculus. For the performance on item 32, about 30% of the students in class 4 scored correctly on this item. Only 4%, .5%, and 2% of students in the other classes (3, 2, and 1 respectively) got this question correct ($\chi^2(3) = 97.43$, $p < .0005$)

Discussion of Classes

Four groups were identified through the LCA, a high performing group, a high average group, a low average group, and a low performing group. As noted in Wang et al. (2005), a major difference between the groups was of a quantitative nature; the divisions were based on magnitude of scores. But there were also qualitative differences that were evident. Each group had different strengths and weaknesses within its primary ranking. Although the scores could be listed in order from highest to lowest, each group still had a distinct profile indicating areas of strength and weakness within their score trajectory.

From considering the four classes together, it appears that success comes first in the area of near transfer. The next factor that is learned is far transfer, which then becomes more familiar than the near transfer. Once these skills are honed, the more general knowledge of the calculations, definitions, and applications from the BKFT comes. It appears that to see mathematics as a coherent body of information, mastery must be achieved in all areas.

The relation of quantitative differences with class ranking applied to more than just the measures taken in this study. In addition, class membership could be used to order students in magnitude of SAT scores and GPAs.

CHAPTER 5

CONCLUSIONS, IMPLICATIONS AND LIMITATIONS

The purpose of this dissertation was to validate the scores of two measures of knowledge of functions. The findings from those instruments were then used to describe college students' knowledge of mathematical functions. The two instruments were intended to measure general knowledge of functions as defined declaratively, procedurally, and conditionally (BKFT), and to measure a more specific ability, that of translating from one representation to another (RTT). The findings from these instruments were complemented by a measure of student beliefs on the nature of mathematics as being either fragmented or cohesive. This study reports initial findings on the validity and reliability of the scores of these instruments and their structures as identified through factor analysis. In addition, profiles of student knowledge patterns were also identified and described using LCA. In doing so, the following research questions were answered:

1) To what extent does the factor structure of the Basic Function Knowledge Test (BFKT) support the validity and reliability of scores?
2) To what extent does the factor structure of the Representational Transfer Test (RTT) support the validity and reliability of scores?
3) Does the factor structure of the Conceptions of Mathematics Scale (CMS) match that of the literature? If not, what structure does seem to exist? Are these scores reliable?
4) What relationships exist between the scales from the RTT, BFKT, and CMS?

5) Given the scales of the three tests, can scale scores be used to form profiles of student knowledge of functions?
6) To what degree are profiles related to important academic performance criteria such as SAT mathematics scores and grades in mathematics courses?

This chapter begins with a review of the instruments and relevant background information about the measures. The factors underlying the tests and the profile composites of the participants are then reported, with discussions of how this informs understanding of student knowledge of function. This dissertation concludes with a discussion of the limitations of this study and implications for future research.

Instrumentation and Procedure

The first measure in this study was designed to measure functions using declarative and procedural knowledge. The Basic Knowledge of Functions Test (BKFT) did this through 31 multiple choice questions that were intended to measure beginning facts, definitions, and early procedures that a student starting calculus would be expected to know. In addition to declarative and procedural knowledge, this first measure also included some items intended to measure conditional knowledge (Alexander et al., 1991). These knowledge types and their application to this measurement are reviewed below.

The term knowledge is broadly defined, so measurement of knowledge requires the concept to be divided into smaller, specifically defined categories. One major division is that of knowing verbally or knowing procedurally (e.g., Anderson, 1996; Hiebert, 1986; Ryle, 1958). When conditional knowledge, knowing when to apply knowledge, is included these three types of knowledge are said to underlie all ways of knowing (Alexander et al., 1991).

For this study, verbal knowledge has various descriptions depending on the field of study and its scope. For example, educational psychologists use the term “declarative knowledge” which usually indicates knowledge that can be stated, for example definitions, terms, laws, simple relations, or directions. Declarative knowledge can be described as “knowing that” (Ryle, 1958). Mathematics education researchers though use the more expansive of term of “conceptual knowledge” that encompasses not only what can be stated, but also how those statements correspond and interact (Rittle-Johnson & Star, 2007). Conceptual knowledge represents a web of knowledge that includes both nodes of information and relations between those nodes. It is often characterized as being a rich, deep, and connected knowledge of a topic (Hiebert & Lefevre, 1986). Although it is contrasted with procedural knowledge, many definitions also include procedures as being part of the conceptual knowledge.

Verbal knowledge was defined as declarative knowledge because the distinct categories of declarative, procedural, and conditional knowledge underlie all other types of knowledge (Alexander et al., 1991). It was expected that the understanding of the three ways of knowing would illustrate aspects of conceptual knowledge of mathematical functions.

Knowledge of how to do tasks is more straightforward. Procedural knowledge is “knowing how” (Ryle, 1958), and is the ability to do something, whether it be the physical work of riding a bike, or the mental work of solving an equation with one unknown (Anderson, 1996). Although procedural knowledge has sometimes been defined as shallow (Hiebert & Lefevre, 1986), if it is drawn into a conceptual web then it can be deep and connected (Star, 2005). Procedural knowledge of functions is essential for anyone who must apply and use knowledge of functions in further mathematics or other fields.

Declarative knowledge was measured through items questions about definitions and recognition of definitions. Procedural knowledge was measured through calculations and other mathematical procedures. Conditional knowledge is primarily manifest through declarative or procedural responses. In this study, the application problems, those that require students to formulate their method of solution were considered conditional knowledge.

While declarative and procedural knowledge were expected to provide a framework for students' knowledge, ignoring the role of representations in this measurement would have overlooked an essential part of students' knowledge of functions. Functions require facility with multiple representations and an understanding of the relations between those representations (e.g., Lesh et al., 1987; O'Callaghan, 1998). To truly understand functions, students must understand that different representations are just different ways of communicating the same ideas. The use of the word "translate" to describe the movement from one representation to another is apt, because they truly are two ways of saying the same thing, just as in language translation (e.g., Elia & Gagatsis, 2008, Janvier, 1987).

Representational translation was measured through the Representation Transfer Test (RTT), using 24 multiple-choice items. Linear or quadratic functions were presented in one representation of four representations, graphical, tabular, algorithmic (equation), or verbal. The correct translation was to be selected from the answer choices.

The two knowledge measures as briefly described, plus an ontological belief measure of the participants' conceptions of the field of mathematics, were administered to 640 undergraduate students. These students came from a variety of majors, and all four years of schooling. They seemed to be a fairly representative sample of the students that attend their university. Their scores on the two tests were submitted to exploratory and confirmatory factor analyses in order to verify the tests measured the intended constructs. Once factors were established, latent class analysis was conducted using the factor scores as grouping variables.

Factors and Classes

Factor analysis was used to determine the structure of the scores of the tests of knowledge and beliefs. Two factors were identified for each test. The factors of the BKFT are *visual* and *symbolic* representations. The transfer factors of the RTT are *far transfer* and *near transfer*. The beliefs instrument contains the *fragmented* and *cohesive* scales as identified by the authors (Crawford et al., 1998b). The LCA conducted with these factors returned four major classes of students: low performers, low average, high average, and high performers. These factors and classes are described in more detail below.

Latent Factor Results

The BKFT has two latent variables, neither corresponding to knowledge factors. Instead they correspond to the major categories of representations: visual and symbolic. The first factor consists of items that use visual representations, such as graphs and tables, or appertained to understanding those representations. This factor was confirmed through three subsets of questions, those about domain and range, about graphical trends, and inverses. The second factor contains items that focused on algebraic symbol notations. Subsets of items about calculation items, interpretation of coefficient items, and definition of function items were used to confirm this factor.

The division of the RTT items is also defined by the visual-symbolic distinction, but the factors did not simply group according to type of representation used, but instead are distinguished by the type of translation that occurred. Translations between representations from the same visual or symbolic family are distinguished from translations that crossed over the boundary. Thus, the two factors are far transfer, in which translations occur across systems, and near transfer, translating within the visual or symbolic system. Alternate terms, according to Paivio's dual coding theory (1991), would be referential and associative translations.

For the beliefs measure, the two factors put forward in the literature (Crawford et al., 1998b) exist in these data as well. It was important to verify these factors because in previous research, the population had been mathematics students. The sample in this research came from many fields of study, so the results may not have been consistent. The two factors are a fragmented conception of mathematics and a cohesive view of the field. A fragmented view is defined as seeing mathematics as arbitrary rules and procedures. A cohesive view sees mathematics as containing connected information useful for solving problems and understanding phenomena. A fragmented view is less sophisticated than the cohesive view that comes with further experience or expertise. Given the descriptions of the factors, it seems that these two factors would be opposite ends of a continuum, and be negatively correlated, but empirically that is not true. These two factors are in fact conceptually distinct, as illustrated by the small negative correlation between the two factors (r = -.09). Thus many students do perceive mathematics as being both fragmented and cohesive, or having neither qualities.

Latent Factor Discussion

Although the format and design of both tests were dissimilar, the four factors from the BKFT and RTT are based on representations. Both tests evidence a divide between visual and symbolic (or verbal) representations. The presence of this distinction, exhibited in different manners, in both tests, indicates that representations hold a critical role in characterizing the mathematics knowledge of students.

The latent factors did not describe knowledge of functions as expected. When this study was conceptualized, the focus was on ways of knowing functions, specifically declarative, procedural, and conditional knowledge. These ways of knowing were to be complemented by an additional look at how students were able to translate between the various representations used in mathematics, which knowledge was assumed to be subsumed within the declarative, procedural, and conditional knowledge framework. The results did not support this knowledge framework, but instead illustrated that mathematics is a field that is dominated by its representations. This became less a study about how people know functions and became a study about how students know representations, using functions as the vehicle for that exploration.

Mathematics is a particularly apt field for this study because of the unique relation that mathematics has with representations. "[Mathematics is an] abstraction – ideas based on experience but independent of any particular experience. Communication about [mathematics] therefore, requires some form of external representations" (NRC, 2001, p. 2). The principles and ideas that constitute mathematics cannot be measured empirically, they exist in thought. However, a person's *knowledge* of those ideas can be measured. A mathematician chooses the representation that is most useful to work with these thoughts and to express them. Often times it is a verbal description. It could also be a diagram or graph, a table, or an equation. The representation is chosen to enable the best use. An expert mathematician looks for elegance and simplicity, and will choose the best medium for that, while recognizing that there are other representations would allow for different understanding. Multiple representations are used when needed or useful.

In contrast, students, or non-mathematicians, understand representations not as means for communication or aids to finding solutions, but as the purpose of mathematics. While they recognize connections between the representations, the underlying concepts are not recognized as being the central tenets of mathematics. In multiple reports on what students know about functions, the highest level of knowledge was postulated to be reification (e.g., O'Callaghan, 1998; Sfard, 1992) in which the representations cease to be the mathematics, but instead are the vehicle to communicate the mathematics. The results from this study suggest that at this level of student knowledge of mathematics, representations still have a tremendous influence.

The visual-symbolic division is logical given the nature of representations. This split is accepted as a natural break in the field of representations (Duval, 2006; NCTM, 2000; NRC, 2001), and as a natural division in the way brains function (Paivio, 1991). Yet this divide was not expected given the nature of the BKFT. The test was designed to measure declarative and procedural knowledge, with representation types being balanced throughout the questions as well as possible. It may be that the classification of knowledge into declarative or procedural is not empirically meaningful at this higher level of mathematics. Research distinguishing those types of knowledge in mathematics is primarily conducted at the elementary and early middle-school level (Leinhardt et al., 1990; Rittle-Johnson & Siegler, 1998) and may not be applicable for older students. At higher levels of mathematics it is conceivable that the split cannot take into account the fundamentally connected nature of both types of knowledge (Baroody et al., 2007; Star, 2005). It is also possible than in the attempts to control for representations effect, it was instead highlighted and thus became the salient factor. A test specifically designed to recognize these ways of knowing (such as the BKFT) would have to be carefully organized in order to truly measure these categories of knowledge.

If representation is the essence of how students know mathematics, then they must take a central role in the education of students. Major documents on mathematics teaching make this point (NCTM, 2000; NRC, 2001) but it is one focus among many. This study indicates that this should be a major, if not the primary, focus in how mathematics understanding is conceptualized.

Belief in Mathematics

Literature both within the fields of mathematics and in other fields demonstrate that beliefs relate to academic performance (e.g. Murphy, 2007; Schoenfeld, 1989; Zeruth, 2006). Yet in this study, beliefs failed to substantially add to the understanding of the classes of participants, even correlations between knowledge and belief measures were small, across all profiles (absolute values ranging from 0 to .30). There are many possible reasons for this. The next few paragraphs will discuss potential measurement explanations.

Measurement issues of both knowledge of mathematics and beliefs may have obscured the relations of the constructs. The knowledge measures were modeled after typical textbook problems; consequently the problems were low affect. Given what may have been perceived as non-engaging material and the research nature of the study (i.e., performance did not result in a grade), it is likely that students felt no need to attend to the problems past the minimum effort required to solve them. As such, it is unlikely that student beliefs were engaged during the application of knowledge. Yet depth of understanding requires an affective engagement of beliefs of the topic at hand as well as the cognitive knowledge, thus to measure understanding in mathematics this relation between knowledge and beliefs should be activated. This could be done by designing an instrument that was intended to be more engaging to students and through measuring engagement in the solving process. This may provide an enhanced sense of the nature of the relation of knowledge and beliefs.

Another possible constraint to better understanding the relation between knowledge and beliefs were the belief measures employed in this research. The two measures applied were intended to measure beliefs about the nature of mathematics and beliefs about the nature of knowledge. The CMS scale was designed to measure fragmented or cohesive views about mathematics. As mentioned earlier, these two factors are conceptually distinct, which seems counterintuitive to the descriptions of the factors. Alternate descriptions of the factors make the lack of correlation more viable, but change the interpretations of the latent variables. Possible alternate descriptions of the factors would be *mathematics as numbers* and *mathematics as systems*, which are both accurate conceptions of mathematics, albeit one view is more limited than the other. The conceptions that match with mathematics as numbers are not false conceptions; they are just not the whole of mathematics. Although this alternate interpretation explains the lack of correlation, it provides little insight as to the plausible engagement of students' beliefs with their knowledge performance. It seems that this measure may not be informative for understanding students' beliefs about doing mathematics.

The second beliefs scale was the Teachers Epistemic Beliefs Instrument (Hennessey, 2007) which was not analyzed in the results. This instrument was intended to measure beliefs about learning, whether it is foundationalist (built on core truths), coherentist (understood through connected truths), or relativist (based on experiments and personally verified truth). The instrument was intended for teachers, but was modified to measure beliefs of students. Correlations between the three scales were large (between .60 – .65; Cohen, 1988) indicating that students perceived much overlap between the three sets of epistemic beliefs. A factor analysis indicated that the items did not load onto a three-factor structure as designed. This could be a result of items that did not measure the epistemic belief frameworks, or a loss of psychometric properties from the instrument modifications or for the new population (students vs. teachers).

Beliefs impact actions (Mandler, 1989), and this is certainly the case within mathematics (Schoenfeld, 1989). Murphy (2007) posits that a hallmark of expertise is believing in one's subject, which suggests a strong affective engagement. Studying knowledge and beliefs together is vital to understanding mathematical achievement and performance (Schoenfeld). Yet to properly gauge this relationship and the effects of beliefs on knowledge, different measurements will be required.

Latent Class Results

Latent class analysis and other clustering methods are intended to find homongeneous subgroups within the population that are in fact separate groups with individual means and variances. The initial division is primarily one of score, those who score high on measures and those who score low (Yang et al., 2006). With enough classes (which require enough participants) other, more qualitative differences, may emerge. Both quantitative and qualitative differences occur in the four classes identified in this study. Simply put, they are low performers, low average, high average, and high performers. Scores on the mathematics measures corresponds with conceptions of mathematics, such that students with high scores view mathematics as a more cohesive field of study while those with low scores view it as more fragmented. These classifications apply to outward measures as well, in terms of SAT scores, GPAs, and grades in calculus courses.

Yet within the ascending ranks of scores, there are individual class profiles of relations between the measures, the qualitative differences. The three lower classes have very similar scores on the visual factor, but then have very different scores on the other three factors. The visual factor is evidently made of the most difficult questions. The lowest performing class did equally poorly on all factors except near transfer, where they performed much better. The low average group had a pattern of ascending scores on the four factors, doing poorest on the visual factor, then the symbolic factor, then far transfer, then near transfer. The high average class had a very mixed profile, where they did best at far transfer, almost equaling the high class, but did worst on the visual factor, just barely outscoring the lower two groups. The high performing students did well on everything, but were better at transfer items as opposed to the knowledge items.

Latent Class Discussion

It was expected that there would be a high and low performing class that were distinguished primarily by scores. But it was also anticipated that the addition of classes would introduce some variation in factor strengths and weaknesses. While this happened to some extent, even at four classes, the division was quantitative. This may be the result of the interrelated nature of the factors; they are all moderately correlated. As a result, improvement in one factor is accompanied by improvement in another. Another way to view the factors though is as a hypothesized growth model.

This study is cross-sectional, thus no conclusions can be drawn about the order in which these topics are mastered. But many possibilities can be conjectured, if the assumption is made that all students follow basically the same learning trajectory to mathematical competence. Under that assumption, these classes can be seen as initial, intermediate and competent stages. The first topic to be mastered is near transfer. Then far transfer is mastered, until it is better than near. The most practiced translation tasks for students is equation to graph (often using the table as a midpoint during the instructional process), which is in the far transfer set, so it is not surprising that this would be their best scale. Finally the more theoretical and applied portions (within the visual and symbolic frameworks) of mathematics are learned. This suggests that procedures of translating among functions are mastered before knowledge of theorems, procedures, and applications are.

As scores increase among the cognitive factors, cohesive conception scores increase and fragmented scores decrease. Although the high performing group is the only class where cohesive scores are higher than fragmented scores, there is also less difference between the two scores here. It suggests that they have only recently gained the requisite mathematics knowledge to see the field as being related, but are not sufficiently convinced to have released their fragmented conceptions. It is also possible that this group of students, being overall high achievers have been told that mathematics is a coherent body of knowledge useful for solving problems, and indicated the “right” answers, as well as endorsing their actual beliefs.

When related to external criterion, the profiles also highlighted the connection between mathematics knowledge and other indicators of academic success. Not surprisingly, the high performing group had the highest calculus grades and math SAT scores. But, they also had the highest verbal SAT scores and GPAs. As measured here, function knowledge is clearly part of a constellation of measures of success at the university level.

Limitations

Although, this study has clearly identified the central role of representations in the study of mathematical functions, there are limitations to this study. The first is that the sample comes from a large-northeastern land grant university that is highly ranked for academic achievement. In addition this sample was 80% female, although students were recruited from both education and science courses. Whether the results from this population are generalizable to other populations remains an empirical question, and efforts should be made in future research to access other populations.

In addition, the concept of functions is vast and varied. Providing a test that measures all aspects of knowing of functions would be prohibitively long for research or classroom purposes. While selection of topics to be measured was made with careful consideration, some mathematicians may feel that other information is more central to the concept.

Another significant limitation is the lack of a picture of expert knowledge in this subject. Studies about knowledge about functions are primarily conducted on high school and college students, and occasionally their teachers of mathematics. These studies (including the present one) show that student knowledge is characterized by representations used to communicate the mathematics. This division of understanding by representations often leads to the conclusion that students do not integrate between the representations. There are other possible explanations though. A different interpretation is that type of representation is the most salient quality of all mathematics. Perhaps even if one can easily move among the representations (i.e., an expert), this may still emerge as the key facet of knowledge. As a picture of student knowledge is developed, a corresponding expert picture could provide direction as to what student knowledge should be. This expert picture could not be measured by the BKFT or RTT, as they would be too easy for the mathematicians. But it would be helpful to devise an instrument that could define the role of representation for experts in mathematics.

This lack of expert group also handicaps the understanding of the CMS. Although fragmented conceptions and cohesive conceptions form distinct factors for undergraduates (both those intending to study mathematics and others), given the descriptions of these items, it seems unlikely that a mathematician could hold both views simultaneously as do many of the students in this study. It is hard to claim what mathematics instruction and beliefs should be without having a better understanding of the abilities and conceptions of the experts.

Another limitation is the restricted types of functions considered in the RTT. Only linear and quadratic functions were used. While this choice was deliberate, it is entirely plausible, and even likely, that a very different picture would have emerged had more types of functions been included. To get a better sense of this, two versions of the instrument would need to be investigated. The first would be a similar instrument with different function types, perhaps exponential and logarithmic. In addition, a longer instrument would need to be tested that used many types of functions. Although type of representation is a salient feature, function type may be more so. Other research has indicated that results from studies of functions and their representations may not be generalizable across types of functions and may only apply to type for which it was measured (Even, 1998; Meltzer, 2004). It is possible that the transfer factors found with the linear and quadratic functions are not applicable to other types of functions.

Future Directions

Despite these limitations, the results provide many interesting conclusions and directions for future research. First, this study should be replicated with a cleaner BKFT instrument, one that builds on the factors found here, and is more careful in balancing representations, function type, and difficulty. The meaning of the components used in the confirmatory analysis should be further investigated and included. A large administration of this type was necessary to investigate the design of the instrument. Now that there are more clear ideas about what is important, a tighter version of the test that more substantially measures the components and factors could be designed, including a version for the experts. It was intended that these instruments serve as diagnostic tools for teachers of relevant mathematics. More study into the meaning and nature of the factors, and how they can be more directly measured must happen first.

In addition, it is important to continue learning about how representations influence learning about functions. One question that should be asked is how understandings of the two major representation systems differ. For example, within a specific representational system, can knowledge then be defined as declarative, procedural and conditional? Or do students understand symbolic systems differently than visual ones? What characterizes these understandings? From this research it appears that functions cannot be defined without representations, but theoretically functions exist beyond representations. Learning more about the composition of knowledge of functions and representations as defined by these latent variables may help identify where functions cease to be representations, and become the underlying construct.

Another important question for research is what happens with more complex functions. It is important to investigate whether representation is still the most prominent feature, or does function type become more salient? It is possible that representation is allowed so much influence because students are basically familiar with the functions used in these measures, and with unfamiliar functions that different results would occur. The interplay between function type and representation is an issue that requires additional research.

An important part of improving student knowledge of function is defining expert knowledge. A comparison of student knowledge with expert knowledge could inform expectations about what knowledge students should possess. Do they need to be able to recognize a function, or just use functions? Many items that required application of knowledge showed basic student mastery, as opposed to items measuring knowledge of formal definitions which were not answered correctly. This indicates that possessing the correct definition is not necessarily related to correctly using functions. Questions of appropriate content for mathematics courses cannot be answered through research, but having an expert picture of this would be provide clues to features of expertise.

Conclusion

The overarching purpose of this dissertation was to validate two tests of function knowledge, and use those results to describe university students' knowledge of functions. Two untested instruments were utilized to accomplish this purpose. The validation process was done through factor analysis. From this analysis, the structures of the scores of the tests were identified. The BKFT was hypothesized to measure declarative, procedural, and conditional knowledge, but instead measured visual and symbolic representations. The RTT was designed to measure students' ability to translate between functions. It did indeed measure this ability by categorizing transfer ability in terms of difference in representation types. In essence, what these instruments demonstrated is that representations are the key to mathematical comprehension of functions. Beliefs were also measured, but because of measurement and implementation issues failed to provide additional understanding about function knowledge. Different tactics should be used in future measurements.

From the factors, profiles of students were created. These profiles suggested that near transfer items, those that require translation between representations that are either both visual or both symbolic, are mastered by most students. Far transfer items, which translate across system types, come next. Students exhibit less proficiency with the more factual and procedural items that are measured in the BKFT, with symbolic items being less complex than visual items. More expertise on functions tends to go with higher SAT scores and higher GPAs.

The results of this research provide evidence that representations may be the most salient factor in teaching and learning mathematics. This is of vital importance to teacher training. This implies that representations should be taught and retaught throughout the teaching of mathematics.

The importance of representations is not unexpected. The NCTM Standards (2000), which is the primary reference for curriculum and policy decisions concerning mathematics in the USA, recognizes that representations are a key element of instruction, as it is one of the ten critical standards for learning and instruction. But given the role of representations in this study, it is troublesome that representation is listed as the last NCTM standard. This could be construed as an implication that it is the least important standard. Even if that is not the intent, given its location, it will most likely receive less attention than the others, as there is never sufficient time in teacher training to focus on all important material. This study indicates that representations may be a superordinate standard, perhaps one from which all other standards are taught.

A statement within the explanation of the importance of representation may provide an indication as to why representations are not given appropriate attention by teachers: “The fact that representations are such effective tools may obscure how difficult it was to develop them, and more important, how much work it takes to understand them”

(NCTM, 2000, p. 67). It is likely that teachers of mathematics understood representations easily as students, and thus expect that all students will similarly “get” representations. Yet this is not the case. Many studies show that students do not understand representations, and especially are unable to move between representations (e.g., Huntley et al., 2000). The results of this study indicate that students foundational understanding is through representations, implying that a focus on representations is essential in the training of mathematics teachers, and then as they teach students how to do mathematics. How to use and connect representations should be taught explicitly multiple times in each student’s mathematical experience.

This study provides a better description of how students currently know functions, but much more is needed. Mathematical competence is seen as very important for all university students, and functions are a vital part of that competence. “Those who understand and can do mathematics will have significantly enhanced opportunities and option for shaping their futures” (NCTM, 2000, p. 5). This study provides a framework for understanding student knowledge of functions, and opens up many possibilities for future research that will enable better measurement and knowledge of functions in the future.

References

Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education, 33*, 131 – 152.

Alexander, P. A., Shallert, D. L., & Hare, V. C. (1991). Coming to terms: How researchers in learning and literacy talk about knowledge. *Review of Educational Research, 61*, 315 – 343.

Alkhateeb, H. M. (2001). University students' conceptions of first-year mathematics. *Psychological Reports, 89*, 41 – 47.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review, 111*, 1036 – 1060.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89*, 369 – 406.

Anderson, J. R. (1996). ACT: A simple theory of cognition. *American Psychologist, 51*, 355 – 365.

Arigbabu, A. A., & Mji, A. (2005). Nigerian undergraduate education majors' conceptions of mathematics. *Psychological Reports, 96*, 273 – 274.

Baroody, A. J., Feil, Y., & Johnson, A. R. (2007). An alternative reconceptualization of procedural and conceptual knowledge. *Journal for Research in Mathematics Education, 38*, 115 – 131.

Borgen, F. H., & Barnett, D. C. (1987). Applying cluster analysis in counseling psychology research. *Journal of Counseling Psychology, 34*, 456 – 468.

Browne, M. W. (2001). An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research, 36*, 111 – 150.

Buehl, M. & Alexander, P. A. (2005). Motivation and performance differences in students' domain-specific epistemological belief profiles. *American Educational Research Journal, 42*, 697 – 726.

Byrnes, J. P., & Wasik, B. A. (1991). Role of conceptual knowledge in mathematical procedural learning. *Developmental Psychology, 27,* 777 – 786.

Carpenter, T. P. (1986). Conceptual knowledge as a foundation for procedural knowledge. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 113 – 132). Hillsdale, NJ: Laurence Erlbaum Associates.

Carpenter, T. P., Lindquist, M. M., Matthews, W., & Silver, E. A. (1983). Results of the third NAEP mathematics assessment: Secondary school. *Mathematics Teacher, 76*, 652 – 659.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, *1,* 245 – 276.

Cattell, R. B., & Vogelmann, S. (1977). A comprehensive trial of the scree and KG criteria for determining the number of factors. *Multivariate Behavioral Research, 12*, 289 – 325.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*, 98 – 104.

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most out of your analysis. *Practical Assessment Research & Evaluation, 10,* Article 7. Retrieved May 15, 2009 from http://pareonline.net/pdf/v10n7.pdf.

Crawford, K., Gordon, S., Nicholas, J., & Prosser, M. (1994). Conceptions of mathematics and how it is learned: The perspectives of students entering university. *Learning and Instruction, 4*, 331 – 345.

Crawford, K., Gordon, S., Nicholas, J., & Prosser, M. (1998a). Qualitatively different experiences of learning mathematics at university. *Learning and Instruction, 8*, 455 – 468.

Crawford, K., Gordon, S., Nicholas, J, & Prosser, M. (1998b). University mathematics students' conceptions of mathematics. *Studies in Higher Education, 23*, 87 – 94.

Darlington, R. B. (1997). *Factor Analysis*. Available at http://www.psych.cornell.edu/Darlington/factor.htm.

Der, G., & Everitt, B. S. (2002). *A handbook of statistical analyses using SAS* (2nd ed.). New York: Chapman & Hall.

Dubinsky, E., & Harel, G. (1992). The nature of the process conception of function. In G. Harel & E. Dubinsky (Eds.), *The concept of function: Aspects of epistemology and pedagogy, MAA Notes, Number 25* (pp. 85-106). Washington, DC: Mathematical Association of America.

Dufour-Janvier, B., Bednarz, N., & Belanger, M. (1986). Pedagogical considerations concerning the problem of representation. In C. Janvier (Ed.) *Problems of representation in the teaching and learning of mathematics* (pp. 109 – 122). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Dugopolski, M. (2003). *College Algebra* (3rd ed.). Boston: Addison-Wesley.

Duval, R. (2006). A cognitive analysis of problems of comprehension in a learning of mathematics. *Educational Studies in Mathematics, 61*, 103 – 131.

Eisenberg, T. (1992). On the development of a sense for functions. In G. Harel & E. Dubinsky (Eds.), *The concept of function: Aspects of epistemology and pedagogy, MAA Notes, Number 25* (pp. 153 – 174). Washington, DC: Mathematical Association of America.

Elia, I., & Gagatsis, A. (2008). A comparison between the hierarchical clustering of variables, implicative statistical analysis and confirmatory factor analysis. *Studies in Computational Intelligence, 127*, 131 – 162.

Elia, I., Panaoura, A., Eracleous, A., & Gagatsis, A. (2007). Relations between secondary pupils' conceptions about functions and problem solving in different representations. *International Journal of Science and Mathematics Education, 5*, 533 – 556.

Elia, I., & Spyrou, P. (2006). How students conceive function: A triarchic conceptual-semiotic model of the understanding of a complex concept. *The Montana Mathematics Enthusiast, 3*, 256 – 272.

Evangelidou, A., Spyrou, P., Elia, I., Gagatsis, A. (2004). University students' conceptions of functions. *Proceedings of the 28th conference of the International Group for the Psychology of Mathematics.* Norway: PME Conference.

Even, R. (1998). Factors involved in linking representations of functions. *Journal of Mathematical Behavior, 17*, 105 – 121.

Everitt, B. S., Landau, S., & Leese, M. (2001). *Cluster analysis* (4th ed.). New York: Oxford University Press.

Farnham-Diggory, S. (1994). Paradigms of knowledge and instruction. *Review of*

Educational Research, 64, 463 – 477.

Feinn, R., & Kulikowich, J. M. (1999). *Representational Transfer Test*. Storrs, CT: University of Connecticut.

Ferrini-Mundi, J., & Graham, K. (1994) Research in calculus learning: Understanding of limits, derivatives, and integrals. In J. J. Kaput & E. Dubinsky (Eds.) *Research issues in undergraduate mathematical learning: Preliminary analyses and results* (pp. 31 – 45). Washington, DC: Mathematical Association of America.

Gabrieli, J. D. E. (1998). Cognitive science of human memory. *Annual Review of Psychology, 49*, 87 – 115.

Gagatsis, A., & Shiakalli, M. (2004). Ability to translate from one representation of the concept of function to another and mathematical problem solving. *Educational Psychology, 24*, 645 – 657.

Garson, G. D. (2008). *Latent Class Analysis, class webpage.* Available at http://faculty.chass.ncsu.edu/garson/PA765/latclass.htm.

Gentner, D. & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist, 52*, 45-56.

Gill, G. (n.d.) *The calculus bible*. Online Text. Retrieved May 27, 2009 from http://www.math.byu.edu/Math/CalculusBible/.

Glass, G. V., & Hopkins, K. D. (1996). *Statistical methods in education and psychology* (3rd ed.). Needham Heights, MA: Allyn & Bacon.

Harel G., & Dubinsky, E. (Eds.). (1992). *The concept of function: Aspects of epistemology and pedagogy, MAA Notes, Number 25* . Washington, DC: Mathematical Association of America.

Hartigan, J. A. (1996). Introduction. In P. Arabie, L. J Hubert, & G. De Soete (Eds.), *Clustering and Classification* (pp. 1 – 4). River Edge, NJ: World Scientific Publishers.

Hennessey, M. E., (2007). *Teacher epistemic beliefs: The development of a psychometrically sound measure.* Pennsylvania State University. Unpublished dissertation.

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

Hiebert, J., & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 1 – 28). Hillsdale, NJ: L. Erlbaum Associates.

Higley, K., Higley, M., & Kulikowich, J. M. (2005). *Basic Knowledge of Functions Test*. University Park, PA: The Pennsylvania State University.

Higley, K., Litzinger, T., Van Meter, P., Masters, C. B., & Kulikowich, J. M. (2007). Effects of conceptual understanding, math and visualization skills on problem-solving in statics. *Proceedings of 2007 American Society of Engineering Education Annual Conference and Exposition*. Honolulu, HI.

Holyoak, K. J., & Thagard, P. (1997). The analogical mind. *American Psychologist*, *52*, 35 – 44.

Hughes-Hallett, D., McCallum, W. G., Gleason, A. M., et al. (2002). *Calculus* (3rd ed.). New York: John Wiley & Sons, Inc.

Huntley, M. A., Rasmussen, C. L., Villarubi, R. S., Sangtong, J., & Fey, J. T. (2000). Effects of Standards-based mathematics education: A study of the Core-Plus

Mathematics Project algebra and functions Strand. *Journal for Research in Mathematics Education*, *31*, 328 – 361.

Janvier, C. (1987). Translation processes in mathematics education. In C. Janvier (Ed.), *Problems of representation in the teaching and learning of mathematics* (pp. 27 – 32). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20,* 141 – 151.

Kaput, J. J. (1989). Linking representations in the symbol systems of algebra. In S. Wagner & C. Keil, F., (Eds.), *Concepts, kinds, and cognitive development*. Cambridge, MA: MIT Press.

Kline, R. B. (2006). Principles and practice of structural equation modeling (2nd ed.). New York City: Guilford.

Koedinger, K. R., Anderson, J. R., Hadley, W. H., & Mark, M. A. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence, 8*, 30 – 43.

Kyriakides, L. (2002). A research-based model for the development of policy on baseline assessment. *British Educational Research Journal*, *28,* 805 – 826.

Larkin, J., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11*, 65 – 99.

LeFevre, J., Smith-Chant, B. L., Fast, L., Skwarchuk, S., Sargla, E. Arnup, J. S., Penner-Wilger, M., Bisanz, J., & Kamawar, D. (2006). What counts as knowing? The development of conceptual and procedural knowledge of counting from kindergarten through Grade 2. *Journal of Experimental Child Psychology, 93*, 285

– 303.

Leinhardt, G., Zaslavsky, O., & Stein, M. K. (1990). Functions, graphs, and graphing: Tasks, learning, and teaching. *Review of Educational Research, 60*, 1 – 64.

Lesh, R., Behr, M., & Post, T. (1987). Rational Number Relations and Proportions. In C. Janiver (Ed.), *Problems of representations in the teaching and learning of mathematics* (pp. 41 – 58). Hillsdale, NJ: Lawrence Erlbaum.

Loehlin, J. C. (2004). *An introduction to factor, path, and structural equation analysis*. Mahwah, NJ: Guilford.

Magidson, J., & Vermunt, J. K. (2002). Latent class models for clustering: A comparison with *K*-means. *Canadian Journal of Market Research, 20*, 37 – 44.

Mandler, G. (1989). Affect and learning: Causes and consequences of emotional interactions. In D. B. McLeod & V. M. Adams (Eds.), *Affect and mathematical problem solving: A new perspective* (pp. 3 – 19). New York: Springer-Verlag.

Mayer, R. E., & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology, 86*, 389 – 401.

Meltzer, D. (2005). Relation between students' problem solving performance and representational format. *American Journal of Physics*, *73*, 463 – 478.

Meuller, G., & Brent, R. I. (2006). *Just-in-time: Algebra and trigonometry for early transcental calculus* (3rd ed.). New York: Pearson-Addison Wesley.

Miller, S. N. (2005). *Teaching college algebra: Reversing the effects of social promotion*. Lanham, MD: Rowman & Littlefield Education.

Mji, A. (1999). Reliability and validity of the Conceptions of Mathematics questionnaire.

Psychological Reports, 85, 579 – 582.

Mji, A., & Klaas, E. N. (2001). Psychometric characteristics of the conceptions of mathematics questionnaire. *Psychological Reports, 88,* 825 – 831.

Monk, S. (1992). Students' understanding of a function given by a physical model. In G. Harel & E. Dubinsky (Eds.), *The concept of function: Aspects of epistemology and pedagogy, MAA Notes, Number 25* (pp. 175 – 193). Mathematical Association of America.

Murphy, P. K. (2007). The eye of the beholder: The interplay of social and cognitive component in change. *Educational Psychologist, 42,* 41 – 53.

Muthén, L. K. (2002, April 4). Mplus discussion > Exploratory factor analysis > Binary data and factor analysis. Message posted to http://www.statmodel.com/discussion/messages/8/50.html#POST962 [importance of factor interpretability].

Muthén, L. K. (2004, March 23). Mplus discussion > Exploratory factor analysis > Binary data and factor analysis. Message posted to http://www.statmodel.com/discussion/messages/8/50.html#POST2417. [issues with Heywood cases].

Muthén, L. K. (2006, May 12). Mplus discussion > Confirmatory factor analysis > Confirmatory factor analysis. Message posted to http://www.statmodel.com/discussion/messages/9/61.html#POST8504. [not positive definite matrix].

Muthén, L. K. & Muthén, B. (1998 – 2007). *Mplus User's Guide*. Fourth Edition. Los Angeles, CA: Muthén & Muthén.

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

National Research Council [NRC]. (1989). *Everybody counts: A report to the nation about the future of mathematics education.* Washington, DC: National Academy Press.

National Research Council [NRC]. (2001). *Adding it up: Helping children learn mathematics.* J. Kilpatrick, J. Swafford, and B. Findell (Eds.). Mathematics Learning Study Committee, Center for Education, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling, 14*, 535 – 569.

O'Callaghan, B. R. (1998). Computer-Intensive Algebra and students' conceptual knowledge of functions. *Journal for Research in Mathematics Education, 29*, 21 – 40.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology, 45*, 255 – 287.

Pintrich, P. R., Marx, R. W., & Boyle, R. A. (1993). Beyond cold conceptual change: The role of motivational beliefs and classroom contextual factors in the process of conceptual change. *Review of Educational Research, 63*, 167 – 199.

Ritter, S., Anderson, J. R., Koedinger, K. R., & Corbett, A. (2007). Cognitive Tutor: Applied research in mathematics education. *Psychonomic Bulletin & Review, 14*, 249 – 255.

Rittle-Johnson, B., & Siegler, R. S. (1998). The relation between conceptual and procedural knowledge in learning mathematics: A review. In C. Donlan (Ed.) *The development of mathematical skills* (pp. 75 – 110). East Sussex, UK: Psychology Press Ltd.

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

Rittle-Johnson, B., & Star, J. R. (2007). Does comparing solution methods facilitate conceptual and procedural knowledge? An experimental study on learning to solve equations. *Journal of Educational Psychology, 99*, 561 – 574.

Ryle, G. (1958). Knowing how and knowing that. In I. Scheffler (Ed.), *Philosophy and education*. Boston: Allyn and Bacon, Inc.

Schoenfeld, A. H. (1989). Exploration of students' mathematical beliefs and behaviors. *Journal for Research in Mathematics Education, 20*, 338 – 355.

Schwartz, S. J., & Zamboanga, B. L. (2008). Testing Berry's model of acculturation: A confirmatory latent class approach. *Cultural Diversity and Ethnic Minority Psychology, 14*, 275 – 285.

Seufert, T. (2003). Supporting coherence formation in learning from multiple representations. *Learning and Instruction, 13*, 227 – 237.

Sfard, A. (1992). Operational origins of mathematical objects and the quandary of reification – The case of function. In G. Harel & E. Dubinsky (Eds.), *The concept of function: Aspects of epistemology and pedagogy, MAA Notes, Number 25* (pp. 59 – 84). Washington D.C.: Mathematical Association of America.

Sierpenska, A. (1992). On understanding the notion of function. In G. Harel & E. Dubinsky (Eds.), *The concept of function: Aspects of epistemology and pedagogy, MAA Notes, Number 25* (pp. 25 – 58). Washington, D.C.: Mathematical Association of America.

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

Steinley, D. (2003). Local optima in *K*-means clustering: What you don't know may hurt you. *Psychological Methods, 8*, 294 – 304.

Tabachneck-Schijf, H. J. M., Leonardo, A. M., & Simon, H. A. (1997). CaMeRa: A computational model of multiple respresentations. *Cognitive Science, 21*, 305 – 350.

Tomarken, A. J., & Waller, N. G. (2003). Potential problems with "well-fitting" models. *Journal of Abnormal Psychology, 112*, 578 – 598.

Tran, U. S., & Formann, A. K. (2009). Performance of parallel analysis in retrieving unidimensionality in the presence of binary data. *Educational and Psychological Measurement, 69*, 50 – 61.

Ubersax, J. (2008). LCA Frequently Asked Questions. Retrieved March 6, 2009, from http://ourworld.compuserve.com/homepages/jsuebersax/faq.htm#ident.

Ullman, M. T. (2004).Contributions of memory circuits to language: The declarative/procedural model. *Cognition, 92*, 231 – 270.

van Streun, A. (2000). Representations in applying functions. *International Journal of Mathematical Education in Science and Technology, 31*, 703 – 725.

Vermunt, J. K., & Magidson, J. (2002). Latent class cluster analysis. In J. A. Hagennars &

A. L. McCutcheon (Eds.), *Applied Latent Class Analysis* (pp. 89 – 106). Cambridge: Cambridge University Press.

Vermunt, J. K., & Magidson, J. (2003). Latent Class Analysis. In *The SAGE Encyclopedia of Social Science Research Methods* (Vol II, pp. 549 – 553). Thousand Oaks, CA: Sage Publications.

Vinner, S. (1983). Concept definition, concept image and the notion of function. *International Journal of Mathematical Education in Science and Technology, 14*, 293 – 305.

Vinner, S., & Dreyfus, T. (1989). Images and definitions for the concept of function. *Journal for the Research in Mathematics Education, 20*, 356 – 366.

Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association, 58,* 236 – 244.

Wikipedia (2009). *Functions (mathematics)*. Retrieved May 15, 2009 from http://en.wikipedia.org/wiki/Function_(mathematics)

Wu, H. (1996). Invited comments on the *NCTM Standards*. Retrieved May 27, 2009 from http://math.berkeley.edu/~wu/inv-paulo.pdf.

Yang, X., Shaftel, J., Glasnapp, D., & Poggio, J. (2006). Qualitative or quantitative differences?: Latent class analysis of mathematical ability for special education students. *Journal of Special Education, 38*, 194 – 207.

Yu, C. (2002). *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes.* University of California – Los Angeles. Unpublished dissertation. Available online at http://www.statmodel.com/papers_date.shtml.

Zazkis, R., & Liljedahl, P. (2004). Understanding primes: The role of representation. *Journal for Research in Mathematics Education*, *35*, 164 – 186.

Zeruth, J. A., Kulikowich, J. M., & Murphy, P. A. (2006, April). *Knowledge, efficacy, and beliefs: Studying emergent expertise in probability*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, California.

Zwick, W. R., & Velicer, W. F. (1984). Comparison of five results for determining the number of components to retain. *Psychological Bulletin, 99*, 432 – 442.

Appendix A

Basic Knowledge of Functions Test

Instructions (printed in header at top of each page):

These questions are intended to measure knowledge that you currently possess. ***Please circle the best response***. If you do not remember how to solve the problem, guessing is acceptable.

(See next page for test.)

1. If y is a function of x, then

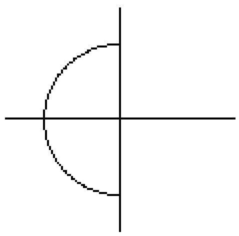
a. the values assigned to y determine the values assigned to x.

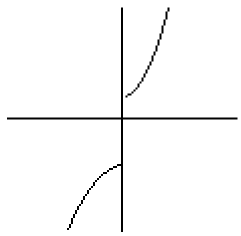
b. the values of both x and y predict one another.

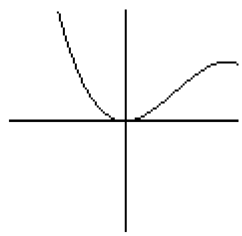
c. each value of x is corresponds to a unique y value.

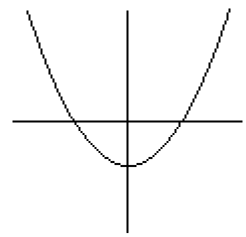
d. y and x are sets of numbers.

2. Which of the following does ***not*** represent a function?

a. 

b. 

c. 

d. 

3. Which of these equations defines ***y*** as a function of ***x***?

a. $(y+2)^2 = x$

b. $(x-2)^2 + (y-1)^2 = 9$

c. $y^3 + 3 = x$

d. $x = 3$

4. What is the domain of a function?

a. the set of possible values of the independent variable

b. the set of values for the dependent variable

c. the values of x chosen for graphing purposes

d. the arithmetic process used to find the dependent values

5. Which of the following must be true for all functions?

a. There must be an equation that defines it.

b. It must be continuous.

c. An equation must exist defining one variable by the other.

d. It is a mapping from one set to another set.

6. What is the range of a function?

a. the values assigned to the independent variable

b. a determination of its continuity

c. the set of possible values of the dependent variable

d. all possible values for either variable

7. What does it mean for a function to be periodic?

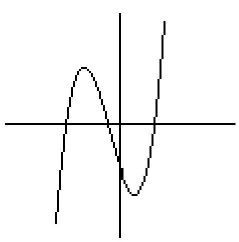
a. the function cannot be defined by an equation

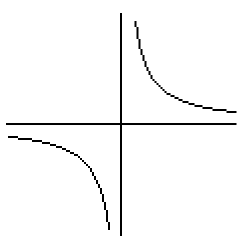
b. the graph of the function repeats itself on regular intervals

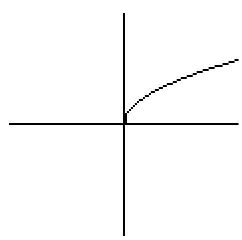
c. the function abruptly ends at one point, and begins somewhere different

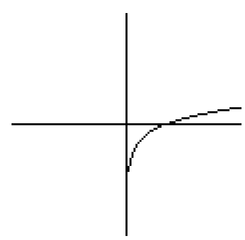
d. the graph of the function is decreasing

8. Which function has a domain of all real numbers?

a. 

b. 

c. 

d. 

9. As x approaches infinity, which function grows fastest?

a. $f(x) = 500x^2$

b. $f(x) = \cos x$

c. $f(x) = 2^x$

d. $f(x) = \ln x$

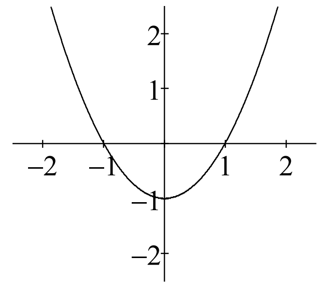
10. The following function is

a. increasing on the interval (- ∞, + ∞)

b. increasing on the interval (- ∞, 0)

c. increasing on the interval (0, ∞)

d. increasing on the interval (-1, ∞)



11. Which of the following is the inverse of the function described by the following set of values?

x	***y***
-2	¼
0	1
3	8
7	124

a.

x	***y***
¼	-2
1	0
8	3
124	7

b.

x	***y***
124	-2
8	0
1	3
¼	7

c.

x	***y***
7	124
3	8
0	1
-2	¼

d.

x	***y***
2	-¼
0	-1
-3	-8
-7	-124

12. If f(x) and g(x) are inverse functions defined on the real numbers, which of the following is NOT always true?

a. f(g(x)) = x, for x in the domain of g.

b. g(x) is a one-to-one function

c. the graphs of f(x) and g(x) are symmetric about the line y = x.

d. the domain of f(x) and g(x) are equivalent

13. Find $f(5)$ if $f(x) = \begin{cases} 2x; & for\ all\ x < 0 \\ x + 4\ ; & for\ all\ x \geq 0 \end{cases}$

a. 10

b. 19

c. 5

d. 9

14. Find f(3) if $f(x) = 2x^2 + 4$.

a. 13

b. 40

c. 10

d. 22

15. Determine which set of values does **not** give *y* as a function of *x*?

a.

x	***y***
2	0
4	1
6	0
8	1

b.

x	***y***
1	9
2	8
1	7
2	6

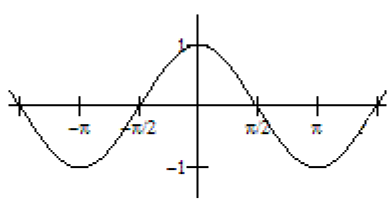
c.

x	***y***
-4	-8
-3	-6
1	-2
2	4

d.

x	***y***
2	.2
20	2.0
8	.8
14	1.4

16. What is the period of this function?



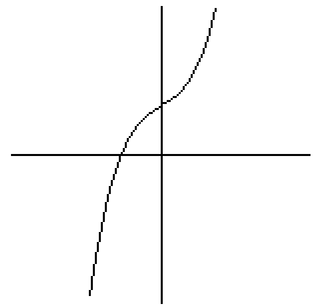
a. 1

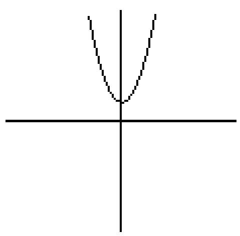
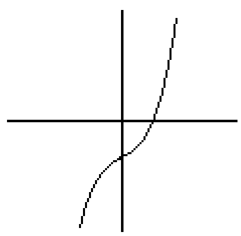
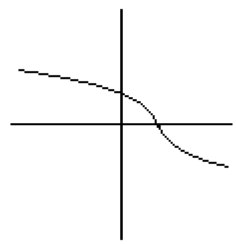
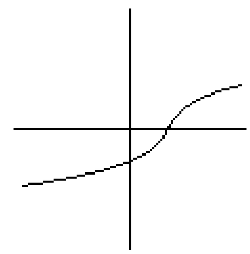
b. π/2, 3π/2,…

c. 2π

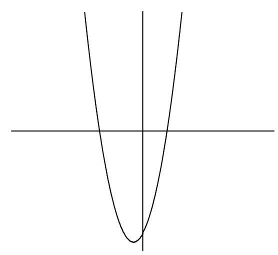
d. 2

17. Given the graph, which of the following shows the inverse?



a  b  c  d 

18. Which expression will shift the graph of the function $f(x) = 4x^2 + 2x - 3$, to the right 3 units?



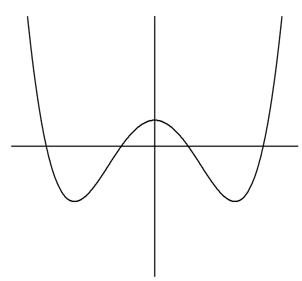
a. $12x^2 + x - 9$

b. $4(x-3)^2 + 2(x-3) - 3$

c. $4x^2 + 2x$

d. $4(x+3)^2 - 3$

19. The following function is symmetric about which line?



a. the *y* axis

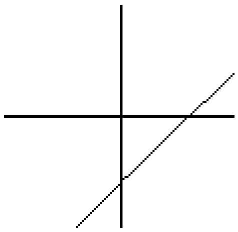
b. the *x* axis

c. not symmetric

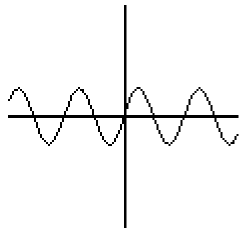
d. the line y = x

20. Which of the following has no inverse function?

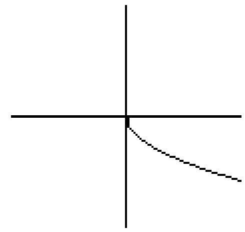
a



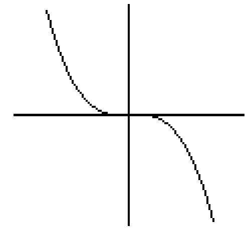
b



c



d



21. Given $f(x) = 4x - 7$, find the inverse.

a. $f^{-1}(x) = \dfrac{x+7}{4}$

b. $f^{-1}(x) = 4x - 7$

c. $f^{-1}(x) = 4y - 7$

d. $f^{-1}(x) = \dfrac{x}{4} + 7$

22. Given $f(x) = x^2 + 7$ and $g(x) = x + 3$, find $f(g(x))$.

a. $f(g(x)) = x^2 - 16$

b. $f(g(x)) = x^2 + 10$

c. $f(g(x)) = x^2 + 6x + 9$

d. $f(g(x)) = x^2 + 6x + 16$

23. What is the range of this function?

x	***y***
2	4
3	9
4	16
5	25

a. {2, 3, 4, 5}

b. {4, 9, 16, 25}

c. {1, 2, 3, 4}

d. {2, 4, 3, 9}

24. Consider the function $f(x) = ax^2 + bx + c$, where a, b, and c are real numbers and a $\neq$ 0. Evaluating the function at x = 1 returns a positive value. Evaluating the function at x = 6 returns a negative value. How many real solutions does this equation have?

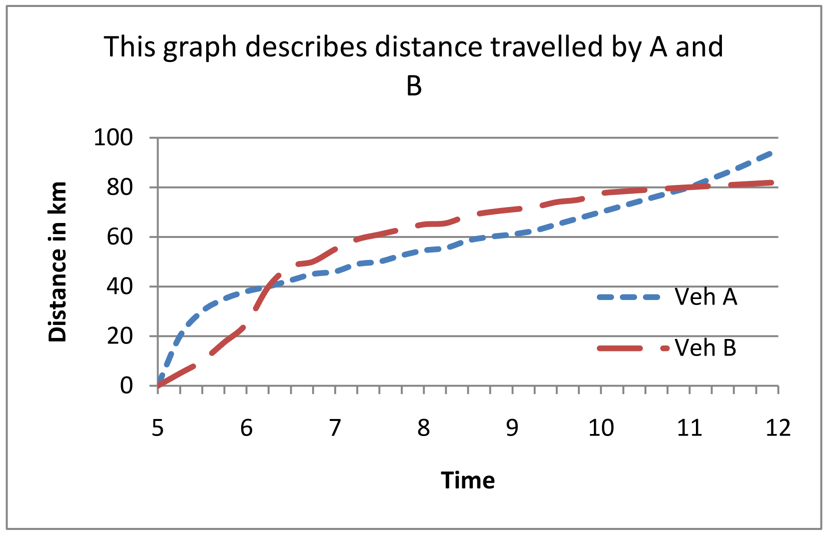
a. 1

b. 2

c. 3

d. 4

Use the diagram (Even, 1998) for the questions 25-27.



25. Approximately how far had vehicle B traveled at 7:15?

a. 30 km

b. 40 km

c. 50 km

d. 60 km

26. Who had traveled farther at 8?

a. vehicle A

b. vehicle B

c. they had traveled the same distance.

d. not enough information on graph.

27. Over what time period is A traveling faster than B?

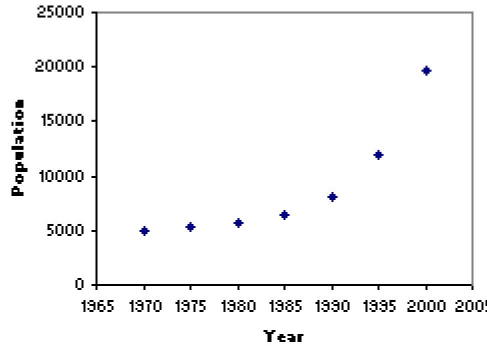
a. 5:00 – 6:00

b. 7:30 – 8:00

c. 9:00 – 12:00

d. 11:00 – 11:30

28. Given the following graph, what sort of function would best model the data?



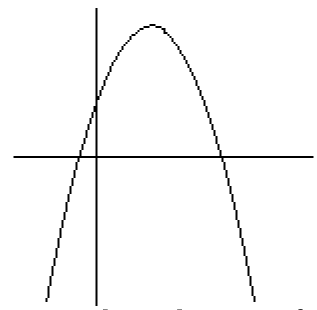
a. linear

b. exponential

c. quadratic

d. inverse

29. This graph represents the function $f(x) = ax^2 + bx + c$, where a, b, and c are real numbers. Which of the following can we say about the sign of *a*?



a. *a* is positive.

b. *a* is negative.

c. *a* equals 0

d. from the graph we cannot determine the sign of *a*.

30. Given the same function as described in question 29, what can we say about the sign of *c*?

a. *c* is positive.

b. *c* is negative.

c. *c* equals 0

d. from the graph we cannot determine the sign of *c*.

31. The population (P) of a town in millions is defined as a function of the number of years since 1950 (t), so f(t) = P. Which of the following is a correct interpretation of f(35) = 12?

a. In 12 years, the population will be 35 million.

b. In 1962, the population was 35 million.

c. In 1985, the population was 12 million.

d. In 1950, the population was less than one million.

32. Please graph the following, and state whether or not it is a function.

$$f(x) = \begin{cases} x^2; \, for \; all \; x \leq 2 \\ 1; \; for \; all \; x > 2 \end{cases}$$

Appendix B

Representational Transfer Test

Instructions (printed in header at top of each page):

For each problem you will be a function in the form of a graph, formula, table, or verbal description. Please select the answer choice that is the ***same function*** in a different representation.

(See next page for test.)

REPRESENTATIONAL TRANSFER TEST

For each problem you will be a function in the form of a graph, formula, table, or verbal description. Please select the answer choice that is the same function in a different representation.

Example: FUNCTION IS PROVIDED AS AN EQUATION: $y = 4x + 7$

CHOOSE THE CORRECT VERBAL DESCRIPTION:

A. a line in which the slope is positive four and the y-intercept is negative seven

B. a parabola opening upward with the vertex at (4, 7)

C. a parabola opening downward with the vertex at (7, 4)

D. a line with a positive slope of four and a y-intercept of positive seven CORRECT

1. $y = \frac{1}{2}x + 3$

a.

x	y
2	-7
4	-5
6	-3
8	-1

b.

x	y
2	4
4	10
6	8
8	13

c.

x	y
2	-7
4	5
6	53
8	245

d.

x	y
2	4
4	5
6	6
8	7

2. A function in which an eight unit increase in the y-coordinate corresponds to a two unit increase in the x-coordinate.

a.

x	y
70	160
71	162
72	164
73	166

b.

x	y
70	160
71	168
72	176
73	184

c.

x	y
70	164
71	168
72	180
73	188

d.

x	y
70	164
71	168
72	172
73	176

3. $y = -3x - 2$
   a. A logarithmic function passing over the y-axis at -2.
   b. A line climbing upward three units for each unit traveled to the right.
   c. A linear function with a negative slope and a y intercept of negative two.
   d. A downward opening parabola with the vertex at (0, -2).

4.

x	y
-2	12
2	0
2	12
4	48

a. $y = 3x^2$
b. $y = 6x$
c. $y = -2e^x + 4$
d. $y = 4x^3$

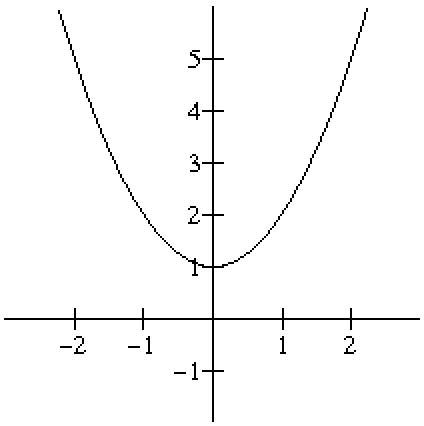
5. $y = -3 + 4x^2$
   a. A narrow upward opening parabola with the y-intercept at negative three.
   b. A line climbing upward four units for each three units traveled to the left.
   c. A function where the y coordinate is the square of the x coordinate.
   d. A wide downward opening parabola.

6.

x	y
4	40
8	30
12	20
16	10

a. A quadratic function in which the y coordinate is 24 plus the square of the x coordinate.
b. A negative quadratic function with the vertex at (2, 50)
c. A linear function in which the y coordinate is 10 times the x coordinate.
d. A linear function in which a change of negative five in the y coordinate corresponds to a change of positive two in the x coordinate.

7.



a. An upward climbing line, passing through the point (0, 1)

b. A cubic function going through the origin

c. An upward facing parabola with the vertex at (0, 1)

d. A line moving downward from left to right, passing though (0, 0)

8. $y = -2x^2 + 3x + 5$

a.

x	y
-3	-4
0	5
1	0
2	3

b.

x	y
-3	32
0	5
1	5
2	19

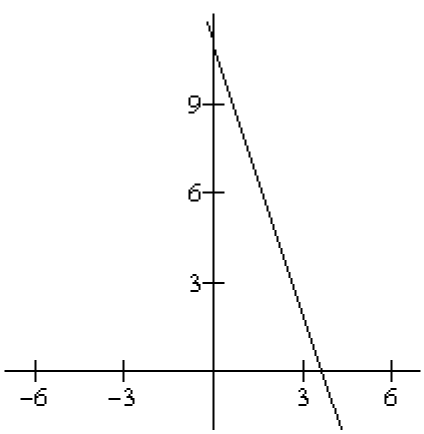
c.

x	y
-3	14
0	5
1	6
2	3

d.

x	y
-3	-4
0	5
1	-3
2	-11

9.



a. $y = 11 - 3x$

b. $y = e^x + 11$

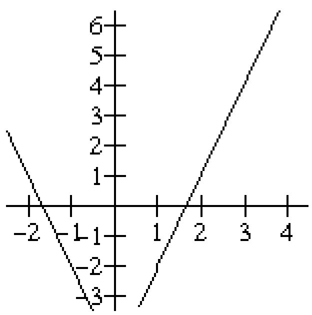
c. $y = 3^x - 11$

d. $y = 11 + x^2$

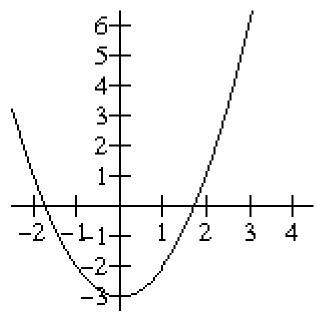
10.

x	y
-2	1
-1	-2
1	-2
3	6

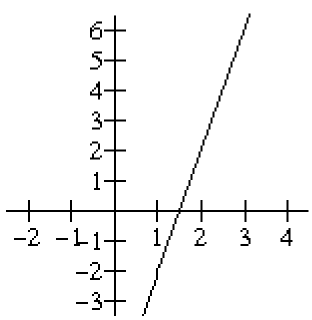
a.



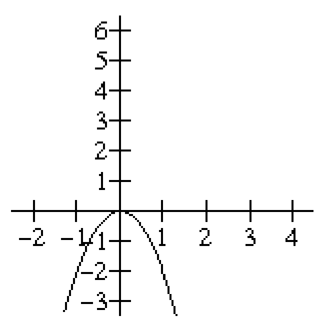
b.



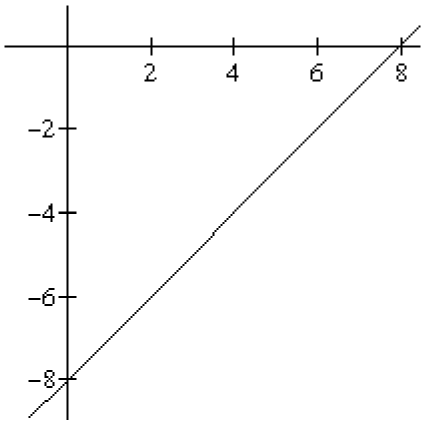
c.



d.



11.



a.

x	y
4	-4
6	-3
8	-2
10	-1

b.

x	y
4	-2
6	2
8	6
10	10

c.

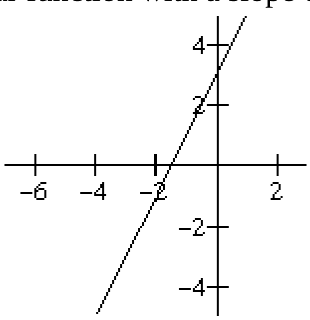
x	y
4	-4
6	-2
8	0
10	2

d.

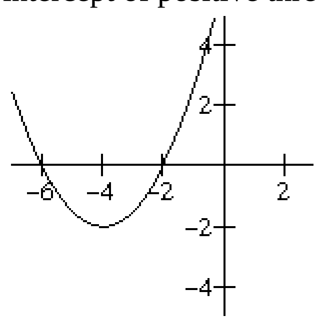
x	y
4	4
6	2
8	0
10	2

12. A linear function with a slope of positive two and y-intercept of positive three.

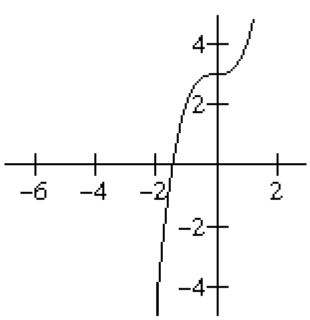
a.



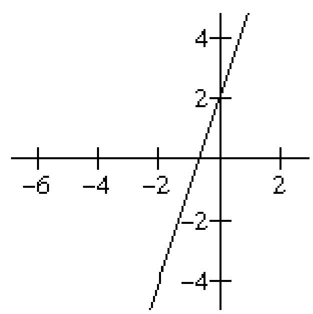
b.



c.

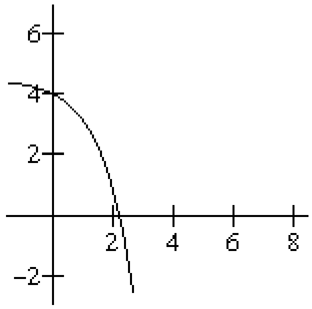


d.

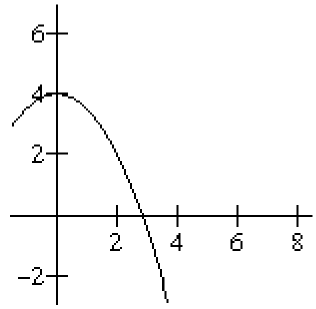


13. $y = -\frac{1}{2}x + 4$

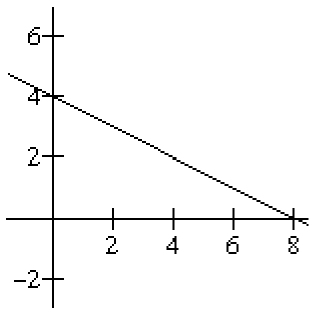
a.



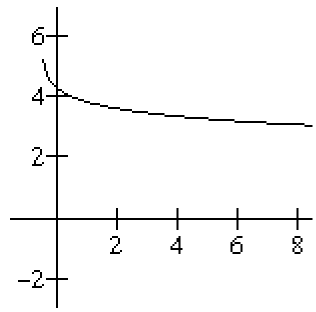
b.



c.



d.



14.

x	y
-4	-16
-1	-1
2	-4
5	-25

a. The y-coordinate is the negative square of the x-coordinate.

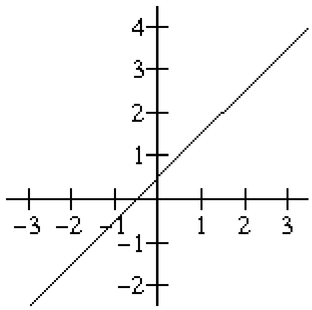
b. The x-coordinate is one-quarter of the y-coordinate.

c. The y-coordinate is the square of the x coordinate.

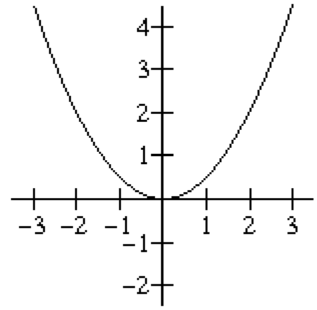
d. The y-coordinate is two times the x-coordinate.

15. The function representing the square of x transposed up one-half unit.

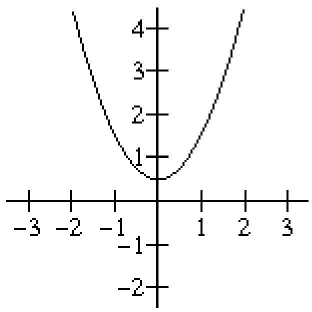
a.



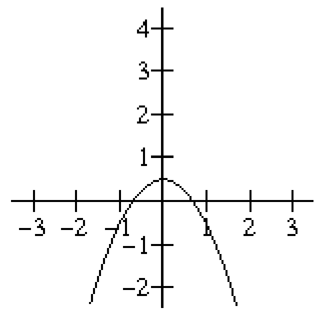
b.



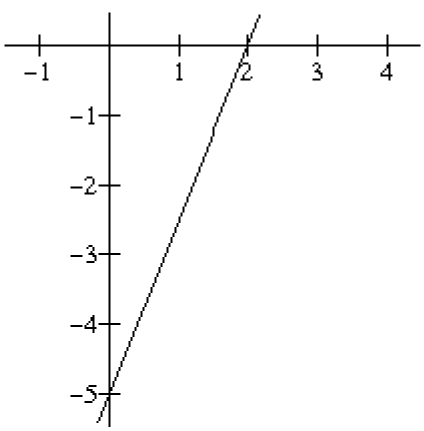
c.



d.



16.



a. A parabola opening downward with the vertex at (0,-5)
b. An upward climbing line, passing through the point (0, -5)
c. A negative exponential function
d. A line moving downward from left to right, passing though (2, 0)

17. An upward opening parabola in which the vertex is one unit down from the origin.

a.

x	y
0	-1
1	-2
2	-3
3	-4

b.

x	y
0	1
1	-1
2	1
3	-1

c.

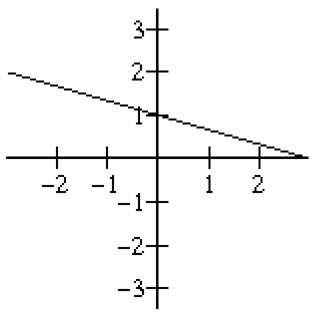
x	y
0	-1
1	0
2	1
3	2

d.

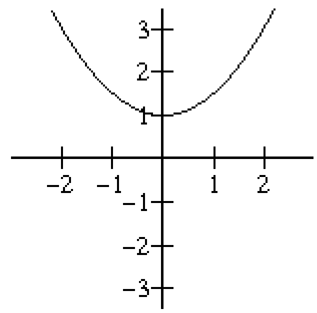
x	y
0	-1
1	0
2	3
3	8

18. $y = \frac{1}{2}x^2 + 1$

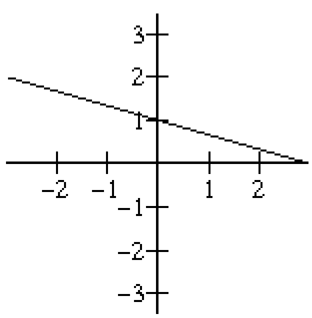
a.



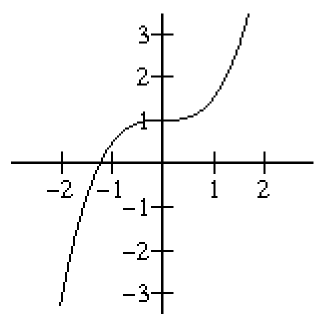
b.



c.



d.



19. A quadratic function with y-intercept of positive two and a negative leading coefficient.

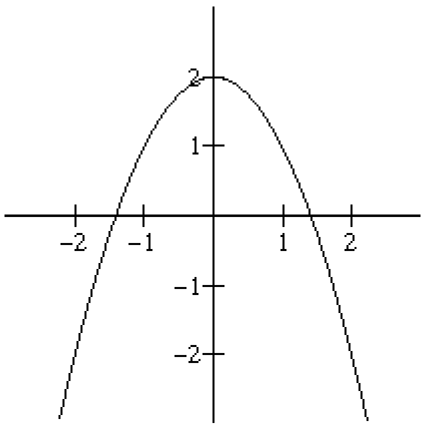
a. $y = 2 - \sqrt{x}$
b. $y = 2 - x^2$
c. $y = 2x^2 + 2$
d. $y = -2^x$

20.

x	y
-3	-5
-1	-1
1	3
3	7

a. $y = 1 + 2x$
b. $y = -3x + 4$
c. $y = 2 + x^2$
d. $y = -x^2$

21.



a.

x	y
-2	-2
0	2
2	6
4	10

b.

x	y
-2	-2
0	2
2	-2
4	-14

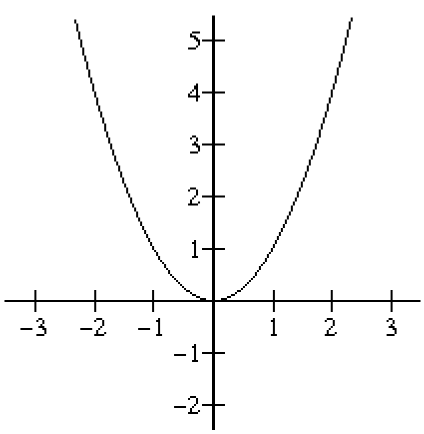
c.

x	y
-2	-4
0	-1
2	-4
4	-16

d.

x	y
-2	0
0	2
2	-2
4	-16

22.

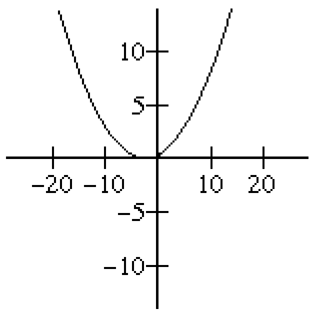


a. $y = x^2$
b. $y = x^3$
c. $y = e^x$
d. $y = x$

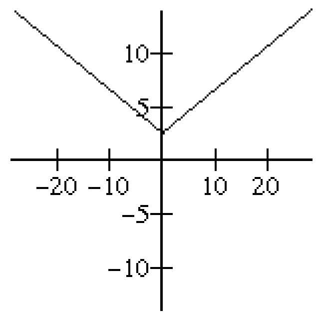
23.

x	y
-20	-5
-5	1
10	7
25	13

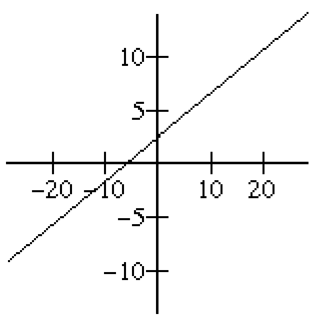
a.



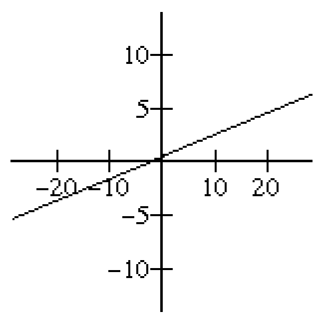
b.



c.



d.



24. A line with a y-intercept at 40 and a slope of positive 12.

a. $y = 12x - 40$
b. $y = 2x + 42$
c. $y = 40x + 12$
d. $y = 12x + 40$

Appendix C

Epistemic Beliefs (modified from Hennessey, 2007)

Please circle the number that indicates your agreement to a statement between a 1 and 5,

1 means "strongly disagree" | 4 means "agree"
2 means "disagree" | 5 means "strongly agree."
3 means "neither agree nor disagree"

	SD	D	N	A	SA
1. Links between as many concepts as possible should be emphasized.	1	2	3	4	5
2. Concepts should be connected with other previously learned concepts.	1	2	3	4	5
3. Most examples should be derived from a few basic understandings.	1	2	3	4	5
4. Demonstrations show how reasoning can be confirmed with data collected as evidence.	1	2	3	4	5
5. Newly learned information should build upon what is known to be true.	1	2	3	4	5
6. New understandings should be verified through the collection of data.	1	2	3	4	5
7. Examples should be supported by evidence collected from the natural environment.	1	2	3	4	5
8. Observations should be based on facts that are always true.	1	2	3	4	5
9. Facts should be based on known truths rather than opinion.	1	2	3	4	5
10. Explanations based on observable evidence should be more viable than explanations not based on observable evidence.	1	2	3	4	5
11. New information should align with what is already understood.	1	2	3	4	5
12. Demonstrations should reinforce basic understanding about a subject.	1	2	3	4	5
13. Thinking should be aligned with observable evidence.	1	2	3	4	5
14. Subject matter in school should be based on a few core concepts.	1	2	3	4	5
15. Explanations should show that new information is related to numerous concepts.	1	2	3	4	5
16. New facts should be explained using facts known to everyone.	1	2	3	4	5
17. Understandings should be evident to everyone.	1	2	3	4	5

Please circle the number that indicates your agreement to a statement between a 1 and 5, 1 means “strongly disagree” 4 means “agree” 2 means “disagree” 5 means “strongly agree.” 3 means “neither agree nor disagree”					
	SD	D	N	A	SA
18. Observations should be collected and described in order to inform understanding.	1	2	3	4	5
19. New facts should be explained by connecting them with existing understanding	1	2	3	4	5
20. New understanding should be explained as being consistent one’s existing knowledge base.	1	2	3	4	5
21. Understanding should be justified with observable evidence.	1	2	3	4	5
22. New facts should be explained by building upon basic understandings.	1	2	3	4	5
23. Conclusions should be checked by using observable evidence.	1	2	3	4	5
24. Evidence should be provided for one’s thinking.	1	2	3	4	5
25. Links among concepts should be examined.	1	2	3	4	5
26. The content taught in school should show that many concepts are related.	1	2	3	4	5
27. The content taught in school requires reasoning to be based on evidence.	1	2	3	4	5
28. Examples should emphasize the ways in which concepts are related.	1	2	3	4	5
29. Premises underlying a topic should be central to acquiring knowledge.	1	2	3	4	5
30. No one understanding is more important than any other.	1	2	3	4	5

Appendix D

Demographic Sheet

DIRECTIONS: Please fill in the blanks or appropriate bubbles below.

Age: _____________

Gender: ____ M ____ F

Ethnicity: ____

Major area of study: ____

Minor area of study: ____

Please indicate your college classification:

Freshman	Certification
Sophomore	Masters
Junior	Ph.D.
Senior	Other____________________________

Please indicate your current grade point average (GPA):___________

Please indicate your SAT score: _______ math ________ verbal

Please indicate the number of years you took math in high school: _________

Did you take calculus? ____yes ____no

What grade did you receive? ____ A ____ B ____ C ____ D ____ F

Please indicate the number of mathematics courses you have taken in college: _________

Did you take calculus? ____yes ____no

What grade did you receive? ____ A ____ B ____ C ____ D ____ F

Please list the names (or course numbers) of those math courses:

Appendix E

Demographics of Exploratory and Confirmatory Samples

	All Students	Exploratory Sample	Confirmatory Sample
Course			
Education	194 (.303)	89 (.296)	105 (.376)
Biology	322 (.503)	148 (.492)	174 (.513)
Missing	124 (.194)	67 (.213)	60 (.177)
Gender			
Female	141 (.220)	224 (.744)	275 (.811)
Male	499 (.780)	77 (.256)	64 (.189)
Year			
Freshman	80 (.125)	41 (.136)	39 (.115)
Sophomore	372 (.581)	171 (.568)	201 (.593)
Junior	117 (.183)	57 (.189)	60 (.177)
Senior	70 (.109)	32 (.106)	38 (.112)
Missing	1 (.002)	0 (.000)	1 (.003)
Ethnicity			
African-American	14 (.022)	7 (.023)	7 (.021)
Asian-American	13 (.020)	6 (.020)	7 (.021)
Caucasian	566 (.884)	270 (.897)	296 (.889)
Hispanic	13 (.020)	7 (.023)	6 (.018)
Foreign	21 (.033)	7 (.023)	14 (.041)
Multi-racial	3 (.005)	0 (.000)	3 (.009)
Missing	10 (.016)	4 (.013)	6 (.018)
High School Math Courses			
Two	6 (.009)	2 (.007)	4 (.012)
Three	76 (.119)	35 (.119)	41 (.121)
Four	547 (.855)	258 (.857)	289 (.853)
Missing	11 (.017)	6 (.020)	5 (.015)
University Math Courses			
Zero	77 (.120)	38 (.126)	39 (.115)
One	192 (.300)	79 (.262)	113 (.333)
Two	232 (.363)	113 (.375)	119 (.351)
Three	96 (.150)	50 (.166)	46 (.136)
Four	27 (.042)	15 (.050)	12 (.035)
Five	9 (.014)	4 (.013)	5 (.015)
Six	3 (.005)	1 (.003)	2 (.006)
Missing	4 (.006)	1 (.003)	3 (.009)
High School Calculus Taken	352 (.550)	169 (.561)	183 (.540)
University Calculus Taken	170 (.266)	82 (.272)	88 (.260)
#32 BKFT total correct	47 (.073)	23 (.076)	24 (.071)
#32 BKFT graph correct	85 (.133)	43 (.143)	42 (.124)

Appendix F

Mean Scores (*SD*) on Measures Taken and External Criteria

	All Students	Exploratory Sample	Confirmatory Sample
N	640	301	339
Basic Knowledge of Functions Test	16.07 (4.82)	16.29 (4.89)	15.87 (4.76)
Representational Transfer Test	16.66 (4.30)	16.86 (4.12)	16.49 (4.45)
Conceptions of Mathematics	70.27 (8.00)	70.34 (8.24)	70.21 (7.80)
GPA	3.36 (0.39)	3.36 (0.37)	3.35 (0.41)
SAT-m	593.9 (82.72)	598.4 (79.13)	590.1 (85.7)
SAT-v	575.6 (74.04)	578.0 (73.10)	573.6 (74.9)
# High School Math Courses	3.86 (0.37)	3.87 (0.36)	3.85 (0.39)
# College Math Courses	1.75 (1.12)	1.80 (1.13)	1.71 (1.11)
High School Calculus Grade	3.52 (0.63)	3.47 (0.68)	3.56 (0.58)
College Calculus Grade	3.03 (0.84)	3.15 (0.82)	2.92 (0.85)

Appendix G

Demographics of Latent Classes

	All Students	Class 1	Class 2	Class 3	Class 4
Course					
Education	194 (.303)	25 (.446)	71 (.397)	73 (.268)	23 (.204)
Biology	322 (.503)	21 (.375)	83 (.464)	136 (.500)	67 (.593)
Missing	124 (.194)	10 (.179)	25 (.140)	63 (.232)	23 (.204)
Gender					
Female	141 (.22)	46 (.821)	151 (.844)	221 (.812)	65 (.575)
Male	499 (.78)	10 (.179)	28 (.156)	51 (.188)	48 (.425)
Ethnicity					
African-American	14 (.022)	3 (.054)	7 (.039)	2 (.07)	0 (0)
Asian-American	13 (.020)	1 (.018)	2 (.011)	6 (.022)	4 (.035)
Caucasian	566 (.884)	47 (.839)	157 (.877)	246 (.904)	98 (.867)
Hispanic	13 (.020)	2 (.036)	5 (.028)	4 (.015)	2 (.018)
Foreign	21 (.033)	2 (.036)	5 (.028)	7 (.026)	7 (.062)
Multi-racial	3 (.005)	0 (0)	1 (.006)	2 (.007)	0 (0)
High School Math Courses					
Two	6 (.009)	2 (.036)	0 (0)	2 (.007)	1 (.009)
Three	76 (.119)	10 (.179)	32 (.179)	22 (.081)	8 (.071)
Four	547 (.855)	44 (.786)	143 (.799)	245 (.901)	101 (.894)
University Math Courses					
Zero	77 (.120)	6 (.107)	16 (.089)	38 (.140)	14 (.124)
One	192 (.300)	24 (.429)	57 (.318)	82 (.301)	19 (.168)
Two	232 (.363)	18 (.321)	77 (.430)	106 (.390)	28 (.248)
Three	96 (.150)	8 (.143)	20 (.112)	36 (.132)	29 (.257)
Four or more	41 (.061)	0 (0)	9 (.050)	9 (.031)	21 (.186)
High School Calculus Taken	352 (.550)	14 (.250)	75 (.419)	155 (.570)	99 (.876)
University Calculus Taken	170 (.266)	3 (.054)	23 (.128)	70 (.257)	67 (.593)

Kelli Higley Vita

EDUCATION

Ph.D. Education Psychology, The Pennsylvania State University, August, 2009
M.A. Educational Psychology, University of Connecticut, December 2003
B.S. Mathematics Education, Utah State University, March 1997

PROFESSIONAL ASSOCIATIONS

American Educational Research Association
American Psychological Association

PUBLICATIONS

Edwards, M. N., Higley, K., Zeruth, J. A., & Murphy, P. K. (2007). Pedagogical practices: Examining preservice teachers' perceptions of their abilities. *Instructional Science, 35,* 443-465.

SELECTED RESEARCH PRESENTATIONS

Lissenden, C., Salamon, N., Van Meter, P., Higley, K., & Vavareck, A. (2008, June). *Design in Mechanics of Material Courses for Deeper Learning*, Poster presented at the annual conference of the American Society for Engineering Education, Pittsburgh, PA.

Higley, K., Litzinger, T., Van Meter, P., Kulikowich, J., & Master, C. (2007, June). *Effects of Conceptual Understanding, Math and Visualization Skills on Problem-solving in Statics,* Paper presented at the annual conference of the American Society of Engineering Education, Honolulu, HI.

Higley, K., & Edwards, M. N. (2006, August). *A New Paradigm for Mathematics Education: Persuasive Pedagogy in Action*. Paper presented at the annual meeting of the American Psychological Association, New Orleans, LA.

Meyer, B. J. F., Wijekumar, K., Middlemiss, W., Meier, C., Spielvogel, J., & Higley, K. (2006, April). Impact of computer-delivered strategy instruction on student outcomes. In E. Albro (Chair), *The many faces of strategy instruction*. Symposium conducted at the annual meeting of the American Educational Research Association, San Francisco, CA.

Higley, K., Edwards, M. N., & Gushka, J. A. (2005, August). The viability of the persuasion metaphor for educators with varying levels of teacher-efficacy. In H. Fives (Chair), *Teaching as persuasion: Is the metaphor viable?* Symposium presented at the annual meeting of the American Psychological Association, Washington, D.C.

Higley, K. (2005, August). *What content specialization in mathematics tells us about teachers' attitudes, efficacy and content knowledge.* Paper presented at the annual meeting of the American Psychological Association, Washington, D.C.