INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI®

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

NOTE TO USERS

This reproduction is the best copy available.

UMI

UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

CONSTRUCTING GRAPHS TO MEET SPECIFIED CONDITIONS:

A GENERALIZABILITY STUDY

A Dissertation

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

Doctor of Philosophy

By

LINDA SUE BRADDY
Norman, Oklahoma
2000

UMI Number: 9961563

UMI®

UMI Microform 9961563

Copyright 2000 by Bell & Howell Information and Learning Company.
All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

Bell & Howell Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

CONSTRUCTING GRAPHS TO MEET SPECIFIED CONDITIONS:
A GENERALIZABILITY STUDY

A Dissertation APPROVED FOR THE
DEPARTMENT OF MATHEMATICS

BY

Curtis C. McKnight

[illegible]

Luther White

[illegible]

Marilyn Breen

© Copyright by LINDA SUE BRADDY 2000
All rights reserved

ACKNOWLEDGEMENTS

It would be impossible to acknowledge everyone who has supported me in this work and my degree, but several people deserve special recognition. My husband, Tommy, deserves first mention. He has been an incredible emotional support to me since I began graduate school in 1992, not to mention the way he has taken up the slack for me in family and household responsibilities. He has served as a sounding board for my ideas as well as my frustrations. He has been a terrific "Mr. Mom" for our children these past few months as I made a final push to finish this writing. Our children, Vanessa and Matthew, have been gracious and understanding, especially the past three months when there were many days they only saw their mom in the mornings before school. Many members of my extended family have offered words of encouragement and support as well, for which I am grateful. Special thanks goes to my older brother, Robbie, for setting a high standard of achievement in mathematics in junior high school for me to live up to. The memory of my Dad's desire to see me further my education has been a source of inspiration for me since his death in 1986.

My advisor, Dr. Curtis McKnight, deserves recognition for his input and guidance on this project, as well as his assistance on other projects undertaken during my years at the University of Oklahoma. Sincere thanks goes to a fellow graduate student of mine, Trisha Bergthold, not only for her work on this research project, but also for the constant support and assistance she offered me since she first came to the University of Oklahoma in 1993.

I would like to express my appreciation to my colleagues in the Mathematics Department at East Central University. Their assistance and support over the past two years has been invaluable to me. Special thanks goes to my friend, Dr. Anne Fine, for her invaluable advice and encouragement, as well as specific input on this project. She spent an incredible amount of time proofreading for me. Her husband, Dennis, also was gracious enough to assist with the proofreading, for which I am grateful. Dr. Ray Hamlett has supported me in numerous ways, from the teaching assignments he gave me to the equipment and supplies he made available. Two of our student workers, Heather and Dustin, have been especially helpful to me this semester as well. I would also like to thank the student participants in my calculus class who made this research possible.

Many people have supported me with their prayers, particularly my dear friends LaDonna Ballard, Margala Rosson, and LaRuth Meeks, and I would like to express my gratitude to them as well. My church family at Trinity Baptist Church have offered many words of encouragement over the past several years, for which I am grateful. Above all, I am thankful for a truly awesome God who has blessed me with the opportunities, abilities, and support necessary to accomplish this goal that He set before me years ago. To Him, I am forever grateful.

TABLE OF CONTENTS

List of Tables viii

Abstract x

Introduction 1

Problem Statement 3

Significance of the Study 4

Research Questions 4

Review of the Literature 6

The Experiment 25

Subjects 25

Course Description 27

Method 28

Design 32

Data Analysis 36

Results 38

Generalizability Analyses 38

Interviews 57

Conclusions 74

Discussion 74

Limitations of the Study 78

Directions for Future Research 81

References 83

Appendix A: Human Subjects Permission 89

Appendix B: Informed Consent Form 92

Appendix C: Background Questionnaire 95

Appendix D: Target Problems 101

Appendix E: Examples of Student Solutions 125

Appendix F: Steve's Solution to Problem 4 on Quiz 3 149

Appendix G: Tom's Solution to Problem 4 on Quiz 3 151

Appendix H: Ben's Solution to Problem 4 on Quiz 3 153

LIST OF TABLES

1. MODEL: p × r × i Type 1 items 39

2. MODEL: p × r × i Type 2 items. 40

3. MODEL: p × r × i All items 41

4. MODEL: p × i × (o:s) Type 1 items 42

5. MODEL: p × i × (o:s) Type 2 items 43

6. MODEL: p × t × s 44

7. MODEL: p × o × (i:t) 45

8. MODEL: p × o × i Type 1 items 46

9. MODEL: p × o × i Type 2 items 46

10. MODEL: p × o × i Type 1 items, Quizzes only 47

11. MODEL: p × o × i Type 1 items, Examinations only 48

12. MODEL: p × o × i Type 2 items, Quizzes only 48

13. MODEL: p × o × i Type 2 items, Examinations only 49

14. MODEL: p × t × (o:s) Quiz 1, Quiz 4 50

15. MODEL: p × t × o Quiz 1 and Quiz 4 only 51

16. MODEL: p × t Examination 2, A students 52

17. MODEL: p × t Examination 3, BCDF students 52

18. MODEL: p × t Quiz 4, All students 53

19. MODEL: p × t Quiz 4, BCDF students 53

20. MODEL: p × t Quiz 1, All students 53

21. MODEL: p × t Quiz 1, Students with average below 95 percent. 54

22. MODEL: p × t Quiz 1, BCDF students 54

23. MODEL: p × t Quiz 1, A students 55

24. MODEL: p × t × o Quiz 1 and Quiz 4 only, A students 55

25. MODEL: p × t × o Quiz 1 and Quiz 4 only, BCDF students 56

26. MODEL: p × t × o Quiz 1 and Quiz 4 only, Students with average 95 percent above 56

ABSTRACT

This study investigated differences in problems involving graphical representations of functions. Specifically, problems requiring students to answer questions about a given graph were compared with problems requiring students to construct a graph to meet specified conditions.

Subjects for the study were students enrolled in a first-semester calculus course during the 1998 fall semester at a medium-sized regional university. Thirty-two students, 13 females and 19 males, completed 24 problems involving graphical representations of functions. Half of these problems required students to answer questions about a given graph, and the other half required students to construct a graph meeting specified conditions. The students' written solutions to these problems were scored by two raters. Generalizability theory was used to study the differences between the two types of problems.

Numerical results suggested true systematic differences in problems types do exist. Information obtained from personal interviews revealed differing opinions regarding difficulty level of the problem types, which supported the numerical findings. Seventy-five percent of the students interviewed thought it was more difficult to construct a graph to meet specified conditions than it was to answer questions about a given graph. Although it might be expected that only higher ability students would think constructing a graph is easier than answering questions about a graph, in fact three of the students who held this opinion had the lowest final course grades. In addition, the numerical results indicated the students prepared differently for examinations than for quizzes. Information from the interviews revealed that this was due in part to the fact

that the examination and quiz scores were weighted differently in the calculation of final grades. The interviews also revealed two basic strategies students used to construct graphs, in addition to several unexpected misconceptions held by the students.

A major result of this study is a warning against assuming comparability in assessment items involving graphical representations of functions. An additional result is a warning against the use of simple tests of differences in mean achievement to determine if items are similar and comparable. When the measurement situation is unclear, studies such as this generalizability study of what factors contribute to variance become especially important.

CHAPTER 1

Introduction

The use of multiple representations of functions (graphical, numerical, and symbolic representations) has been a major theme in recent calculus reform (e.g., Douglas, 1986; Leitzel, 1991; NCTM, 1989; Roberts, 1996). The assumption is that students benefit from exposure to different representations of the same function. Presumably, these benefits include either (1) students mastering content in more powerful, flexible forms, (2) students mastering content more quickly than would be the case without deliberate attention to multiple representations, (3) students mastering content who would not normally be able to do so, or (4) some combination of the first three (Braddy and McKnight, 1997). This idea is not unique either to the current reform movement or to the subject of calculus. Dienes may be regarded as the father of multiple representations (Ahmadi, 1995). According to Dienes (1960, p. 3), there is no single representation that can lead to abstraction and variety is essential. He asserted that the use of multiple representations is a prerequisite to understanding a concept. Numerous other researchers and practitioners have made similar claims, based on either empirical evidence or their own classroom experiences (e.g., Garofalo and Durant, 1991; Piez and Voxman, 1997).

The validity of the notion that students benefit from exposure to different representations of the same function, particularly graphical representations, is rarely questioned, and seldom tested. In order to test the validity of this idea, we need to know what types of questions are appropriate to assess understanding of functions and their graphical representations. This assessment often is assumed to be unproblematic and already understood. Is this assumption warranted? If in fact the appropriateness of such assessment methods is more complex than assumed, then empirical studies incorporating the use of graphical representations might have more ambiguous results than suggested at first glance. Effective, well understood assessment methods that produce less ambiguous research results are also essential for developing better curricula, textbooks, and instruction.

This study is one step in an agenda to clarify what is measured in assessment items centered on graphical representations and to develop additional new approaches. In particular, there seems to be a clear distinction between assessments that question students about *given* graphical representations and assessments that ask students to *construct* graphical representations meeting specified conditions. These conditions may include both general characteristics of a function, such as the requirement that the function pass through a certain point, as well as characteristics more closely tied to ideas from calculus, such as the requirement that it have a local extrema or inflection point at a certain location. This latter item type, requiring students to construct a graph of a function, is a cognitive task that does not depend on multiple representation ideas or on the increased use of graphing calculators, but its relevance has been increased by these more graph-intensive approaches that have been characteristic of recent calculus reform.

Further, the use of this sort of assessment item may increase student attention to instructional efforts on more powerful, fundamental uses of graphical representations. Appropriate assessment helps set students' learning agendas.

This study is an empirical analysis designed to clarify issues in answering the question, "For typical calculus students, do assessment items with questions asked about a *given* graph and items that require *constructing* a graph to meet specified conditions draw on different student abilities and achievements?". If it can be demonstrated that these two types of graphical representation items are different, the way is opened for further pertinent questions: Are the differences due to student ability or achievement? Do the differences correlate with other differences in student performance? Are there differing parts of calculus instruction that are needed to improve student performance for each area? This study is only the first step in an agenda of related questions.

Problem Statement

The problem addressed in this study is ascertaining whether or not systematic differences exist in two different types of assessment items involving graphical representations of functions. Specifically, this study will investigate students' written solutions to problems given on quizzes and examinations as part of the normal assessment activities in a first-semester, graphing-calculator-based calculus course. Information obtained through personal follow-up interviews will further confirm the quantitative results.

Significance of the Study

Numerous studies have investigated the function concept. Much of the research literature for functions focuses on the definition of function and how the understanding of the function concept develops. Many studies have focused on multiple representations of functions, and in particular, graphical representations. However, no research exists that addresses the question of whether or not different types of problems involving graphical representations of functions actually measure different student abilities. According to Romberg, Fennema, and Carpenter (1993), research in the area of graphical representations of functions is "sparse" but "a coherent body of knowledge about how the connections are developed among tables, graphs, and the algebraic expressions related to functions is desperately needed" (p. ix). This study will make a significant contribution to the research, particularly at the undergraduate level of mathematics education. By providing evidence of differences (or lack of differences) between assessment items incorporating the use of graphical representations, this study has the potential to help clarify what types of questions are appropriate for assessing understanding of functions and their graphical representations. In addition, a better understanding of these types of assessment items gained from this study has the potential to inform instruction, as well as to aid the development of curricula and textbooks.

Research Questions

1. Differences in problem types. Are there systematic differences in graphical representation problems involving questions about a *given* graph of a function and problems requiring the *construction* of a graph to meet specified conditions?

2. Differences in student performance. Do students perform differently on these two types of problems involving graphical representations of functions? If so, to what can these differences in performance be attributed? How much can these differences be attributed to systematic differences in the students? How much can these differences be attributed to differing levels of preparation of the students? How much can these differences be attributed to differences in rater evaluations?

3. Differences in student strategies, perceptions and misconceptions. What are the thought processes of students as they solve the two different types of problems involving graphical representations of functions? What types of strategies do they use to solve these two types of problems? What are students' perceptions of these two types of problems? Do they think one type is more difficult than the other? Do they think one type requires different mathematical abilities than the other? Do they enjoy one type more than the other? Are there common misconceptions that arise?

CHAPTER 2

Review of the Literature

The purpose of this chapter is to discuss how previous research influenced the research questions in this study as well as the design of the study. Literature is reported that addresses the relevance of calculus reform and multiple representations of functions to this study, that suggests the importance of graphical representations of functions in general, and that describes the statistical analysis tool referred to as generalizability theory.

Calculus Reform and Multiple Representations of Functions

The function concept is undoubtedly one of the most fundamental concepts in all of mathematics. Much research has been done regarding students' understanding of functions (Davis, 1984; Markovits, Eylon, and Bruckheimer, 1986; Monk, 1992; Orton, 1983; Tall, 1996; Vinner and Dreyfus, 1989). Advocates of calculus reform have recommended the use of multiple representations as a way to improve students' understanding of functions. The three representations emphasized are graphical, numerical, and symbolic or algebraic forms. Leinhardt, Zaslavsky, and Stein (1990) assert that the use of multiple representations is important partly because of "the increased recognition of the organizing power of the concept of functions from middle school mathematics through more advanced topics in high school and college" (p. 1). Other researchers have made recommendations regarding the advantage of using a multiple representations approach as well (Dreyfus and Eisenberg, 1987; Moschkovich, Schoenfeld, and Arcavi, 1993; Schwarz and Dreyfus, 1993; Yerushalmy and Schwartz, 1993). The assertion is that providing students with the opportunity to view functions from various perspectives increases their conceptual understanding.

The availability of technology has focused increased attention on the use of multiple representations of functions in many areas of mathematics education. This accessibility has important implications for mathematics curricula. For instance, representing functions with graphs could allow mathematics to become more a study of classes of functions than a study of symbol manipulation techniques (Philipp, Martin, and Richgels, 1983).

A tremendous amount of research literature for functions and their various representations exists. There are published studies that have investigated the effect of different kinds of technology on students' graphing and visualization skills as well as their understanding of different representations of functions (Borba and Confrey, 1996; Caldwell, 1995; Goldenberg, 1988; Shoaf-Grubbs, 1994; Yerushalmy, 1990), the effect of different instructional approaches on students' abilities to use and understand connections between different representations when solving problems (Porzio, 1995), student's graphing skills (Brasell and Rowe, 1993; Clement, Mokros, and Schultz, 1986), common graphing conceptions and misconceptions (Barclay, 1987; Mevarech and Kramarsky, 1997), and ways in which calculus students use graphs to solve problems (Patterson, 1983).

Others disagree with the idea that linking multiple representations of functions is beneficial for students. Varying degrees of importance are placed on the development of a student's ability to move flexibly among the three fundamental representations. Thompson (1994) admits he "jump[ed] on the multiple-representations bandwagon early" (p. 39), but that he was mistaken and it may be wrong to focus on graphical, numerical, and algebraic representations. He asserts,

> [T]he core concept of "function" is not *represented* [author's emphasis] by any of what are commonly called the multiple representations of functions, but instead our making connections among representational activities produces a subjective sense of invariance (p. 39).

Although some researchers do not support the notion that an emphasis on multiple representations of functions benefits the development of students' understanding of functions, the majority of the research literature does support this idea.

Graphical Representations of Functions

As important as the three fundamental representations of functions and the connections among them may be, graphical representations are important as an entity in their own right. Encouraging the use of a graphical approach in mathematics classes, Philipp et al. (1983) assert,

> [A] graphical approach makes it easier to work with many existing topics, greatly increases the range of problems which are accessible to students, and reduces the prerequisite of extensively developed manipulative skills, so that students can spend more time modeling and solving problems.

Functions and graphs should be studied and researched as an important content area in itself (Williams, 1993). Graphical representations have unique characteristics not evident from other representations of functions (e.g., they provide a view of global function behavior) and in certain cases, the graphical representation of a function might be more important than its algebraic form. In particular, applications of graphing technology in which data is input through a sensor of some sort (e.g., a motion detector) and a graph produced do not require an algebraic representation of the function in order for students to make the connection between the graph and the situation it represents (e.g., the position of a bouncing ball). The importance of graphs can also be seen in the process of building a velocity graph from a position graph (Kaput, 1993). This process does not require an algebraic formula for the function represented by either graph. The same process would be used to produce a graph of the acceleration function. This kind of task (which will later in the current study be labeled a "graph-to-graph Type 2" task) is becoming increasingly important as advances in technology provide the ability to manipulate dynamic graphical representations. These types of interactive representations require that more importance be placed on the graph than on the algebraic formula for the function.

As important as graphical representations of functions may be to the study of elementary mathematics, relatively little research has been done in this area. Williams (1993) asserts, "[T]he treatment of function and graph, viewed as mathematical...objects, remains sketchy in the research as a whole" (p. 325). And according to Leinhardt et al. (1990), "The teaching of graphs, graphing, and functions is understudied in comparison to other areas of early mathematical instruction (addition, fractions, decimals, algebra work problems)" (p. 54). Certainly, more research is needed in the area of functions and their graphical representations.

Kieran (1993) reports that the traditional emphasis regarding the use of graphical representations has begun to shift. The practice of using graphs as merely an additional way to depict an algebraic representation of a function is being replaced with the practice of using graphs as a tool to examine global features of a function as well as a tool for solving problems in applied settings. A result of this shifted emphasis is the need to focus efforts on improving students' global or qualitative graph interpretation skills, rather than remaining focused on skills involving the interpretation of local behavior (Dugdale, 1993; Krabbendam, 1982; and Phillips 1986). Interpreting local behavior includes such things as determining intercepts, slopes, extrema, and points of discontinuity (this kind of task will later in the current study be labeled as a "Type 1" task).

Student difficulties with interpretation of qualitative representations of functions are well documented in the literature (Bell and Janvier, 1981; Ferrini-Mundy and Graham, 1994; Janvier, 1978; Karplus, 1979; Peters, 1982; Shaw, Padilla, and McKenzie, 1983). Qualitative representations of functions shift focus away from numerical values to the overall relationship of the variables described by the graph. Leinhardt et al. (1990) distinguish between categories of *interpretation* of graphs (here Type 1 problems) and *construction* of graphs (here Type 2 problems), referring to qualitative interpretation as "trend detection" (p. 9). For example, students might be asked to examine a graph and describe events that could have produced the graph or a situation modeled by the graph. Alternatively, students might be asked to construct a graph of a function based on available information about the derivative of the function

(Heid, 1988), the same kind of task as the one described previously (building velocity and acceleration graphs from position graphs).

Assessment Issues

With so much emphasis being placed on improving qualitative interpretation skills and understandings, it seems more necessary than ever to insure the use of appropriate assessment methods with respect to graphical representations of functions. Markovits et al. (1986) assert that problems asking students to sketch the graph of a function that meets specified conditions (Type 2) are a type of task that probes the students' understanding of qualitative aspects of the function; in particular, it probes their understanding of the relationship between the variables. Empirical research is needed to establish effective assessment methods that can be used to confirm such assertions.

Graphical representations of functions have become more widely used in mathematics instruction as the use of graphing technology has become more prevalent. In order to investigate the validity of the assumption that exposure to graphical representations is of benefit to students, it is necessary to determine what types of questions are appropriate to assess their understanding of functions and graphs. According to Philipp et al. (1993), assessment must change in response to the use of a graphing approach in any mathematics class. The Research Advisory Committee of the National Council of Teachers of Mathematics (1990) asserts, "There is a clear need for research in the development and implementation of new approaches to the assessment of mathematics learning…" (p. 290). There is only limited research on assessment and evaluation of student learning in calculus and more generally in all college mathematics courses (Tucker and Leitzel, 1995). In assessing the present state of research in the content area of graphical representations of functions, Williams (1993) contends,

> [W]e as researchers have failed to provide a comprehensive analysis of the subject matter with an eye toward informing research on learning or teaching....Research has also failed to deal effectively with understanding the graphical representation of functions. Specifically, we have failed to engage in the sort of careful **analysis of tasks** [emphasis added] and structures that would allow for modeling the understanding, learning, and teaching of graphs and functions (p. 313).

He goes on to assert that, while gaining a better understanding of cognitive processes relative to functions and graphs may require careful analysis of the tasks themselves, such analyses of graphing tasks and graph interpretation tasks have not yet been synthesized.

In particular, the question of whether or not different types of problems involving graphical representations of functions actually measure different student abilities has not been investigated. It seems that this could be a starting point in the quest for appropriate assessment techniques.

Research Issues

If it can be established that different types of graphing problems do in fact measure different abilities, this would suggest that the practice of using a mixed variety of these types of problems as research instruments in empirical studies should be reconsidered.

Markovits et al (1986) investigated junior high school students' understanding of the function concept. Two particular components of the students' understanding that was studied was their ability to identify which of several given functions satisfied specified conditions and their ability to give examples of functions satisfying specified conditions. There were three kinds of conditions: (1) those specifying global properties of the function (e.g., increasing, constant), (2) those placed on the domain and range, and (3) those requiring the function to pass through specific points. Problems given were designed to identify student difficulties as well as to suggest causes of the difficulties. One type of problem asked the students to choose which of three graphs satisfied a specified requirement for the domain and range, which would fit into the category of interpreting graphs (similar to the Type 1 here). Another type of problem used in the study asked students to sketch the graph of a function meeting specified conditions regarding function behavior (Type 2). A third type of problem asked students to produce an algebraic representation of a function that satisfied specified conditions on the domain and range. It was found that students had difficulty constructing graphs that satisfied given constraints and even more trouble finding algebraic representations for functions satisfying the constraints. The authors acknowledge their inability to draw any conclusions based on these results, in part because many of the problems used were "not standard" and may well have been measuring a level of understanding "quite high" for the subjects. As a result, they focused on types of difficulties the students experienced rather than the students' overall success. At no time did the study examine differences among the different types of graphing problems used.

Ruthven (1990) compared the performance of two groups of upper secondary school students on different types of graphing problems. In the first type, called *symbolism* items, students were asked to give an algebraic representation of a given graph. In the second type, called *interpretation* items, students were asked interpretive questions about a given graph (a Type 1 task). For example, they were given a graph of a cyclist's speed in meters per second over a 10-minute period and asked at what point the speed was greatest. The experimental and control groups differed only in terms of their access to graphing calculators or graphing software. The performance of the two groups on the symbolism items was compared separately from their performance on the interpretation items. It was found that the experimental group performed better than the control group on the symbolization items, but not on the interpretation items. This treatment effect was attributed to the experimental group's use of graphing technology.

The following explanation was given for the pattern of performance observed:

> Although, superficially, both sets of items are concerned with graphs, they depend, in fact on very different competences [sic]. The symbolisation items depend on expertise in recognizing graphic forms, and relating them to appropriate symbolic forms....[while the interpretation items depend] on synthesising verbal, contextual, and graphic information, unmediated by symbolism (p. 448).

The author asserts that the regular use of graphing technology is likely to have improved the experimental group's performance on the symbolization items, while the same technology will not, in itself, help develop expertise with the interpretation items. However, at no time did this researcher investigate differences in the types of items in order to produce empirical evidence that they depend on different competencies; he simply compared the performance of the two groups on the items.

Ferrini-Mundy and Graham (1992) report the results of a qualitative study that explored four calculus students' understandings of calculus concepts, including functions, limits, continuity, derivatives, and integrals. These students were enrolled in a traditional calculus course taught in a large lecture format without the use of technology. Each student participated in four interviews during which they completed tasks involving graphical representations of functions. Several of the tasks were of the *interpretation* variety in which students were given a graph and asked questions about the graph (Type 1). Other tasks were of the *construction* variety (Type 2). There were two categories of construction tasks: (1) those asking the students to sketch the graph of a function that met specified conditions, and (2) those asking the students to sketch the graph of the derivative of a function that was given graphically (graph-to-graph Type 2). The study focused on probing students' procedural knowledge as well as their conceptual knowledge of the aforementioned calculus concepts. Differences among the types of graphing problems were not investigated. One conclusion was, "Startling inconsistencies exist between performance, particularly on procedural items, and conceptual understanding. Traditional means of assessment in calculus are almost certain to mask the nature of student understanding" (p. 43).

A study using similar types of problems was also reported in 1994 by Lauten, Graham, and Ferrini-Mundy. Tasks were described as those requiring students to "generate examples of functions with certain properties" (*construction*, Type 2), and those requiring students to "answer questions about functions using their graphs" (*interpretation*, Type 1). Again, no attention was given to establishing the existence of systematic differences between the two types of tasks.

Eisenberg and Dreyfus (1994) conducted a study to investigate the effects of a teaching unit designed to enhance students' visualization of function transformations. The subjects were 16 boys who were 17 or 18 years of age from a boys' senior high school in Israel. These students were given identical pre-tests and post-tests and were interviewed as well. The tasks involved were categorized as "standard" and "nonstandard". The students' success rates on the two categories of tasks were compared, and conclusions were drawn based on these comparisons. However, the "standard" tasks consisted of both *interpretation* and *construction* problems. For example, several "standard" tasks asked the students a question about a given graph (Type 1). Other "standard" tasks asked the students either to draw a graph that met specified conditions (Type 2) or to draw a graph that was related to a given graph (graph-to-graph Type 2). In their conclusions, the authors did state their belief that "the transformations used in this study are hierarchically ordered with respect to difficulty in understanding and mastering" (p. 58) and that "certain classes of transformations seemed to be easier for students than others" (p. 62). No empirical evidence was offered to confirm the correctness of their assertions, pointing again to the need for such evidence.

Clement, Mokros, and Schultz (1986) examined middle school children's graphing skills and misconceptions in a study. The research instruments used were graphing problems, some of which involved *interpretation* of a graph (Type 1), and some of which involved *construction* of a graph (Type 2). The authors assert that "we need to know more about the development of graphing competencies [authors' emphasis]" (p. 3). How can level of such competency accurately be measured if the instruments used do not measure the *same* competencies? Again, the necessity of comparability of the measurement items is required.

Preece (1983) investigated qualitative interpretation strategies and interpretation errors. Subjects for the study were 122 students who were 14 or 15 years old. Tasks were all of the *interpretation* variety (Type 1), although they included multiple curve graphs as well as single curve graphs. An example of a task that involved multiple curves is one in which time-and-distance graphs of three different cars were presented on the same set of coordinate axes. This type of task proved to be especially difficult for the students. However, no comparison was made of the differences between the different types of problems.

Other studies have been published (e.g., Yerushalmy, 1991) in which research instruments included tasks requiring students to match a graphical representation to its algebraic formula (given among several algebraic formulas), as well as tasks requiring students to produce an algebraic formula themselves, based on a given graph. No empirical evidence exists that establishes the comparability of these types of items.

These and similar research strategies appear to be perfectly reasonable strategies, but if there are in fact differences between the types of graphing problems used as research instruments, then this kind of approach is one that would need to be refined.

A study that investigated calculus students' strategies for sketching a graph of a function that met several specified conditions (Type 2) was conducted by Baker, Cooley, and Trigueros (in press). They developed a matrix in accordance with the APOS theory (Asiala, Brown, DeVries, Dubinsky, Mathews, Thomas, 1996) which depicted graphing schema used by the students. They also enumerated the four areas that were most problematic for the students. However, no comparison was made between this type of graphing problem and any other type.

A modest amount of research has been done in the area of graphical representations of functions but much more remains to be done. Research designed to develop appropriate methods for assessing students' understanding of functions and their graphs is certainly lacking. In particular, careful analyses of tasks involving graphs is practically nonexistent.

Generalizability Theory

Several analytical methods of data analysis were available for this study. A simple one-way analysis of variance to compare the two item types would have been appropriate if the variability in student responses had been well understood to be primarily due to student differences. However, little was known about the relative effects of several other factors affecting the variability of student responses, including effects due to raters, occasions, and stakes (i.e., quizzes vs. examinations). Therefore, generalizability theory was chosen as the primary analytical method for this study because it provides a separate estimate of the magnitude of multiple sources of variance in a single analysis.

Generalizability theory essentially decomposes variance into components due to numerous sources. With generalizability theory, it is not necessary to restrict decomposition of variance to only the component due to differences among individuals and the residual composed of systematic and random sources labeled "error" variance. This theory allows the proportion of overall variance due to differences in individual performances on a behavioral measurement to be compared with that due to various other factors in the measurement situation. A multi-facet study can be used to estimate interactions inaccessible with more commonly used methods, which in turn enables researchers to design more efficient data collection procedures (Cronbach, Gleser, Nanda, and Rajaratnam, 1972). Hence a generalizability study should incorporate into its design as many potential sources of variance as are of interest.

Generalizability theory is fairly new, with its beginnings in the work of Cronbach, Gleser, Nanda, and Rajaratnam in 1957, which resulted in the publication of their book *The Dependability of Behavioral Measurements: Theory of Generalizability for Scores and Profiles* in 1972 (Brennan, 1997). The theory is still evolving, but it has been recommended over classical test theory, which encompasses all common statistical procedures such as ANOVA, MANOVA, multiple regression, factor analysis, discriminate function analysis, and canonical correlation analysis (Arnold,1996; Brennan, 1992; Cronbach, et al., 1972; Eason, 1991; Shavelson and Webb, 1991; Shavelson, Webb, and Rowley, 1989). In particular, generalizability theory provides more accurate reliability estimates than classical theory as well as a method for analyzing more complicated, and thus *more realistic*, models than analyses such as ANOVA. Eason (1991) contends,

> Generalizability theory considers the multiple sources of error that may influence scores, as well as the **interaction effects** of error influences. This more complex model better honors the reality to which the researcher wishes to generalize, since reality usually involves multiple sources of error variance that can also interact with each other (boldface emphasis in original, p. 84).

Her recommendation is based partially on the fact that with classical theory it is impossible to consider multiple sources of error variance in a single analysis, which she classifies as a severe limitation of the theory.

Shavelson et al. (1989) characterize generalizability theory as "broader and more flexible" than classical test theory and go on to say, "generalizability theory asks how accurately observed scores permit us to generalize about persons' behavior in a defined universe of situations" (p. 922). Many researchers wish to make inferences that generalize over several factors (for instance, occasions, items, forms, raters, observers, etc.), and generalizability theory is precisely what is needed in order to accomplish that.

Generalizability theory has not typically been used to study issues in undergraduate mathematics education. One study that used generalizability theory at this level of education is *An Investigation in the Use of Performance Assessment Tasks in Undergraduate Mathematics* (Fine, 1995). This study compared scores on performance assessment tasks with scores on a series of multiple choice questions targeting roughly the same knowledge domains. The subjects were students enrolled in college algebra. The main finding was that performance assessment tasks probably measure a different aspect of mathematical understanding than that tapped by short-answer tests. Additionally, a moderately large amount of the total variation in scores (17 percent) was actually attributed to the rater facet. The raters were student workers employed by the mathematics department where the study was conducted, suggesting the need to include raters as a facet in the current study.

Interpretation of Results

To develop an understanding of the interpretation of generalizability results for a two-facet, fully crossed design, consider an example from Shavelson and Webb (1991, p. 22-23): *A study of children's help-seeking behavior.* On two occasions, 13 children were tape-recorded while working mathematics problems. Two raters coded all children's behavior on both occasions. The design has persons (p), the object of measurement, crossed with raters (r) and occasions (o), denoted $p \times r \times o$, and has seven variance components that can be estimated and interpreted using generalizability theory, which are displayed in the following table (the percentages do not sum exactly to 100 due to rounding):

Source of Variation	Estimated Variance Component	Percentage of Total Variance
p	0.3974	35
r	0.0096	1
o	0.1090	10
pr	0.0673	6
po	0.3141	28
ro	0.0064	1
pro,e	0.2244	20

In the "Source of Variation" column, "pr" denotes the interaction between persons and raters; "po" denotes the interaction between persons and occasions; "ro" denotes the interaction between raters and occasions; and "pro,e" denotes the residual, due to the three-way interaction between persons, raters, and occasions and/or other random or systematic sources that were not measured in this study.

The substantial variance component for persons (35 percent of the total variance) indicates that there are systematic differences in the frequency with which the children sought help. The variance component for raters (1 percent of the total), as well as the component for the rater by occasion interaction (1 percent of the total) are both small, indicating that the raters were consistent in general, and in particular, they were consistent on each occasion. In contrast, the component for the person by rater interaction is slightly larger (6 percent of the total), indicating that even though raters used the same standards overall, they disagreed somewhat in the relative standing of the children. This component shows the extent to which one rater "saw" more help-seeking behavior than the other for each child. The slightly larger variance component for occasions (10 percent of the total) indicates that the level of the children's help-seeking behavior is higher on one occasion than another. The large component for the person by occasion interaction (28 percent of the total) indicates that the relative standing of the children differed from one occasion to another. For example, the child who sought the most help on one occasion may have sought the least amount of help on another occasion.

The residual component accounts for a substantial portion (20 percent) of the total variance. To develop an understanding of the interpretation of generalizability results for a two-facet, nested design, consider one additional example from Shavelson and Webb (1991, p. 52-54): *A study of teacher behavior.* On three occasions, 5 teachers were videotaped in the classroom. Then three raters coded each of the 15 tapes in terms of the number of times the teacher directed a student to "try again". For this design, the object of measurement is persons (p), the teachers. The facets are raters (r) and occasions (o). The occasion facet is nested within teachers since there are three occasions per teacher and the occasions differ from teacher to teacher. Rater is crossed with occasions and teachers since each rater coded all teachers on all occasions. This design is partially nested because there are both crossed and nested facets, denoted $(o{:}p)\times r$. It has only five variance components that can be estimated and interpreted using generalizability theory (note that the percentages do not sum exactly to 100 due to rounding):

Source of Variation	Estimated Variance Component	Percentage of Total Variance
p	8.228	32
r	2.613	10
o:p	4.622	18
pr	1.537	6
o:pr,e	9.072	35

The substantial variance component for persons (32 percent of the total variance) indicates that the teachers differed in their behavior. The variance component for raters was also substantial (10 percent of the total), indicating that the raters differed in the amount of behavior they "saw," averaging over teachers and occasions. Occasions are nested within teachers so it is not possible to separate the occasion effect from the effect for the interaction between teachers and occasions. The source "o:p" denotes these combined effects, and the variance component for the combined effects (18 percent of the total) indicates that the teacher behavior differed from one occasion to another, but it is impossible to know whether one occasion produced more behavior than another (the occasion effect), whether the relative standing of teachers differed from once occasion to another (person by occasion interaction), or both. The variance component for the person by rater interaction (6 percent of the total) suggests that the relative standing of the teachers differed somewhat from one rater to another. The rater by occasion interaction is confounded with the three-way interaction between persons, raters, and occasions, and/or other random or systematic sources that were not measured in the study. This large residual component, denoted "o:pr,e" accounts for 35 percent of the total.

Generalizability theory provides a way to estimate multiple sources of variance in a single analysis, and it can be used to estimate interactions that are inaccessible with more commonly used statistical methods. This become particularly important when the sources of variance in a given situation are not well understood.

Summary

There is a tremendous amount of research literature that emphasizes the importance of multiple representations of functions and in particular, the importance of graphical representations. The literature does suggest, however, that more work must be done in the area of assessing understanding of functions and their graphs. Specifically, little work has been done to investigate whether assessment items involving different types of graphical representations actually measure the same abilities. Generalizability theory provides a valuable tool for investigating these potential differences in types of graphing problems.

CHAPTER 3

The Experiment

This chapter describes the key elements of the experiment that was performed as a means to address the questions in Chapter 1 of this study.

Subjects

Subjects were virtually all students enrolled in a section of a first-semester calculus course during the 1998 fall semester at East Central University, a medium-sized regional university in Ada, Oklahoma.

The Sample

The prerequisites for this course are college algebra and trigonometry courses, or equivalent high school courses. Enrollment in College Algebra at this university is contingent upon an ACT math subject score of 19 or higher or a score of 75 or higher on the Elementary Algebra Computerized Placement Test administered by the university.

Thirty-two students, 13 females and 19 males, completed all of the items included in this study. Seventy-nine percent of these students had completed three or more years of high school mathematics, while 47 percent had completed four or more. Seventy-nine percent had completed a trigonometry course in college. Nineteen percent were "non-traditional" students in the sense that they completed high school mathematics courses during the 1980's or earlier. Two were international students, with English not their native language. Sixty-nine percent had an ACT math score of 23 or higher, while 19 percent had a score of 29 or higher.

Protection of Human Subjects

The use of human subjects in this research was approved (see Appendix A) by the University of Oklahoma's Institutional Review Board, as well as East Central University's Professional Development and Research Committee.

Subjects in this study were volunteers. To assure them that refusal to participate would involve no penalty to their course grade, students were informed of this twice in writing: on the informed consent form (see Appendix B) and on the background questionnaire (see Appendix C). Both of these documents were distributed to the students for completion during regular class periods at the beginning of the semester.

Students were given adequate time during class to study the consent form and ask questions pertaining to the research and their participation. They were asked for permission to allow the researcher to include in this study data from their written work that would be required as part of the normal assessment activities for the course. Students were at least 18 years of age, and thus could legally consent to participate in the study. They were informed that they were under no obligation to participate, but that refusing to participate would not exempt them from the class work. They were assured that they could discontinue their participation at any time with no penalty to their grade. In addition, they were assured that their identity would be kept confidential and that they would never be identified in any publication or presentation associated with this research. Each student was assigned a pseudonym to be used in this dissertation and any subsequent publications or presentations.

Students were also informed that time devoted to a voluntary personal follow-up interview might cost them study time, but that the interviews were optional and that they could choose not to participate in them. During the course of the semester, 28 of the 32 subjects chose to participate in the interviews.

The students' grades were never at risk as a result of the collection and analysis of data for the study. However, the researcher anticipated that the students might become concerned that their course grades would be penalized unfairly as a result of being required to perform tasks they may perceive as unusual or non-standard. Care was taken to protect against the potential perception of unfair grading policies. Problems from the textbook similar in nature to the target problems included on examinations and quizzes were assigned as regular homework and in-class assignments. The students were told at the beginning of the semester that the study would involve no unusual tasks and that the target problems would simply be problems they would work during the normal course of the semester.

Course Description

This course basically followed a lecture format but incorporated the use of small group activities many times during the semester. The class met daily, Monday through Friday, for 50 minutes each day. The researcher was the instructor for the course. The textbook used was *Calculus: Graphical, Numerical, Algebraic* by Ross L. Finney, George B. Thomas, Franklin Demana, and Bert K. Waits (1994). During the semester, differential calculus was covered in its entirety. Additionally, an introduction to integral calculus was included.

Graphing calculators were required of all students in the course and were used extensively throughout the semester. A particular emphasis was placed upon demonstrating the shortcomings of the calculators. Students were encouraged to use them during class as well as outside of class but were challenged to consider the appropriateness and/or necessity of their use on any given problem. The textbook supported the use of graphing calculators, providing examples and exercises that highlighted their strengths and weaknesses. Calculator-based homework problems were assigned on a regular basis to reinforce topics covered in class. Students were allowed to use their calculators for all homework, quizzes, and examinations.

Method

This study examined written data from quizzes and examinations given during the normal assessment activities of the course. All assessments were administered during regularly scheduled class periods. The study was designed to cause no interruption in the regular activities of the course, and there were no special or unusual interventions related to the study. Data gathered were ordinary problem solutions collected during the semester (see Appendix E for examples of student solutions).

Items

Two types of problems were targeted for evaluation in this study. The first type (Type 1) was a traditional type of graphing problem in which a graph is given, and several questions are asked about that graph. Questions covered concepts such as the coordinates of points in the plane, limits, asymptotes, and properties of the first and second derivatives of the given function. The second type of problem (Type 2) was less traditional. Type 2 problems asked students to sketch a graph that met several specified conditions at the same time. The conditions presented in Type 2 problems closely paralleled the kinds of questions asked in Type 1 problems.

These conditions fell into 4 categories:

(i) Those calling for graphs passing through specific points in the plane;

(ii) Those calling for certain kinds of limiting behavior, including asymptotic behavior;

(iii) Those based on or related to the first and second derivative (e.g., critical points, relative extrema, concavity, increasing or decreasing)

(iv) Those calling for points of discontinuity or nondifferentiability.

Additionally, students often had to deal with the *implicit* interaction of two or more specified conditions. For example, one target problem asked students to sketch the graph of a function f that satisfied 6 different conditions. One condition was "$f'(-1)=0$", which implies the function f either has a relative maximum, minimum, or inflection point at $x=-1$. Another condition was "$f''(-1)=-9$", which implies the function f must be *concave down* at $x=-1$. The *interaction* of the two conditions implies the function f must have a relative *maximum* at $x=-1$.

In other Type 2 problems (graph-to-graph Type 2), students were required to generate the "specified conditions" for themselves, based on another graph. For example, one target problem gave the graph of a function f and asked the students to sketch a possible graph of the derivative f'.

Two problems of each type were included on four different quizzes and on two different hour examinations for a total of 24 target problems (see Appendix D). These problems were similar to textbook problems as well as traditional examination and homework problems that might arise in any calculus course. Two problems of each type were required on each quiz and examination in order to allow an investigation of the consistency in items of the same type. Both types of problems were required on each quiz and examination in order to allow the researcher to estimate any variance due to differences between the types. One type may be more difficult than the other, or the two problem types may draw on different student abilities.

Partial Credit Rubrics

Partial credit rubrics were designed for each item. These rubrics were based on analyses of the questions asked in the Type 1 items or the conditions and interactions in the Type 2 items. Points were assigned for each question or condition. This resulted in considerable variation in the number of possible points for the items. Therefore, in the actual data analysis, the scores were converted to proportion of possible points received so that each item would be scored on a comparable metric. Small amounts of incomparability were unavoidable given the different "grain sizes" of the rubrics.

Raters

The performance of each of the 32 students on each of the 24 items was independently evaluated by two different raters (the researcher and a fellow graduate student) to allow an investigation of consistency in scoring. Immediately after the students had completed each quiz or examination containing target problems, two copies of each student's work was made, one for each rater, before the problems were graded by the instructor as part of the course grade. Each rater independently scored each item using the partial credit rubric for that item. These scores constituted the primary data for analyses.

Stakes

Target problems were included on quizzes as well as on hour examinations in order to allow the researcher to estimate any variance in scores due to attitudinal differences that students might have toward a quiz compared to an examination. Students were informed on the first day of class that their quiz scores would be scaled to account for 10 percent of their final grade, their examination scores would be scaled to account for 65 percent of the their final grade, and the remainder of their grade would be determined by homework scores. It might be expected that the difference in weights of quiz grades verses examination grades might cause students to prepare differently for a quiz than they would for an exam.

Occasions

Target problems were included on two different hour examinations in order to allow the researcher to estimate any variance in scores due to the students' writing the examinations on different days. Target problems were included on four different quizzes for a similar reason.

Interviews

In addition to the student work that was collected, voluntary personal interviews were conducted with a small number of subjects following each quiz or examination that included target problems. Students were contacted individually concerning their participation in an interview, and it was made clear to them that they were under no obligation whatsoever to participate. The only students in the class who experienced anything different from what they would experience in a non-experimental calculus class were the students who agreed to be interviewed. The purpose of the personal interviews was to enable the researcher to develop a better understanding of the students' thought processes and the general strategies they used as they were solving the problems. By interacting with them and asking them questions about their work, the researcher gained a better perspective on important aspects of their thought processes. In addition, the researcher was able to ascertain whether they felt one type of problem was more difficult than the other and whether or not they liked these types of problems. To avoid influencing their comments, the researcher informed each student at the beginning of the interview that any questions regarding "correct" problem solutions would be addressed *after* the interview was completed. Interviewees were encouraged to voice what they were thinking and feeling, without worrying about whether they were "right" or "wrong" since there were no incorrect answers in this context.

Design

Several analytical methods were available. If the variability in student responses had been well understood and was primarily due to student differences, a simple one-way analysis of variance to compare the two item types would have been appropriate. Since little was known about the relative effects of several factors affecting the variability of student responses, generalizability theory was chosen as the primary analytical method for this study. This theory was first developed and published in 1972 (Cronbach, Gleser, Nanda, and Rajaratnam). Shavelson and Webb (1991) provided an accessible introductory treatment. The computer program *GENOVA: A Generalized*

Analysis of Variance System (Crick and Brennan, 1983) was used for all generalizability analyses in this study.

Generalizability theory essentially decomposes variance into components due to numerous sources. With generalizability theory, it is not necessary to restrict decomposition of variance to only the component due to differences among individuals and the residual composed of systematic and random sources labeled "error" variance. In particular, this theory allows the proportion of overall variance due to differences in individual performances on a behavioral measurement to be compared with that due to various other factors in the measurement situation. Hence a generalizability study should incorporate into its design as many potential sources of variance as are of interest. These multiple sources of variance can be estimated separately in a single analysis. Since at least two instances of any source (student, rater, etc.) must be included and these must be combined in the design, the number of instances grows quickly through combinatorial "explosion." This makes it necessary to limit variance sources considered to those most important in the situation and can lead to degrees of freedom difficulties if sample size is not large. In this study, the sources of variation of interest were: differences among students, items, problem types, raters, stakes (examination vs. quiz), occasions, and all interactions.

In variance decompositions using generalizability theory, a final residual component remains which confounds random error and systematic influences from sources not explicitly included or controlled in the design with the first level of unexamined interactions (for example, three-way interactions). When the relative contribution of sources of variation in a measurement situation are unknown, generalizability theory offers a powerful analytic approach for examining the role of each source (Shavelson and Webb, 1991).

Variability among persons being studied reflects systematic differences in their knowledge, skills, and abilities. This source of variance is called the *object of measurement.* Each major source of variance is called a *facet* in generalizability theory, analogous to a *factor* or *main effect* in an ANOVA design. Any facet may become the object of measurement. The *conditions* of a facet are analogous to the *levels* of a factor in an ANOVA design.

Facets may be *crossed* or *nested.* Facets A and B are crossed when all conditions of Facet A are observed with all conditions of Facet B. Facet A is nested within Facet B if two or more conditions of Facet A appear with one and only one condition of Facet B. With nested facets, it is not possible to estimate all variance components separately, so crossed facets are preferred over nested facets whenever possible. A design should be as fully crossed as possible to estimate as many sources of variability as possible (Shavelson and Webb,1991, p. 55).

Pilot Study

The researcher conducted a pilot study during the 1997 spring semester using "found" data that had been generated in the normal assessment activities of a first semester calculus course during the 1996 fall semester at the University of Oklahoma. This pilot study served as a valuable tool for clarifying the experimental design needed for the current study. In the pilot study, a sample of 53 students were evaluated by two raters (the researcher and the instructor for the course) on 8 target problems (four Type 1 and four Type 2) for a completely crossed, two-facet design with student as the object of measurement:

$$\text{student} \times \text{problem type} \times \text{rater.}$$

The effect of the interaction of students with problem types was overwhelming, accounting for almost 60 percent of the total variance, indicating the need to modify the design (Braddy and McKnight, 1997).

There appeared to be a clear explanation for the huge interaction component. Since the pilot study used "found" rather than designed data, different problem types were taken on different occasions throughout the course. From the data, students apparently were differentially prepared on different occasions. The problems were completed in different circumstances (quizzes, hour examinations, final examination), which seems clearly to have affected student performance, and presumably, preparation. The occasion of measurement was confounded with differences among problem types (Type 1 vs. Type 2) and stakes (quiz vs. hour examination). Using two problems of each type on each occasion for this current study should allow the separation of the effect of problem types and the effect of occasions within stakes. This calls for a partially nested, five-facet design with student as the object of measurement:

$$\text{student} \times \text{rater} \times \text{(item within problem type)} \times \text{(occasion within stakes).}$$

Nesting items within problem type causes the item effect to be confounded with the effect for the item by type interaction. As a result, the item effect cannot be estimated separately from the effect for the item by type interaction. Similarly, the occasion effect cannot be estimated separately from the effect for the occasion by stakes interaction.

Occasionally, an estimate of a variance component may be negative. According to Shavelson and Webb (1991, p. 37), such negative estimates may arise because of a misspecification of the model or because of sampling error. If negative estimates are large in relative magnitude, chances are that they are a result of a misspecification of the model, and the model should be re-examined. If negative estimates are small in relative magnitude, near zero, they probably are a result of the true value of the component being at or near zero.

Since negative variance is conceptually impossible, Cronbach et al. (1972) recommends setting the negative estimate to zero. For any further calculations of variance components using that negative component, a value of zero, rather than the negative value, is used. The strength of this method is that it immediately eliminates the theoretical impossibility of a negative variance component. Brennan (1992) also recommends setting the negative estimate to zero but recommends using the original negative estimate in further calculations. The strength of this method is that it provides unbiased estimates of the variance components. Either method is commonly used, but Cronbach's method "produces biased estimates of variance components because they were calculated with the negative estimate set to zero" (Shavelson and Webb, 1991, p. 38). The GENOVA computer program (Crick and Brennan, 1983) used for this study provides two different estimates, one using Cronbach's method and one using Brennan's. This study reports the estimates generated using Brennan's method.

Data Analysis

Two raters independently evaluated the performance of each of 32 students on the 24 different items. These original scores were each converted to a proportion of possible points and these proportions constituted the data set used for all generalizability analyses. Due to the final sample size, the researcher anticipated poor results from the analysis of the full five-facet partially nested model that had originally been specified for this study. As expected, the analysis resulted in a large number of negative variance components, many of which were also large in relative magnitude. Of the 35 different main and interaction effects that could have been estimated by such a model, only 18 had non-negative variance components. Clearly, the small sample size and degrees of freedom could not provide an accurate analysis with such a complicated model.

As a result, the researcher conducted a series of analyses using smaller, less complicated models. Many of these models excluded one or both of the nested facets by collapsing over the conditions. Many of the models that provided significant results were fully crossed designs, which provide estimates of more sources of variability than nested designs with the same facets. Some of these simpler models were the result of eliminating the stakes facet by examining quizzes separately from examinations. Most of the models excluded the rater facet after it was established that there indeed was no effect due to differences in raters. In addition, the researcher re-examined the analysis of many of these models in the context of subgroups based on each student's final course grade. This was done to increase homogeneity in an attempt to reduce the variability due to systematic differences among the students.

Information obtained from audio-tapes of the personal interviews was used as additional confirmation of effects suggested by the quantitative analyses. This information was also used to assess students' attitudes toward and perceptions of the different types of graphing problems investigated in this study.

CHAPTER 4

Results

The data analyses were carried out as described in Chapter 3. The computer program used for all generalizability analyses was *GENOVA: A Generalized Analysis of Variance System* (Crick and Brennan, 1983).

Generalizability Analyses

The initial analysis using the full five-facet, partially nested model

student × rater × (item within problem type) × (occasion within stakes)

originally specified for this study yielded poor results, as anticipated. Seventeen of the 35 main and interaction effects had estimated variance components that were negative, many of which were large in magnitude, confirming the fact that there are degrees of freedom problems associated with the small sample size relative to such a complicated model. This indicated the need for a series of analyses using smaller, less complicated models.

Rater Effects

One way to simplify the model was to examine the rater effect, and in the absence of any effect due to raters, the rater facet could be eliminated from further consideration. The next set of analyses incorporated a two-facet, fully crossed design with students crossed with raters and items. The goal of these analyses was to examine consistency in scoring by the two raters. In the first analysis, students were crossed with raters and the 12 Type 1 items, which yielded the results shown in Table 1.

Table 1 **MODEL: p x r x i**
Type 1 items

Source of Variation	Estimated Variance Component	Percentage of Total Variance
p	0.0145319	22.4
r	0.0000029	0
i	0.0121790	18.8
pr	0*	0
pi	0.0377621	58.2
ri	0*	0
pri,e	0.0004223	0.6

**A negative component was set to zero*

The estimated variance components for the student by rater and rater by item interactions were negative but very small in relative magnitude, indicating the true value of the components are at or near zero. Following the procedure described by Brennan (1992), these estimates were set to zero only after the remaining variance components were calculated using the negative values. This convention will be used for the remainder of the presentations as well. A point will be made when any of the negative components were large in relative magnitude.

The results in Table 1 show no main effect due to raters, indicating that raters were consistent in their stringency. There is no effect due to the interaction of raters and students or due to the interaction of raters and items. This indicates that the raters were consistent from student to student as well as from item to item for all Type 1 items. The residual component is negligible, indicating that very little variance is due to the three-way interaction and sources that are not included in this model.

In the next analysis, students and raters were again crossed with items, but only the 12 items of Type 2. This analysis yielded the results shown in Table 2. These results follow the same pattern as those for the Type 1 items. The major point in both Tables 1 and 2 is the lack of differences due to raters.

In the final analysis involving students crossed with raters and items (see Table 3), all 24 items were included making no distinction between the different types.

Table 2 **MODEL: p x r x i**
Type 2 items

Source of Variation	Estimated Variance Component	Percentage of Total Variance
p	0.0125327	23.0
r	0.0000004	0
i	0.0132233	24.3
pr	0*	0
pi	0.0278394	51.2
ri	0*	0
pri,e	0.0008019	1.5

******A negative component was set to zero*

The results shown in Table 3 follow the same pattern as those for the separate analyses of Type 1 and Type 2 items, again with the major observation being the lack of differences due to raters.

Table 3 **MODEL: p x r x i**
All items

Source of Variation	**Estimated Variance Component**	**Percentage of Total Variance**
p	0.0138909	23.4
r	0*	0
i	0.0123546	20.8
pr	0*	0
pi	0.0325024	54.7
ri	0*	0
pri,e	0.0006790	1.1

**A negative component was set to zero*

These three analyses confirm the consistency of the raters across both types of items. Based on these results, the researcher felt confident in the correctness of the decision to collapse the scores over raters for the remainder of the analyses, thereby eliminating the need for a rater facet in later models.

The extremely large component for the student by item interaction shown in each of the Tables 1 through 3 (from 51.2 to 58.2 percent of the total) indicates the relative standing of the students changed from item to item, which is the same kind of result obtained in the pilot study. This confirms the need to include an occasion and/or stakes facet in the design.

Occasions and Stakes

In an attempt to decompose the item main effect shown in the previous results, the researcher incorporated occasions within stakes as a facet in the design. It was suspected that the variance due to items in the previous results was due to factors other than differences among the items within types. Thus, the next analyses were based on a model that had students crossed with items and occasions nested within stakes. The separate results for Type 1 and Type 2 items are shown in Tables 4 and 5.

Table 4 **MODEL: p x i x (o:s)**
Type 1 items

Source of Variation	Estimated Variance Component	Percentage of Total Variance
p	0.0389180	13.3
s	0*	0
o:s	0.0378591	13.0
i	0*	0
ps	0*	0
po:s	0.0509612	17.5
pi	0.0231258	7.9
si	0.0239721	8.2
oi:s	0.0059966	2.1
psi	0.0075123	2.6
poi:s,e	0.1033773	35.4

**A negative component was set to zero*

The nesting of the occasions facet within the stakes facet is denoted "o:s". The notation "po:s" denotes the interaction of students with occasions along with the three-way interaction of students, occasions, and stakes. The results in Tables 4 and 5 suggest the items within each type are comparable but that there are differences due to occasions within stakes.

Table 5 **MODEL: p x i x (o:s)**
Type 2 items

Source of Variation	Estimated Variance Component	Percentage of Total Variance
p	0.0387819	19.8
s	0*	0
o:s	0.0057838	3.0
i	0*	0
ps	0.0077687	4.0
po:s	0.0057451	2.9
pi	0.0048601	2.5
si	0.0000000	0
oi:s	0.0237520	12.2
psi	0.0000000	0
poi:s,e	0.1085902	55.6

**A negative component was set to zero*

The occasion effect cannot be estimated separately from the occasion by stakes interaction due to the nesting in the designs shown in Tables 4 and 5.

However, it seems likely that systematic differences due to occasions would exist. Students might prepare differently for a quiz than they would for an examination, given the fact that quizzes and examinations carry different weights in the calculation of final grades. In each set of results, the three negative variance components were large in relative magnitude, and a large amount of variation (55.6 percent) is accounted for by either random error or systematic differences not accounted for in this model. As a result, additional models were analyzed.

Table 6 displays the analysis results of a model that crossed students with types and stakes, averaging over all quizzes and both examinations. This eliminates occasion from separate consideration in this analysis.

Table 6 **MODEL: p x t x s**

Source of Variation	Estimated Variance Component	Percentage of Total Variance
p	1.5809301	8.6
t	0.0544126	0.3
s	15.2066723	83.1
pt	0.0708687	0.4
ps	0.9542808	5.2
ts	0*	0
pts,e	0.4383512	2.4

**A negative component was set to zero*

An overwhelming amount of variance (83.1 percent of the total) is attributed to differences due to the stakes main effect, with an additional 5.2 percent due to the interaction of students and stakes. The main effect indicates that stakes were systematically different, meaning stakes were consistently different for all students. The interaction component indicates an additional small amount of variance is due to differences in relative standings of the students from quizzes to examinations. While this analysis confounds occasion effects with others by ignoring occasions, the large effects for stakes are still suggestive. This is confirmation of the notion that students in this study did prepare differently for quizzes than they did for examinations.

Table 7 displays the results of the analysis of a model in which students are crossed with occasions and items nested within types using all 6 occasions and all 24 items. Only one of the negative variance component estimates was large in relative magnitude.

Table 7 **MODEL: p x o x (i:t)**

Source of Variation	Estimated Variance Component	Percentage of Total Variance
p	0.0519397	19.8
o	0.0399110	15.2
t	0.0018384	0.7
i:t	0*	0
po	0.0313709	12.0
pt	0.0007517	0.3
pi:t	0*	0
ot	0*	0
oi:t	0.0219198	8.4
pot	0*	0
poi:t,e	0.1147935	43.7

**A negative component was set to zero*

In Table 7, a substantial[1] occasion main effect appears, as expected, along with a substantial student by occasion interaction. There is a moderate interaction of occasions with items within types. This time stake effects are confounded with other effects since stakes are ignored. The researcher suspected the occasions effect might be masking any effect due to types. Therefore, separate analyses were run for the two different item types. Tables 8 and 9 contain the results of the separate analyses of Type 1 and Type 2 items. The main effect of items is 0 in both analyses, indicating the Type 1 items are comparable and the Type 2 items are comparable. This serves as additional confirmation of the previous results suggesting items within each type are comparable.

---

[1] Components that are around 10 percent or more are referred to as "substantial" by Cronbach (1972), p. 176, and by Shavelson and Webb (1991), p. 54.

The substantial effect for the interaction between items and occasions (11.6 percent of the total) shows that some items varied in difficulty from occasion to occasion. Thus there were no consistent differences over all items of one type or the other, regardless of occasion, there were only differences among items caused by differences in occasions.

Table 8 **MODEL: p x o x i**
Type 1 items

Source of Variation	Estimated Variance Component	Percentage of Total Variance
p	0.0535626	20.4
o	0.0218032	8.3
i	0*	0
po	0.0271996	10.3
pi	0.0048369	1.8
oi	0.0303922	11.6
pro,e	0.1250359	47.6

**A negative component was set to zero*

Table 9 **MODEL: p x o x i**
Type 2 items

Source of Variation	Estimated Variance Component	Percentage of Total Variance
p	0.0517060	22.7
o	0.0449600	19.7
i	0*	0
po	0.0129809	5.7
pi	0*	0
oi	0.0135273	5.9
pro,e	0.1047295	46.0

**A negative component was set to zero*

Based on the information in Tables 8 and 9, the researcher decided to analyze quizzes and examinations separately using a model that crossed students with occasions and items. Types again were analyzed separately. Tables 10 and 11 contain the results of the separate analyses of the Type 1 items on quizzes and Type 1 items on examinations, while similar results for the Type 2 analyses are listed in Tables 12 and 13.

The item effect for the Type 1 quiz items (8.6 percent) shown in Table 10 indicates that these items varied somewhat in level of difficulty. This was a surprising finding, suggesting that graphing problems may not be as similar as they might appear at first glance.

Table 10 **MODEL: p x o x i**
Type 1
Quizzes only

Source of Variation	Estimated Variance Component	Percentage of Total Variance
p	0.0528428	19.4
o	0.0359206	13.2
i	0.0232964	8.6
po	0.0205242	7.5
pi	0.0013952	0.5
oi	0.0091255	3.4
pro,e	0.1291011	47.4

**A negative component was set to zero*

There is a small effect due to the variability of Type 1 examination items (4.3 percent), shown in Table 11, suggesting these differed little in difficulty.

Table 11 **MODEL: p x o x i**
Type 1
Examinations only

Source of Variation	Estimated Variance Component	Percentage of Total Variance
p	0.0139163	4.8
o	0.0473669	16.3
i	0.0123800	4.3
po	0.0662268	22.8
pi	0.0356512	12.3
oi	0.0088960	3.1
pro,e	0.1058345	36.4

**A negative component was set to zero*

The results in Tables 12 and 13 suggest that Type 2 items are comparable based on the fact that there is no main effect for Type 2 items on quizzes and a negligible effect (1.3 percent) for those same type items on examinations.

Table 12 **MODEL: p x o x i**
Type 2
Quizzes only

Source of Variation	Estimated Variance Component	Percentage of Total Variance
p	0.0594269	21.9
o	0.0570701	21.0
i	0*	0
po	0.0108057	4.0
pi	0*	0
oi	0.0213034	7.8
pro,e	0.1229425	45.3

**A negative component was set to zero*

Table 13 **MODEL: p x o x i**
Type 2
Examinations only

Source of Variation	Estimated Variance Component	Percentage of Total Variance
p	0.0409940	24.5
o	0.0367093	22.0
i	0.0022339	1.3
po	0.0126016	7.5
pi	0*	0
oi	0*	0
pro,e	0.0746709	44.7

**A negative component was set to zero*

The original design of the full model for this study was *unbalanced* because the stakes facet has a different numbers of levels or occasions for quizzes (four occasions) and examinations (two occasions). A *balanced* design is required in order to use the GENOVA program for data analysis. A similar requirement exists for other computer programs such as BMDP3V in the BMDP statistical package and the VARCOMP procedure in the SAS package. For an unbalanced design, the recommendation is to randomly delete levels of a facet to create a balanced design (Shavelson and Webb, 1991). According to Shavelson, Webb, and Rowley (1989), such random deletions have little effect on the estimated variance components. Shavelson and Webb (1991) additionally recommend comparing the results of several random deletions to ensure that any particular selection is not unusual. For this study, all possible combinations of 2 quizzes with the 2 examinations were analyzed. Each model had students crossed with types and occasions within stakes. For this set of analyses, the results using different pairs of quizzes were similar, with the analysis using Quiz 1 and Quiz 4 showing the largest amount of variability due to types. The results of this particular analysis are given in Table 14. Only one of the negative variance components was large in relative magnitude. The main effect due to type accounted for 3.9 percent of the total variation, indicating a slight difference in types for all students. An additional 4.2 percent of the variation was due to the interaction of students and types, indicating that the rank ordering of students differed from Type 1 to Type 2. This 4.2 percent is above and beyond the systematic differences due to students, which accounted for 28.2 percent of the total variation. Note the large amount of variability due to occasions within stakes: 14.8 percent with an additional 13.2 percent due to the interaction of students with the occasions within stakes.

Table 14 **MODEL: p x t x (o:s)**
Quiz 1, Quiz 4

Source of Variation	Estimated Variance Component	Percentage of Total Variance
p	0.1934602	28.2
t	0.0267852	3.9
s	0*	0
o:s	0.1018069	14.8
pt	0.0291710	4.2
ps	0.0259882	3.8
po:s	0.0905728	13.2
ts	0.0034835	0.5
to:s	0*	0
pts	0*	0
pto:s,e	0.2230545	31.4

**A negative component was set to zero*

Since the occasion effect could not be estimated separately from the occasion by stakes interaction for the model shown in Table 14, the next model analyzed was one that crossed students with types and occasions, using only Quiz 1 and Quiz 4. Table 15 contains the results of this analysis. Note that the main effect due to type is 6.9 percent of the total, the effect due to occasion as a main effect is reduced to 3.7 percent, with an additional 8.1 percent variation coming from the interaction of students with occasions. Similar analyses were run using the same model with different subsets of occasions. The results of these analyses were similar to, though not as marked as those shown in Table 15.

Table 15 **MODEL: p x t x o**
Quiz 1 and Quiz 4 only

Source of Variation	Estimated Variance Component	Percentage of Total Variance
p	0.2974413	45.3
t	0.0453198	6.9
o	0.0245092	3.7
pt	0*	0
po	0.0534080	8.1
to	0*	0
pto,e	0.2356845	35.9

**A negative component was set to zero*

Student Subgroups

Due to the large main effect due to students with this model (45.3 percent of the total variability), the researcher made the decision to conduct analyses on several sets of subgroups of students. The subgroups were formed based on the students' final grade in the course in an attempt to form more homogeneous groups, and thereby reduce the variability due to students. Four subgroups were focused upon: (1) the 17 students whose final grade in the course was a B, C, D, or F; (2) the 15 students whose final grade in the course was an A (final average 90 percent or above); (3) the 26 students whose final average in the course was below 95 percent; (4) the 6 students whose final average in the course was 95 percent and above.

First of all, a model that crossed students with types was analyzed for each occasion individually. The most important results of these analyses are presented in Tables 16 through 23. In the results presented in these eight tables, the percentage of variance due to types is larger than the percentages seen previously. However, it should be born in mind that there are only two facets in this design, which makes the proportion due to students appear larger.

Table 16 **MODEL: p x t**
Examination 2
A students

Source of Variation	Estimated Variance Component	Percentage of Total Variance
p	0.0471429	52.7
t	0.0038571	4.3
pt,e	0.0384762	43.0

Table 17 **MODEL: p x t**
Examination 3
BCDF students

Source of Variation	Estimated Variance Component	Percentage of Total Variance
p	0.2954706	60.8
t	0.0266544	5.5
pt,e	0.1638279	33.7

Table 18 **MODEL: p x t**
Quiz 4
All students

Source of Variation	Estimated Variance Component	Percentage of Total Variance
p	0.3465373	59.7
t	0.0292976	5.0
pt,e	0.2047040	35.3

Note that there are no negative variance components in any of these results in Tables 16 through 23, and in all cases, the model accounted for the majority of the variance.

Table 19 **MODEL: p x t**
Quiz 4
BCDF students

Source of Variation	Estimated Variance Component	Percentage of Total Variance
p	0.4678664	68.8
t	0.0433691	6.4
pt,e	0.1686809	24.8

Table 20 **MODEL: p x t**
Quiz 1
All students

Source of Variation	Estimated Variance Component	Percentage of Total Variance
p	0.3551613	55.0
t	0.0489966	7.6
pt,e	0.2411347	37.4

Table 21 **MODEL: p x t**
Quiz 1
Students with average below 95 percent

Source of Variation	Estimated Variance Component	Percentage of Total Variance
p	0.3535385	52.0
t	0.0596168	8.8
pt,e	0.2668563	39.2

The percent of variance that can be attributed differences between Type 1 and Type 2 problems on Quiz 1 is 7.6 percent when all students are analyzed together (see Table 20), but increases to 8.8 percent when the six students with an average of 95 percent and above are removed from the analysis and only the 26 students with averages below 95 percent are considered (see Table 21).

Table 22 **MODEL: p x t**
Quiz 1
BCDF students

Source of Variation	Estimated Variance Component	Percentage of Total Variance
p	0.3473529	48.3
t	0.0859967	11.9
pt,e	0.2864121	39.8

When all of the A students are removed from the analysis of Quiz 1 and only the 17 students who made a B, C, D, or F in the course are considered, the variance due to differences between the two types of problems jumps to a substantial 11.9 percent (see Table 22). Table 23 shows 11.6 percent of variance is due to differences in types on Quiz 5 for the A students.

Table 23 **MODEL: p x t**
Quiz 5
A students

Source of Variation	Estimated Variance Component	Percentage of Total Variance
p	0.2580757	51.2
t	0.0581829	11.6
pt,e	0.1875771	37.2

A final design of persons crossed with type and occasion was examined for three of the subgroups of students. The results for these analyses using only Quiz 1 and Quiz 4 are shown in Tables 24 through 26.

Table 24 **MODEL: p x t x o**
Quiz 1 and Quiz 4 only
A students

Source of Variation	Estimated Variance Component	Percentage of Total Variance
p	0.0411286	12.5
t	0.0186607	5.7
o	0.0434902	13.2
pt	0.0250893	7.6
po	0.0032664	1.0
to	0*	0
pto,e	0.197281	60.0

**A negative component was set to zero*

The amount of variance due to types is almost twice as large for the B, C, D, and F students (10 percent; see Table 25) as for the A students (5.7 percent; see Table 24).

Table 25 **MODEL: p x t x o**
Quiz 1 and Quiz 4 only
BCDF students

Source of Variation	Estimated Variance Component	Percentage of Total Variance
p	0.3096805	40.0
t	0.0772417	10.0
o	0.0074825	1.0
pt	0*	0
po	0.0979292	12.6
to	0*	0
pto,e	0.2821118	36.4

**A negative component was set to zero*

Table 26 shows that a substantial amount of variance (12.6 percent) is due to differences in item types on Quiz 1 and Quiz 4 when only the top six students in the class are considered.

Table 26 **MODEL: p x t x o**
Quiz 1 and Quiz 4 only
Students with average 95 percent and above

Source of Variation	Estimated Variance Component	Percentage of Total Variance
p	0*	0
t	0.0470267	12.6
o	0.0055233	1.5
pt	0*	0
po	0.0865600	23.2
to	0*	0
pto,e	0.2334967	62.7

**A negative component was set to zero*

For the results in each of the Tables 25 and 26, one negative variance component was large in relative magnitude.

The results reported in this study were generated using models in which all facets were considered *random.* If type is considered a *fixed* facet, similar differences result, although in some cases, the effects are slightly smaller. All facets were considered random in order to allow the results to be generalized as broadly as possible, and in particular, to graphing problems of any sort.

Interviews

Personal interviews were conducted throughout the semester. These follow-up interviews were conducted as quickly as possible after an examination or quiz was given that contained target problems (see Appendix E for examples of student solutions). Twenty-four of the 32 students in the class chose to participate in at least one interview. Five students agreed to be interviewed twice, and five other students agreed to participate in an interview after *all* quizzes and examinations with target problems were given. These students provided the qualitative data for five "case studies."

One of the more interesting and surprising findings was in the answers to the question, "Are both types of problems (Type 1, Type 2) equally difficult for you?" Based on the belief that it is a more complex task to produce a graph meeting several specified conditions at once than to simply answer questions based on a given graph, it might be expected that the Type 2 problems would be more difficult for everyone. Not surprisingly, 18 of the 24 students interviewed confirmed that indeed Type 2 problems were more difficult for them. Two students said the different problem types were about the same level of difficulty.

Surprisingly, four students said the Type 2 problems were *easier* for them, even "enjoyable" and "fun." The natural assumption at this point might be that these four students were the students with the highest grades in the class. Interestingly, the highest final average among these four students was 88 percent, certainly *not* among the highest grades in the class. More interesting was the fact that *three* of these four students had the three *lowest* averages in the class. After all the grades were curved, these four students' final averages were 59 percent, 63 percent, 64 percent, and 88 percent. Even though these students claimed the Type 2 problems were easier than the Type 1, there was no significant difference in their performance on the two types of problems.

Additionally, the two students who felt the target problems were about the same level of difficulty also did not have the highest averages in the class. Their averages were 90 percent and 94 percent.

Reactions to Ambiguity

The researcher investigated the students' reasons behind their opinions regarding the difficulty levels of the two types of target problems. During her first interview, Emma[2] responded to the researcher's follow-up question, "Why are Type 1 problems easier for you?" by saying,

> Well, for me, I'm a visual person, and so if I already have it in front of me...then it's pretty easy for me to look at it and figure it out, but then if I have to come up with something on my own, using things that are given, I just find that a lot harder. And also, because it's a lot of times open-ended. You know, there's [*sic*] several ways you can draw it. And *that's* [3] difficult for me because it's easier to me if there's an exact answer...Also, I think it's hard especially like when you have like a through d (conditions), and so you have to match each one up, it's harder to get it right that way... I noticed that I have to go back and put check marks to make sure, ok, this doesn't contradict this, and that, you know,

---

[2] Not her actual name. Pseudonyms were used to protect the anonymity of the students.
[3] Italics reflect recorded oral emphasis.

> that they all go together. That makes it more difficult... You have to keep checking each one to make sure that they don't contradict each other. And I think that makes it more difficult. You don't have to worry about each one of these (questions asked when a graph is given) contradicting each other because they're all right there in front of you, and you have the whole graph.

Other students' opinions coincided with Emma's in the sense that the possibility of more than one right answer was disconcerting to them. These students often indicated during the interviews that they were more comfortable with the Type 1 problems because questions on these problems did have "one right answer". Kristen expressed her opinion this way:

> There's [*sic*] too many variations [in solutions of Type 2 problems]. When you have a graph and are asked questions, there's just a right answer and a wrong answer. So I think [Type 1 problems are] easier. That's why I like math, better than some classes like chemistry, because in math, a rule is a rule is a rule, and it stays the same, and never changes.

When asked if she thought Type 1 problems were easier for everybody, she replied,

> Kristen: I think it's easier for most people when they have a graph and are asked questions. Because I think it's harder to, for most people, I think that it's harder to make one up themselves.
>
> R: Do you have any idea why that is?
>
> Kristen: No. Not really. But I've heard people, just like sitting in class, [say] "Ok, it's concave up" and they think, well then am I supposed to put it over here? I mean, they don't realize that if it's concave up and you don't have any other conditions, it doesn't matter how skinny or fat you draw it. You know, I think that's hard for them to comprehend, that it doesn't matter what it looks like concave up, and it doesn't even matter if it's increasing or decreasing, if it's concave up, you know. I think that's harder for most people.

Doug said the Type 2 problems were harder for him because,

> It's very hard to graph for me....You are just left by yourself, but on the [Type 1 problems], you can have some help...The graph is there. [On

> the Type 2 problems] there's so much to know. I can't get all of it. I can't remember for more than a minute... because it's not like a formula.

Steve, who also said Type 1 problems were easier, contrasted the two types of problems by comparing Type 2 problems to a jigsaw puzzle:

> Steve: It's easier to understand something when you can actually *see* it. But when you have to put it together, it's kind of like a jigsaw puzzle.
> R: So you're saying, having to draw the graph yourself is like a jigsaw puzzle?
> Steve: Yea. You have to – you can do one thing, and it's no problem, but then you have to get it so they all connect, so it's kind of, a little more difficult.

Ben felt that Type 1 problems were easier as well, but he likened the Type 1 problems to a road map:

> R: Were all of these problems equally difficult for you?
> Ben: No. No, the ones where you provide the graph and ask about the behavior are much easier because it's just like reading a road map.

Macy had an interesting viewpoint. When questioned about her opinion of the difficulty level of the Type 1 problems compared to the Type 2 problems, she said they were about the same after Quiz 1, then continued,

> I like [drawing the graph] better, just because I like to try to meet the qualifications. Because there's [*sic*] so many different ways of doing it. It's kind of like a game that way. Ok, you have to do *this*. Can you do it? (laughing) It's just more of a challenge than [a Type 1 problem in which] you've either got a right or a wrong answer. There *is* a right or wrong answer [on a Type 2 problem], but...there are usually several different ways you can do it.....It's just more of a challenge than [the Type 1 problems].

Macy *liked* the fact that there isn't just one right answer for a Type 2 problem. Todd was the other student who felt the problems were basically about the same level of difficulty. His opinion was:

Todd: Sometimes actually I think, if anything, the ones [for which] you draw the graph yourself are easier....Sometimes they're the same, it just depends on how hard the graph is...

R: So no matter what the questions are like, the ones where you draw a graph are never any harder than having to answer questions?

Todd: No, I don't think so.

Interpretation vs. Creation

David felt the Type 2 problems were easier. During an interview following Quiz 5, he made the following comments:

David: I like the ones where you gave a list. I can do pretty good on those. [Compared to the Type 1 problems], I like the ones where you give us the conditions and we draw the graph.

R: You like those better. Does that mean they're easier?

David: They're easier for me. I understand them better. I like those better.

Katie explained her preference for Type 2 problems by saying,

> I'm pretty ok with drawing stuff, like if you give us *conditions*, and then you ask us to draw a graph, you know. I don't mind doing that. I might mess up sometimes, but I'm *better* at that than when you *give* us something and say, ok, find this and this and this.... I guess [it's] because I can draw it how I *want* to draw it, and [with a Type 1 problem] it's just, like a set thing...it's just, I had this graph, and I'm thinking, what am I doing? But if you just give me conditions, I can draw my *own* and I *understand* what I've done. I like drawing, all the time, better. I *like* it. It makes me think about everything that I do. When *you* [give us the graph], I don't know why, I just don't get it most of the time. I like to do my *own* stuff.

Preston had a similar preference for Type 2 problems and he remarked,

> The easier ones for me are the ones where I have to draw a graph because for *some* reason I can figure out how to draw graphs, but I can't just look at one sometimes and figure out what you're looking for. I've been graphing stuff for so long, it's just kind of sometimes easier for me to do that. We had to graph a lot of stuff in high school.

The student who made one of the most profound statements regarding a potential reason for the difference in level of difficulty of the two problem types was Kyle. After commenting that Type 2 problems were more difficult, he continued,

> Where you're drawing the graph, and there's [*sic*] a lot of vague [conditions], where you can go with it. For me, it creates complications. I'm not sure how I want to draw it. Whereas, when I'm given a graph, I can just interpret it...We talked after this quiz, and I noticed a couple of people I was talking to agreed, that when you're drawing a graph, and there's a lot of vagueness on how the functions can go. You run into a few problems. It's a lot easier to *interpret* than to *create*, that's the way I look at it. When you're drawing a graph, there's an infinite number of right answers, and you have to worry about meeting all these specific conditions as you read down through them. Because you start with these conditions, and you try to meet each condition. Well, you get down farther, and one of these conditions may flip your whole [idea] of exactly what you're looking at.

These issues of *interpretation* and *construction* (i.e., creation) have been defined and discussed in the literature. "*Interpretation* refers to the action by which a student makes sense or gains meaning from a given graph or a portion of the graph.... Construction is quite different than interpretation. Whereas interpretation requires reaction to a given graph, construction requires the student to generate new graphs or equations that are not given" (Demana, Schoen, and Waits, 1993, p. 13). It is worth noting here that up to this point in the semester, the instructor had *never* used these words (interpret, construct, create) in class nor during any of the interviews with the students.

During another interview, in response to the question, "Is one problem harder than another?", Josh made reference to the idea of interpretation as well, saying,

> I think drawing the graph is harder than just looking at a graph and figuring stuff out, because there could be an unlimited number of possibilities on that. Whereas, if you *give* us the graph, all you have to do is be able to interpret it.

Stability of Opinion

Another interesting fact is that none of the students who were interviewed more than once ever changed their basic opinions regarding difficulty level of the problems. Macy, who seemed to enjoy all of the target problems more than any other student in the class, was the only one who wavered a bit in her opinion. Initially, she said the all the target problems on Quiz 1 were about the same level of difficulty. She modified this slightly when she was interviewed after Quiz 3, saying the Type 1 problems were "probably the easiest". However, near the end of the semester during an interview after Quiz 5, she said, "I think they were all about the same to me. Even when we got into the antiderivatives...it still all made sense."

Graph to Graph Problems

There was one exception made by the students in their ranking of the difficulty of all the target problems given during the entire semester. Quiz 5, which was given two weeks before the end of the semester, contained problems that were rated "most difficult of the semester" by all ten students who were interviewed shortly after the quiz. In fact, the Type 2 problems on Quiz 5 *were* of a slightly different nature than the typical Type 2 problems given on all previous quizzes and examinations (See Appendix D). The Type 2 problems on Quiz 5 are classified as "graph-to-graph Type 2" problems because the students were given the graph of a function f and on one problem were asked to sketch a possible graph of the derivative f'. The other problem asked them to sketch a possible graph of the antiderivative $\int f$.

It appears that the added difficulty of these two problems was caused by the fact that the students essentially had to construct the specified conditions for themselves. Emma described it this way:

> This [graph-to-graph Type 2 problem] is a *lot* harder than this [Type 2 problem with specified conditions] because you have to figure out the conditions for the graph before you can think about drawing your own graph.

Kinds of Conditions

Within the rankings of the difficulty of the problems discussed above, a clarification is necessary. Several students made distinctions among the specified conditions in Type 2 problems and similar distinctions among the questions in Type 1 problems. Those distinctions were made between conditions and questions involving (a) the function itself, (b) derivatives of the function, and (c) antiderivatives of the function.

The students referred to conditions in category (b) as "forward" conditions and those in category (c) as "backward" conditions. This terminology was based on the fact that taking the derivative and higher derivatives of a function seemed to be moving in a "forward" direction, but taking antiderivatives seemed to be moving in a "backward" direction. For example, given the graph of f', a question about f'' is a "forward" question, but a question about f is a "backward" question. In terms of level of difficulty, all students who made these distinctions ranked the categories in the following way: (a) was easiest, (b) was more difficult, and (c) was most difficult, regardless of whether they appeared on a Type 1 or a Type 2 problem.

During Emma's interview following Quiz 3, she said the Type 2 problems were more difficult than the Type 1 problems (See Appendix D), but went on to make a distinction between the two Type 2 problems based on the kind of conditions that were given. She said problem number 4 was more difficult for her than problem number 1. She helped clarify the differences in these categories of conditions by saying,

> I think it's because [problem 4] was more straightforward, and all of [the conditions] just deal with f.... And then on this one, on the first problem, it talked about f prime, and the second derivative of f, so you had to think about what that *meant* for the graph of f itself before you could do anything with the information. And [on problem 4], you just talked about f, so it was kind of, look at what's here, and then deal with it, but the first problem, you had to process what was there before you could do anything with it.

Emma made a distinction between conditions in category (a) and category (b). The conditions specified in problem 1 were category (b) conditions, and those in problem 4 were category (a) conditions. She said conditions in category (a) were more "straightforward", but those in category (b) required her to "process" them and determine their relationship to the graph of f before she could produce the graph.

During an interview with Macy, it was revealed that she also thought problem 4 was easier than problem 1. When questioned about her reasoning, Macy responded,

> Um, I don't think you asked much about the second derivative, and stuff like that, which I don't mind the second derivative, I like it. I mean, I *like* those problems.

This indicates the reason problem 4 was easier for Macy was because the conditions did not involve the second derivative. "[S]tuff like that" presumably refers to conditions involving the first derivative or higher derivatives, possibly even antiderivatives. So in

Macy's opinion, it appears that conditions in category (a) were easier than other types of conditions.

During a discussion concerning what kinds of information can be extracted from different graphs, Kristen made a distinction between conditions in category (b) and category (c):

> Kristen: I get confused, because I think, ok if there's a kink in the graph of the first derivative, does that say *anything* about the function...?
>
> R: Well, if there's a kink in the graph of the [first] derivative, does it say anything about the second derivative?
>
> Kristen: Yes, that the second derivative doesn't exist there. I have no problem going *that* way, it's going *backwards* that I think, does it say anything? I can't think if it *says* anything about the function.

Kristen continued to say that going "forward" was no problem for her, but that going "backward" was problematic.

The Number of Conditions

The students made one additional distinction that was a complete surprise. The students who were interviewed said consistently that Type 2 problems with specified conditions varied in difficulty level depending upon the *number* of conditions specified as well as the number of *interactions* involved.

The researcher's expectations initially were that fewer conditions would allow more freedom, thereby making the task easier for the students. Fewer conditions actually make a solution more ambiguous, but this ambiguity caused additional problems for many students. During Kristen's first interview with the researcher, which took place after Quiz 3, she indicated the difficulty level of a Type 2 problem was related to the number of conditions given in the problem.

Kristen: I think it's easier when you have more boundaries.

R: More boundaries in terms of, like, conditions?

Kristen: Yea...I guess it's just because there are so many, especially when you don't have very many conditions, there are so many different graphs you could make out of it, and I think it's hard to keep straight on what's important and what's not.

During Kristen's second interview, which took place near the end of the semester after Quiz 5, she elaborated on the importance of the number of conditions specified on a given problem. In response to the question, "Which of the problems was more difficult?", she said the Type 2 problems with specified conditions were harder, but continued with:

Kristen: ...Unless there were a lot of conditions to where it was a lot more concrete. To where we knew what was happening at every point. Because sometimes you'd give us conditions where we didn't really know what was happening on a certain interval. So then, it makes you stop and think, Am I *supposed* to know what's happening here? That's what I think is harder about those...when you give like a list of six conditions and there's this one little interval where I don't know if it's increasing or decreasing or concave up or concave down, then I sit there and look at it for five minutes going, am I *supposed* to know this? And I look at all the other conditions and think, am I *supposed* to know this?

R: So when I give you lots of conditions, if every interval is specified, that makes [a Type 2 problem] easier?

Kristen: Right, right. Because sometimes there are two other conditions that will tell you what's happening there, and maybe that's not evident when you first look at it. But I think that's probably just me. I don't like math problems that don't have numbers in them! (laughing) I want numbers! I want an answer! So that might just be my personality. I think I'm more comfortable with boundaries.

Kristen's feeling appears to be that Type 2 problems with fewer conditions are more difficult than Type 2 problems with more conditions.

During Fred's first interview, which took place after Quiz 3, he was asked if he felt the problems on Quiz 3 were all equally difficult. He said Type 2 problems were harder, then confirmed Kristen's claim:

> Fred: There's [*sic*] just different variations and ways you could do things, so it's a little more confusing...The more conditions you have, then it gets easier. It's just when you have a lot of things that are left open, then it becomes more difficult.
>
> R: So the *more* conditions I give you, the easier it is –
>
> Fred: The easier it is, because then you know these points have to exist, and it's easier to determine what the graph is gonna look like.

Effect of Stakes

An informal conversation with Macy provided confirmation of the numerical results that suggested there was a large effect due to occasions and stakes. For Macy, the quizzes served as useful practice for the examinations. She did not study for the quizzes because they only accounted for 10 percent of the final grade in the course, and she simply used them to ascertain whether or not she knew how to solve the problems.

Over-interpreting Conditions

There was a quite unexpected misconception revealed during interviews with several of the students. These students were over-interpreting many of the conditions. They were assuming these conditions were much stronger than they actually were. During an interview with Steve, conducted after Quiz 3, when asked why he had included in his solution for problem 4 an inflection point at $x = 1$ (see Appendix F for his solution), he responded:

> Steve: So it has a local max at negative one, and from negative infinity to negative one, it's increasing, so it would have to start coming down from negative one to one, and if from one to three it's concave up, then from negative infinity to one would have to be concave down.

R: It would *have* to be?
Steve: Well, yeah, because it says it's concave up just from 1 to 3.
R: So even though I didn't tell you it was concave down on that entire interval from negative infinity to negative one, you're saying it would *have* to be?
Steve: Right.
R: Ok.
Steve: Because there's a local max right here, on negative one, and uh, unless there were some more — there was no specific instructions between negative one and one, so unless there was some more curves in there somewhere, or concave up and concave downs, the way I see it, that's the way it would have to be.

Steve read too much into the given condition that f is concave up on the interval from one to three, as some of the other students did. He believed the condition specifying the interval from one to three meant the *only* place where f could be concave up was on that interval. He also incorrectly concluded that the function had to be concave down on the whole interval from negative infinity to negative one. In reality, he could have drawn the graph any way he liked on that interval, provided he included a local maximum at negative one.

The interview with Ben revealed a similar misconception. Ben also was questioned regarding his reason for drawing an inflection point at $x=1$. He hesitantly admitted it "could have varied a little bit", but he had trouble accepting the fact that the function correctly could have been concave up on the interval from zero to three as well. When asked if it would have been correct to draw the function concave up somewhere between the domain values of zero and one with an additional minimum and maximum between zero and one (see Appendix H for his solution), he insisted,

> If I'm to match these criteria, and introduce no *new* criteria, no new maximums or minimums, ...that necessitates it being concave down there [from negative infinity up to one].... If we had other maximums

> and minimums in there that you cared to see, then you would have stated them as conditions of the problem!

Ben was guilty of making the same kind of incorrect assumption that Steve had made; he assumed that function behavior which was not explicitly *required* by a specified condition was actually *prohibited.*

Familiar Information Presented in an Unfamiliar Way

There was one other notable difficulty the students had with the Type 2 problems. When information was presented in a new way, students had trouble understanding its significance. Several students indicated they were completely at a loss how to interpret the condition "$f''(-1) = -9$" given on a Type 2 problem on Quiz 3 (see Appendix D, p. 132). The students had been given conditions involving the second derivative on homework problems that were phrased in terms of an interval, such as "$f'' < 0$ on $(-\infty, 0)$". Both of these conditions convey the fact that the function f is concave down on some interval including negative one, even though the second condition certainly conveys additional information. However, six of the ten students who were interviewed after Quiz 3 reported they indeed had difficulty with that condition. Their troubles ranged from the level of having to stop and consider it for a while, to leaving the problem and returning to it later, to completely missing that part of the problem. Natalie said that she simply did not recognize the significance of that condition because she had never seen one like it. As a result, she plotted the point $(-1,-9)$ as part of the graph of f. She also plotted the point $(-1,0)$ in response to the given condition "$f'(-1) = 0$". She commented, "I knew that they were different, but I just plotted them the same." She knew f, f', and f'' were different functions, but she was at a loss as to what she needed to draw that would meet those conditions involving f' and f''.

Two Basic Strategies for Type 2

There were two basic strategies used by the students to draw the graphs required by the Type 2 problems. One strategy students used was to draw part of the graph after they read each given condition and to erase and redraw when another condition called for some particular refinement or change in the graph. Tom's comments about his strategy for problem number 4 on Quiz 3 appear below. (The dashed line that appears in his solution, shown in Appendix G, was added in order to highlight his erasure marks.)

> Tom: I start with what I can put down that doesn't depend on anything else, like if f of 2 is zero, I can put that point down, it doesn't depend on anything else. That's what I call a "given"... You see that I drew two or three things and then had to erase it. I was trying to figure out [how condition f] depends on [conditions a through e], so I need to do this first, that sort of thing...I think they all kind of went together sort of. One thing always depended on something else.

The other strategy students used to solve Type 2 problems was to make notes along the top or bottom of the grid regarding intervals over which the function behaved a certain way and points at which certain function behavior occurred. After all conditions were considered and duly noted, the student would then draw the entire graph of the function in one stroke. Ben gave a description of this type of strategy during an interview following Quiz 3:

> R: When you saw, on problem number 1, the first condition, that the first derivative does not exist at four and two ...what did you do to try to meet that condition?
>
> Ben: Well, if it doesn't exist, then you've got some sort of sharp point there,

so I just, I made note of the fact, put a little mark there that, to help me remember that something odd happened there, and uh, you know, went down and looked at the rest of the conditions.

R: Ok, so you didn't draw anything immediately, you just made a note to yourself that there was something going on there?

Ben: Yeah.

R: Ok. When did you decide to draw what you drew?

Ben: After I'd already looked at all of them.

R: Ok.

Ben: What I do is, I – I hate these, [Kyle] and I agree on that, we hate these, because there's no definite way that it *has* to be, you just have to – the only way that I can approach something like that is just try to get in my head all at once what's supposed to be happening and make little notes on the graph about where special behavior is occurring at [*sic*], and then just – throw something in there. And I usually I go back and check that I satisfied everything... They always take longer, to me, because I actually have to try to get a mental image of it before I can draw it, and I guess, putting marks on the graph as I do helps a *little* bit with that, but you've got to go back and say, ok what was this mark, what was that one! (laughing) So that makes it more complex... You actually have to, well, you have to plan it out. You can't just, I guess you could scribble something down, and then have to erase it, but I prefer to take one shot at it and make it look as smooth and nice as I can.

Ben's solution to problem 4, the other Type 2 problem on Quiz 3 (see Appendix H) clearly shows several notes he made at the bottom of the graph.

Todd gave the following explanation of his strategy:

> If there's a lot of different ways to [draw the graph], it's real easy for me because I just go through step by step. I just mark a point if it's a min or a max, or if it's undefined, I always mark those. And then [after I draw the graph], I make sure [all the conditions are satisfied] afterwards.

Impact on Students' Understanding

Kristen made an unsolicited comment during her second interview regarding the impact of being required to work on these types of problems:

R: So [the target problems] were not your favorite things to do.

Kristen: No, but they helped. They helped me to understand concepts, because

I'm a visual learner. So some of the concepts that I didn't really maybe understand, they did help.

R: So would it have helped as much if I had *only* given you problems where I gave you graphs and asked you questions, or maybe if I had *only* given you problems where you had to draw a graph?

Kristen: (pause) No, I think it wouldn't have helped as much as you giving us both.

R: Why?

Kristen: Well... I think it's important to go at things from different directions because you have so many people in the class that learn different ways, and can understand one concept, but can't another. If you just explain a thing one way, you may have half the class that don't [*sic*] understand it. So, that's just my personal opinion about math teachers. I don't think math teachers should explain things one way. I think it's really helpful to do everything you can think of.

R: So you see [the two types of problems] as two different ways to actually help people—

Kristen: Understand the concept. Because some people might be backwards from me. [Type 2 problems] might be easier for them than the [Type 1 problems].

CHAPTER 5

Conclusions

The purpose of this study was to determine whether systematic differences exist between problems requiring students to interpret a given graph and problems requiring students to construct a graph to meet specified conditions.

Discussion

The main result from this study is a warning against assuming comparability in assessment items involving graphical representations of functions. The data examined here suggest that the assumption that all graphing problems assess the same sorts of skills and abilities is not warranted. The results in this study indicate that all types of graphing problems are not interchangeable; there are noticeable differences even among sub-domains of items that *appear* comparable.

Target problems for the quizzes and examinations were designed presuming that the Type 1 items were comparable and that the Type 2 items were comparable. This study sought to investigate any differences in the two types. In fact, differences due to the two different types of items were present, suggesting that successful completion of the two item types might require different abilities in the students. Consequently, it appears that problems involving graphical representations of functions may measure different kinds of skills and abilities. These results indicate there is a need for a closer examination of problems that require students to answer questions about a given graph (an *interpretation* task) and problems that require students to *construct* a graph, as well as other types of graphical representation problems. Consequently, common assessment methods as well as research practices may deserve reconsideration.

The variance decomposition from the generalizability analyses suggests that given the nature of the causes of the variance, a simple ANOVA examining differences in the mean scores for the two item types would likely be misleading. In fact, these results highlight the value of generalizability and variance decomposition analyses in assessment situations for which the nature of the variance is not well understood. Generalizability analyses may eliminate spurious conclusions based on a misleading comparison of means.

Qualitative data from interviews with students support the numerical results that suggested there are true systematic differences between these two types of problems. No matter how difficult it was to discover differences in item types in the quantitative analyses, it is obvious from a qualitative standpoint that there is a *clear* difference between Type 1 (interpretation) and Type 2 (construction) items in the minds of the students. Seventy-five percent of the students interviewed said that problems requiring them to construct a graph to meet specified conditions were more difficult than the problems requiring them to interpret a graph. Although several students, including the four lowest scoring students in the class, claimed that they found Type 2 problems easier than Type 1, there was no difference in their performance on these problems.

Interestingly, several students said they liked the Type 2 problems better, even though they thought these were more difficult than the Type 1 problems. In their interviews, students described the Type 2 problems as challenging, fun, and enjoyable, while others likened them to a puzzle or a game. On the other hand, some students absolutely hated the Type 2 problems, describing them as much more complex than the Type 1 problems. Their complaints included difficulty in processing all of the information presented in the specified conditions in order to produce a valid graph, frustration with the need to deal with interactions among the conditions, and a feeling of discomfort when confronted with problems having more than one correct solution. It is interesting to note that the characteristic of having infinitely many correct solutions was one of the most appealing aspects of the Type 2 problems for the students who liked them.

Students made distinctions in level of difficulty even within the two types of problems. This corroborates the unexpected small differences that were evident in the quantitative results among the Type 1 items themselves, particularly among those on the quizzes. According to ten of the students interviewed, conditions and questions involving information about the derivatives or antiderivatives of the function were more difficult to deal with than those conditions and questions involving direct information about the function itself. Distinctions have been made in the research literature between tasks involving local or quantitative interpretations and tasks involving global or qualitative interpretations (Leinhardt et al., 1990). Direct information or questions regarding the function itself are often local or quantitative in nature, whereas information involving derivatives or antiderivatives of the function are certainly global or qualitative in nature.

Additionally, dealing with information about derivatives was not as difficult for the students in the current study as dealing with information about antiderivatives. Similar findings have been reported in published literature. Several studies have found that students in various age groups have more difficulty producing an algebraic formula from a given graph, than producing a graph from a given formula (Carpenter, Corbit, Kepner, Lindquist, and Reys, 1981; Markovits et al., 1986; Wagner, Rachlin, and Jensen, 1984). It should not be assumed that students will automatically be able to *reverse* any of the processes they have mastered.

Several students in the current study agreed that the Type 2 problems with *more* specified conditions were easier than those with *fewer*. In addition, many agreed that any question in a Type 1 problem or condition in a Type 2 problem that was related to the first or second derivative was more difficult than a question or condition related directly to the function itself.

Difficulty with cusp points, vertical tangents, conditions involving the second derivative, and coordinating the information in the specified conditions in Type 2 problems has also been documented by Baker, Cooley, and Trigueros (in press).

The quantitative results presented in this study confirmed the suspicion that students might prepare differently for quizzes than for examinations if the scores were weighted differently in the calculation of final grades. Many of the analyses produced results containing a large amount of variation due to the effect of different occasions or stakes (quizzes vs. examinations). The results were further confirmed by a conversation with one student who said she indeed did not study for the quizzes because they accounted for such a small percentage of her final grade. Additionally, the numerical results showed that with minimal training of raters, effects due to raters are absent.

The interview process revealed unexpected misconceptions. For example, an error some students made was to over-interpret the conditions. Some students assumed that function behavior which was not explicitly *required* by a specified condition was actually *prohibited*.

One notable difficulty examined was the trouble students had with information that was presented in a new way. Many students experienced difficulty in determining significance of familiar conditions or in determining answers to familiar questions when they were presented in an unfamiliar way.

The interviews revealed two basic strategies for solving Type 2 problems. One strategy is to draw part of the graph after reading each given condition, and to erase and redraw when another condition calls for some particular refinement or change in the graph. The other strategy is to make notes along the top or bottom of the page concerning intervals over which the function behaves a certain way and points at which certain function behavior occurs. After all conditions are considered and duly noted, the student draws the entire graph of the function in one stroke.

Limitations of the Study

The sample available for this study was too small for the complicated model originally specified. Better results might have been obtained with the original model if a larger sample had been available. Certainly, given a large enough sample, a single analysis could have been run that incorporated all the facets of interest in this study.

More attention could have been given to the distinctions in different types of graphing problems if time in class had allowed more assessments. Better results might have been obtained if further distinctions could have been made between graph-to-graph Type 2 problems like the ones given on Quiz 5 in which one graph is given and another must be constructed, and the Type 2 problems in which explicit conditions are specified. For the Type 2 problems on Quiz 5, the students had to construct the specified conditions for themselves based on the graph that was given and then sketch a graph that met those conditions. For example, one problem gave the graph of a function f and asked the students to sketch a graph of the derivative of f. More distinction in item types could have produced more of an effect due to differences in types of items.

More distinction in categories of conditions might have yielded more powerful results as well. During the interviews, many students said that conditions and questions related to the first or second derivative or the antiderivative were more difficult to deal with than conditions and questions concerning the function itself. Better results might have been obtained if conditions from only *one* category had been given on each problem. For instance, if the given conditions on one problem involve only the function itself, and the given conditions on a different problem involve only the first and second derivatives, it might possible to show that *these* two problems measure different abilities.

More care could have been taken with respect to problems on which students received points by default. For example, for a Type 2 problem on Quiz 1 (see Appendix D, p. 104) there was no condition requiring the function be defined for all real numbers. As a result, there were many instances in which the student drew nothing at all on a particular interval, but received credit for meeting some of the conditions by default. Information obtained during interviews confirmed that some of these students were indeed thinking incorrectly as they drew the graph for this problem. These default points artificially inflated many of the scores on Type 2 problems, and could have served to partially mask the sought-after type differences.

On Quiz 4, there was a Type 1 problem in which the graph of f' was given, rather than the graph of f (see Appendix D, p. 111). This was the first time such a problem appeared on a quiz or examination, and many students obviously answered the questions based on the graph being the graph of f. Six of the thirteen students who were interviewed after Quiz 4 confirmed this fact. This error caused them to miss all the questions on that problem, which served to artificially deflate the scores on the Type 1 problems. This could have served to partially mask the sought-after type differences as well.

Additional contact with the instructor could have caused a slight effect. Twenty-four of the 32 students were interviewed at least once during the course of the semester. If all students had been interviewed exactly once, the case could be made that this would have served to reduce any effect caused by additional contact with the instructor. However, five students were followed as case studies, and were interviewed after every quiz and examination containing target problems. For those five students, there could have been a slight effect caused by the additional contact. However, it seems reasonable to assume the effect would have been somewhat constant across types of problems. Since the results were based on a comparison of performance on two different types of problems, additional contact with the instructor should be less of a threat to validity.

In a study such as this, there is perhaps the issue of the students' gaining experience and proficiency over time with the target problems as a threat to validity. Since the study is comparing the two types of problems, it seems a reasonable assumption that similar kinds of experience are gained with *both* types of target problems, which would reduce the threat to the validity of the results.

Mutual learning could also be a threat to the validity of the results of this study. It is possible that working one type of target problem could affect the amount learning that takes place regarding the other type of problem. This seems unlikely since the students worked on both types of problems throughout the semester. Even if mutual learning did occur, it was not enough to completely mask the differences in the two types of problems. There was still evidence of an effect due to the different types of problems.

Directions for Future Research

The results of this study suggest that there are systematic differences between the two types of problems examined. The natural questions at this point are: Can these results be replicated and strengthened? Are the differences in problem types due to ability or achievement of the student? Do the differences correlate with other differences in student performance? Can calculus instruction be modified to improve student performance on graphing problems? Would improvement in performance on graphing problems result in better understanding of functions and graphs in general?

Answering the questions regarding differences due to ability or achievement of the student and the correlation of ability with other differences in student performance might involve an analysis of two subgroups of students based on their scores on each type of problem. If a comparison of means for the two groups yielded a significant result, further analyses could be conducted involving such factors as the students' final course grades, their GPA's, their ACT scores, or other background factors. These analyses could reveal a correlation between achievement on the different problem types and some indicator of ability or background factor.

A qualitative study could be undertaken to examine in more detail the strategies and methods used for solving Type 2 problems. Such a study might serve to answer the question: Do students solution methods *evolve* as they gain more experience?

Additional questions that arise are related to other types of problems involving graphical representations. For instance, are there systematic differences between problems requiring students to match a graph to its algebraic formula and problems requiring students to produce the algebraic formula to match a given graph on their own? It seems reasonable to suspect such differences might exist in light of the results of the current study.

REFERENCES

Ahmadi, D. C. (1995). *A comparison study between a traditional and reform calculus II program.* Doctoral dissertation, University of Oklahoma.

Asiala, M., Brown, A., DeVries, D. J., Dubinsky, E., Mathews, D., Thomas, K. (1996). A framework for research and curriculum development in undergraduate mathematics education. *Research in Collegiate Mathematics Education II, 3,* pp. 1-32.

Baker, B., Cooley, L., and Trigueros, M. (in press). A calculus graphing schema. *Journal for Research in Mathematics Education.*

Barclay, T. (1987). A graph is worth how many words? *Classroom Computer Learning, 7* (5), 46-50.

Bell, A., and Janvier, C. (1981). The interpretation of graphs representing situations. *For the Learning of Mathematics, 2*(1), 34-42.

Borba, M. C., and Confrey, J. (1996). A student's construction of transformations of functions in a multiple representational environment. *Educational Studies in Mathematics, 31* (3), 319-337.

Braddy, L., and McKnight, C. (1997, August). *Generating graphs to meet required conditions: A generalizability study.* Paper presented at Mathfest 97, the annual summer meeting of the Mathematical Association of America, Atlanta, GA.

Brasell, H. M., and Rowe, M. B. (1993). Graphing skills among high school physics students. *School Science and Mathematics, 93* (2), 62-70.

Brennan, R. L. (1992). *Elements of generalizability theory* (Rev. ed.). Iowa City, IA: American College Testing Program.

Brennan, R. L. (1997). A perspective on the history of generalizability theory. *Educational Measurement: Issues and Practice, 16*(4), 14-20.

Caldwell, F. W. (1995). *Effect of graphics calculators on college students' learning of mathematical functions and graphs.* Paper presented at the annual conference of the American mathematical Association of Two-Year Colleges, Tulsa, OK.

Carpenter, T. P., Corbit, M. K., Kepner, H. S., Jr., Lindquist, M. M., and Reys, R. E. (1981). *Results from the second mathematics assessment of the National Assessment of Educational Progress.* Reston, VA: National Council of Teachers of Mathematics.

Clement, J., Mokros, J. R., and Schultz, K. (1985). *Adolescents' graphing skills: A descriptive analysis.* Paper accepted for presentation at the 1986 annual meeting of the American Educational Research Association.

Crick, J. E., and Brennan, R. L. (1983). *GENOVA: A generalized analysis of variance system* (Fortran IV computer program and manual.) Iowa City, IA: American College Testing Program.

Cronbach, L. J., Gleser, G. C., Nanda, H., and Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles.* New York: John Wiley.

Davis, R. B. (1984). *Learning mathematics: The cognitive science approach to mathematics education.* Norwood, NJ: Ablex Publishing Corporation.

Demana, F., Schoen, H. L., Waits, B. (1993). Graphing in the K-12 curriculum: The impact of the graphing calculator. In T. A. Romberg, E. Fennema, and T. P. Carpenter (Eds.), *Integrating Research on the graphical representation of functions* (pp. 11-39). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Dienes, Z. P. (1960). *Building up mathematics.* London: Hutchinson Educational.

Douglas, R. G. (1986). *Toward a lean and lively calculus.* Stony Brook, NY: The Mathematical Association of America.

Dugdale, S. (1993). Functions and graphs – Perspectives on student thinking. In T. A. Romberg, E. Fennema, and T. P. Carpenter (Eds.), *Integrating Research on the graphical representation of functions* (pp. 101-130). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Eason, S. (1991). Why generalizability theory yields better results than classical test theory: A primer with concrete examples. In B. Thompson (Ed.), *Advances in educational research: Substantive findings, methodological developments* (vol. 1, pp. 83-98). Greenwich, CT: JAI Press.

Eisenberg, T. and Dreyfus, T. (1994). On understanding how students learn to visualize function transformations. *CBMS Issues in Mathematics Education, 4,* 45-68.

Ferrini-Mundy, J. and Graham, K. (1994). Research in calculus learning: Understanding of limits, derivatives, and integrals. In J. J. Kaput and E. Dubinsky (Eds.), *Research Issues in Undergraduate Mathematics Learning: Preliminary Analyses and Results.* MAA Notes No. 23. Washington DC: The Mathematical Association of America.

Fine, A. E. (1995). *An investigation in the use of performance assessment tasks in undergraduate mathematics.* Doctoral dissertation, University of Oklahoma.

Finney, R. L., Thomas, G. B., Demana, F., Waits, B. K. (1994). *Calculus: Graphical, Numerical, Algebraic.* Reading, MA: Addison-Wesley Publishing Company.

Garofalo, J., and Durant, K. (1991). Applied functions and graphs: A necessary topic for developmental mathematics. *Research and Teaching in Developmental Education, 8* (1), 51-55.

Goldenberg, E. P. (1988). Mathematics, metaphors, and human factors: Mathematical, technical, and pedagogical challenges in the educational use of graphical representation of functions. *Journal of Mathematical Behavior, 7,* 135-173.

Heid, M. K. (1988). Resequencing skills and concepts in applied calculus using the computer as a tool. *Journal for Research in Mathematics Education, 19,* 3-25.

Janvier, C. (1978). *The interpretation of complex Cartesian graphs: Studies and teaching experiments.* Unpublished doctoral dissertation, University of Nottingham, UK.

Kaput, J. (1993). The urgent need for proleptic research in the representation of quantitative relationships. In T. A. Romberg, E. Fennema, and T. P. Carpenter (Eds.), *Integrating Research on the graphical representation of functions* (pp. 279-312). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Karplus, R. (1979, October-December). Continuous functions: Students' viewpoints. *European Journal of Science Education, 1,* 397-415.

Karplus, R., Lawson, A. E., Wollman, W. T., Appel, M., Bernoff, R., Howe, A., Rusch, J. J., and Sullivan, F. (1977). *Science teaching and the development of reasoning.* Berkeley, CA: University of California.

Kieran, C. (1993). Functions, graphing, and technology: Integrating research on learning and instruction. In T. A. Romberg, E. Fennema, and T. P. Carpenter (Eds.), *Integrating Research on the graphical representation of functions* (pp. 189-237). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Krabbendam, H. (1982). The non-quantitative way of describing relations and the role of graphs: Some experiments. In G. van Barneveld and H. Krabbendam (Eds.), *Proceedings of the Conference on Functions* (pp. 125-146). Enschede, The Netherlands: National Institute for Curriculum Development.

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

Leitzel, J. R. C. (1991). *A call for change: Recommendations for the mathematical preparation of teachers of mathematics.* Washington, DC: The Mathematical Association of America.

Markovits, Z., Eylon, B., and Bruckheimer, M. (1986). Functions today and yesterday. *For the Learning of Mathematics, 6* (2), 18-25.

Mevarech Z. R., and Kramarsky, B. (1997). From verbal descriptions to graphic representations: Stability and change in students' alternative conceptions. *Educational Studies in Mathematics, 32,* 229-263.

Monk, G. S. (1992). Students' understanding of a function given by a physical model. In G. Harel and E. Dubinsky (Eds.), *The concept of function: Aspects of epistemology and pedagogy.* MAA Notes and Reports, No. 25. Washington, DC: Mathematical Association of America, 175-194.

Moschkovich, J., Schoenfeld, A. H., and Arcavi, A. (1993). Aspects of understanding: On multiple perspectives and representations of linear relations and connections among them. In T. A. Romberg, E. Fennema, and T. P. Carpenter (Eds.), *Integrating Research on the graphical representation of functions* (pp. 69-100). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

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

Norman, F. A. (1993). Integrating research on teachers' knowledge of functions and their graphs. In T. A. Romberg, E. Fennema, and T. P. Carpenter (Eds.), *Integrating Research on the graphical representation of functions* (pp. 159-187). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Orton, A. (1983). Students' understanding of differentiation. *Educational Studies in Mathematics, 15,* 235-250.

Patterson, D. L. (1983). Calculus students' utilization of graphs as visual representations. *Dissertation Abstracts.* University Microfilms, Inc.

Peters, P. C. (1982). Even honors students have conceptual difficulties with physics. *American Journal of Physics, 50*(6).

Philipp, R. A., Martin, W. O., and Richgels, G. W. (1993). Curricular implications of graphical representations of functions. In T. A. Romberg, E. Fennema, and T. P. Carpenter (Eds.), *Integrating Research on the graphical representation of functions* (pp. 239-278). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Phillips, R. J. (1986). Micro primer maths – There's more than meets the eye. *Mathematics in Schools, 15*(2), 29-32.

Piez, C. M., and Voxman, M. H. (1997). Multiple representations – Using different perspectives to form a clearer picture. *The Mathematics Teacher, 90* (2), 164-166.

Porzio, D. T. (1995). *Effects of differing technological approaches on students' use of numerical, graphical and symbolic representations and their understanding of calculus.* Paper presented at the annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education, Columbus, OH.

Preece, J. (1983). Graphs are not straightforward. In T. R. G. Green and S. J. Payne (Eds.), *The Psychology of Computer Use* (pp. 41-56). London: Academic Press.

Research Advisory Committee of the National Council of Teachers of Mathematics (1990). Mathematics education reform and mathematics education research: Opportunities, obstacles, and obligations. *Journal for Research in Mathematics Education, 21*(4), 287-292.

Roberts, A. W. (1996). *Calculus: The dynamics of change.* Washington, DC: The Mathematical Association of America.

Romberg, T. A., Fennema, E., and Carpenter, T. P. (1993). *Integrating research on the graphical representation of functions.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Ruthven, K. (1990). The influence of graphic calculator use on translation from graphic to symbolic forms. *Educational studies in mathematics, 21*, 431-450.

Shavelson, R. J., and Webb, N. M. (1991). *Generalizability theory.* Newbury Park, CA: Sage Publications.

Shavelson, R. J., Webb, N. M., and Rowley, G. L. (1989). Generalizability theory. *American Psychologist, 44*, 922-932.

Shaw, E. L., Padilla, M. J., and McKenzie, D. L. (1983). *An examination of the graphing abilities of students in grades seven through twelve.* Paper presented at the annual meeting of the National Association of Research in Science teaching, Dallas.

Schwarz, B. and Dreyfus, T. (1993). Measuring integration of information in multirepresentational software. *Interactive Learning Environments, 3*(3), 177-198.

Shoaf-Grubbs, M. M. (1994). The effect of the graphing calculator on female students' spatial visualization skills and level of understanding in elementary graphing and algebra concepts. *CBMS Issues in Mathematics Education, 4,* 169-194.

Tall, D. (1996). Functions and calculus. In A. J. Bishop et al. (Eds.), *International Handbook of Mathematics Education,* (pp. 289-325). The Netherlands: Kluwer Academic Publishers.

Thompson, B. and Crowley, S. L. (1994, April). *When classical measurement theory is insufficient and generalizability theory is essential.* Paper presented at the annual meeting of the Western Psychological Association, Kailua-Kona, Hawaii.

Thompson, P. W. (1994). Students, functions and the undergraduate curriculum. *CBMS Issues in Mathematics Education, 4,* 21-44.

Tucker, A. C. and Leitzel, J. R. (1995). *Assessing calculus reform efforts: A report to the community.* Washington, DC: The Mathematical Association of America.

Vinner, S. and Dreyfus, T. (1989). Images and definitions for the concept of function. *Journal for Research in Mathematics Education, 20*(4), 356-366.

Williams, S. R. (1993). Some common themes and uncommon directions. In T. A. Romberg, E. Fennema, and T. P. Carpenter (Eds.), *Integrating Research on the graphical representation of functions* (pp. 313-337). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Yerushalmy, M. (1991). Student perceptions of aspects of algebraic function using multiple representation software. *Journal of Computer Assisted Learning, 7* (1), 42-57.

Yerushalmy, M. and Schwartz, J. L. (1993). Seizing the opportunity to make algebra mathematically and pedagogically interesting. In T. A. Romberg, E. Fennema, and T. P. Carpenter (Eds.), *Integrating Research on the graphical representation of functions* (pp. 41-68). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

APPENDIX A

Human Subjects Permission

The University of Oklahoma

OFFICE OF RESEARCH ADMINISTRATION

August 5, 1998

Linda Braddy
Route 3, Box 263C
Ada, Oklahoma 74820

Dear Ms. Braddy:

Your research proposal, "Generating Graphs to Meet Required Conditions: A Generalizability Study," has been reviewed by Dr. E. Laurette Taylor, Chair of the Institutional Review Board, and found to be exempt from the requirements for full board review and approval under the regulations of the University of Oklahoma-Norman Campus Policies and Procedures for the Protection of Human Subjects in Research Activities.

Should you wish to deviate from the described protocol, you must notify me and obtain prior approval from the Board for the changes. If the research is to extend beyond 12 months, you must contact this office, in writing, noting any changes or revisions in the protocol and/or informed consent form, and request an extension of this ruling.

If you have any questions, please contact me.

Sincerely yours,

Karen M. Petry

Karen M. Petry
Administrative Officer
Institutional Review Board

KMP:pw
FY99-19

cc: Dr. E. Laurette Taylor, Chair, IRB
Dr. Curtis McKnight, Mathematics

EAST CENTRAL UNIVERSITY

Duane C. Anderson
Vice President for Academic Affairs

ADA, OKLAHOMA 74820-6899

INTER OFFICE

MEMO

To: Linda Braddy, Assistant Professor
From: Duane C. Anderson DCA
Subject: Human Subjects
Date: September 2, 1998

The Human Subjects Committee has submitted to me its review of your proposal entitled "Generating Graphs to Meet Required Conditions: A Generalizability Study." The committee expressed concern about the tagging of the raw data with numerical codes which will be attached to the subjects' names (see item 3.4 of the proposal and items 1 and 2 of the questionnaire). Since questions 11 and 12 on the Background Questionnaire elicit personal academic information, there is some concern about confidentiality.

Unless there is an essential reason for tagging the raw data with numerical codes that identify the subjects' names, I recommend deletion of the tag. With this change I will approve the proposal.

If you need to visit with me, or the committee, about this issue, please make an appointment with myself, or speak with Dr. Anita Walker, Chair of the Professional Development and Research Committee.

cc: Anita Walker, Chair

APPENDIX B

Informed Consent Form

Informed Consent Form for Participation in a Research Project
Being Conducted Under the Auspices of the
University of Oklahoma – Norman Campus

Study Title: Generating Graphs to Meet Required Conditions: A Generalizability Study
Sponsor: Dr. Curtis McKnight, Mathematics Department, University of Oklahoma
Principal Investigator: Linda Braddy

Description of My Research Study

The purpose of this research is to gain a better understanding of what types of questions may be appropriate for assessing Calculus I students' understanding of functions and their graphical representations. The data for this research will come from problems on quizzes and exams you will be taking as part of the normal activities of this course. Following each quiz or exam that includes a problem I select for my research, I would like to conduct personal interviews with a few of you outside of class. These interviews will be scheduled at your convenience. Each interview will last approximately 20 minutes and will be recorded on audio cassette tape. During the interview, I will ask you about the kinds of things you were thinking as you worked through the selected problem.

Your Potential Risks and Benefits from Participating

You should have little risk or discomfort from taking part in this study. The only additional time commitment will be the personal interviews and it is up to you whether or not you take part in them. Agreeing to let me interview you could cost you some study time (an hour or two altogether for the entire semester) for this or other classes, which could negatively affect your course grade(s). If, at any point, you feel that the time required to participate in this study is having a negative affect on your course work, you are free to discontinue your participation.

On the other hand, your participation in the personal interviews may potentially result in you acquiring a better understanding of graphical representations of functions than you would have gained otherwise, since you will be spending additional time thinking about the selected problems. The knowledge that we gain from this study may potentially result in better instruction for other students in future semesters.

My Guarantees to You

Your participation is voluntary. You won't be paid. You may choose from the beginning not to participate. You may decline further participation at any time. No penalty will be associated with such action. Consenting to participate will not add any work to your load. *Not* consenting to participate will not *exempt* you from any of the quizzes or exams. Your grade in this course will be determined only by the numerical guidelines listed in the course syllabus whether or not you participate.

Your identity will be kept confidential, and you will never be identifiable in any publication or presentation associated with this research.

If you have questions either about the research itself or about your rights as a research subject, you may contact Linda Braddy by phone at (580) 332-8329 or (580)332-8000, ext. 283 or by e-mail at telsbrad@chickasaw.com.

By signing below, you indicate your willingness to participate in this study as described above.

Signature ____ Date:________________

APPENDIX C

Background Questionnaire

Background Questionnaire
Calculus I - Fall 1998

Please complete the following short questionnaire. This will help your instructor have some background information on the students in the class. This information will not be used to determine any part of your grade in the course.

1. Name ______________________________

2. ID Number ___________________________

3. Gender: ____ Male ____ Female

4. Classification: ____ Freshman ____ Sophomore ____Junior ____Senior

5. What is your major? ____________________

6. Why are you taking this course?____
____

7. How many hours per week are you employed?____________

8. When did you graduate from high school? 19____

9. How many students were in your high school graduating class? ______ students

10. Which of the following mathematics classes did you have **in high school** and when?

____ Algebra I in 19____
____ Geometry in 19____
____ Algebra II in 19____
____ Trigonometry in 19____
____ Advanced Math in 19____
____ Precalculus in 19____
____ Calculus in 19____
____ Other (________________) in 19____
____ Other (________________) in 19____

11. Which of the following mathematics classes did you have **in college**, where, when, and what was your course grade?

____ Beginning Algebra at ________________ in the __________ semester of 19____
Course Grade: A B C D F W

____ Intermediate Algebra at ______________ in the __________ semester of 19____
Course Grade: A B C D F W

____ College Algebra at __________________ in the __________ semester of 19____
Course Grade: A B C D F W

____ Survey of Math at ___________________ in the __________ semester of 19____
Course Grade: A B C D F W

____ Trigonometry at ____________________ in the __________ semester of 19____
Course Grade: A B C D F W

____ Calculus I at _______________________ in the __________ semester of 19____
Course Grade: A B C D F W

____ Other (______________) at ___________ in the __________ semester of 19____
Course Grade: A B C D F W

____ Other (______________) at ___________ in the __________ semester of 19____
Course Grade: A B C D F W

12. Which graphing calculator(s), if any, do you already know how to use? __________
____

13. Which of the following best describes your experience with graphing calculators?

____ None
____ I have seen one or two demonstrations
____ I have seen more than two demonstrations
____ I have used graphing calculators once or twice
____ I have used graphing calculators frequently in (how many?) ____ of my classes
____ Other ____
____

14. How helpful do you expect the graphing calculator to be to your understanding and learning of calculus in this class?

____ Very unhelpful
____ Somewhat unhelpful
____ Undecided
____ Somewhat helpful
____ Very helpful

15. What is your general feeling about calculators and computers?
____ I like them a lot
____ They're OK
____ I can take them or leave them
____ I don't like them
____ I really hate them

16. Compared to other students in my high school and college math courses, the amount of time I spend studying is ____more than they spend
____about the same as they spend
____less than they spend

17. I expect to spend ____ (how many?) hours per week studying calculus this semester.

18. How much time did you spend last week doing calculus homework?
____

19. How much time did you spend last week studying calculus (including studying the book, learning to use your calculator, and all other out-of-class time you spent)?____________

20. How do you feel about math?
____I like it a lot
____I like it
____I don't know how I feel about it
____I don't like it
____I hate it

21. I find math . . . .
____very easy
____somewhat easy
____about the same difficulty level as other subjects
____somewhat difficult
____very difficult

22. How do you think of yourself at a mathematics student?
____very weak
____ somewhat weak
____ I'm undecided
____ somewhat strong
____ very strong

23. How do you feel about studying calculus as you start this course?

_____ very shaky
_____ a little afraid
_____ I don't know how I feel about it
_____ confident
_____ very confident

24. What grade do you expect to get (I mean, **really** expect to get)?

____A ____B ____C ____D ____F

25. Circle the appropriate number on the 5 point scale for the following statements:

(NOTE: The phrase "*studying math*" refers to activities such as reading the textbook, studying your notes, studying for tests, doing homework, and anything else you do to learn the material and prepare for class.)

	Never	Rarely	Sometimes	Often	Always
I do my math homework.	1	2	3	4	5
In math, I'm a visual thinker.	1	2	3	4	5
I attend my math classes regularly.	1	2	3	4	5
I work hard at studying math.	1	2	3	4	5
I find math courses challenging.	1	2	3	4	5
I draw graphs when studying math.	1	2	3	4	5
I find math courses boring.	1	2	3	4	5
In general, I'm a visual thinker.	1	2	3	4	5
I find math courses interesting.	1	2	3	4	5
I find math courses intimidating.	1	2	3	4	5
I can figure things out by looking at a graph.	1	2	3	4	5

26. Circle the appropriate number on the 5 point scale for the following statements:

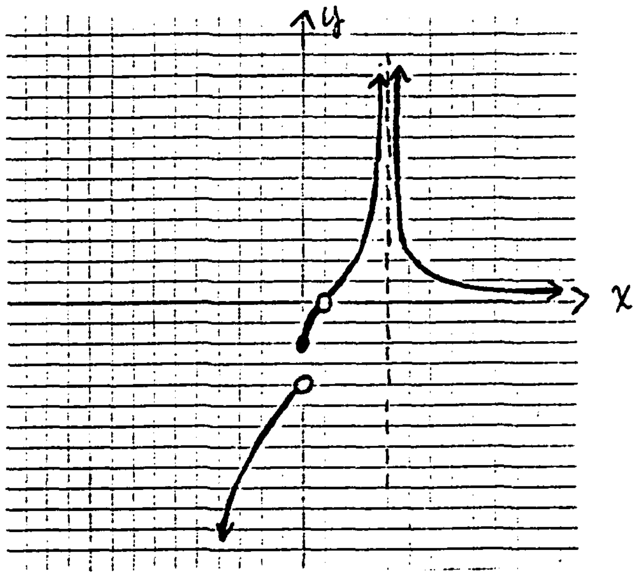
	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
Math courses take more time than other courses.	1	2	3	4	5
It is important for me to learn to use graphs in order to get a good grade in math.	1	2	3	4	5
Math courses are my favorite courses.	1	2	3	4	5
I learn math best by following examples.	1	2	3	4	5
I am good at math.	1	2	3	4	5
The ability to understand graphs is an important skill for studying math.	1	2	3	4	5
I think most math courses are a waste of my time.	1	2	3	4	5
I learn math best by memorizing and using rules and definitions.	1	2	3	4	5
I learn anything better when I can see a picture of what I am studying.	1	2	3	4	5
I find it difficult to draw graphs.	1	2	3	4	5
Math is an important subject.	1	2	3	4	5
It is helpful for me to draw graphs, even when I'm not told to, when studying math.	1	2	3	4	5
I think math homework is a waste of my time.	1	2	3	4	5
I enjoy math.	1	2	3	4	5

APPENDIX D

Target Problems

APPENDIX D
Quiz 1
Type 1 problem

1.



(a) $\lim_{x \to \infty} f(x) =$ ____________

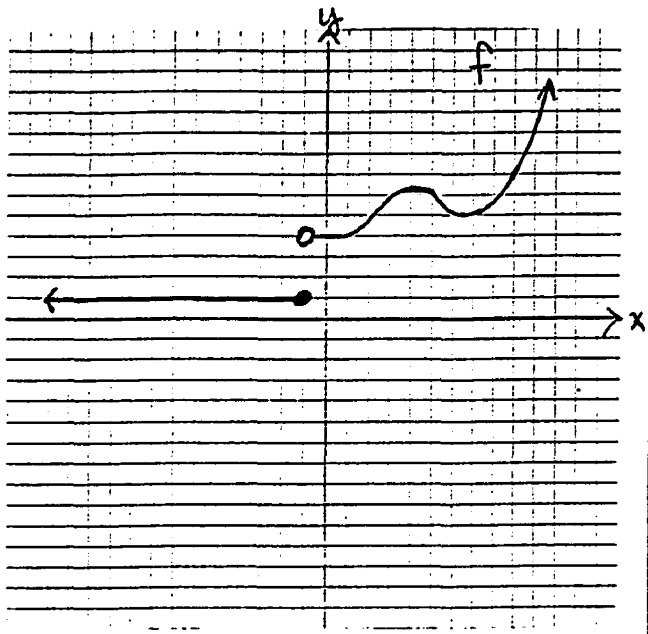
(b) $f(1) =$ ________________

(c) $\lim_{x \to 0^-} f(x) =$ ____________

(d) At what values of x, if any, is f discontinuous? ______________________________

APPENDIX D
Quiz 1
Type 1 problem

3.



(a) $f(-1) =$ ____________

(b) For what values of x, if any, is f continuous? ________________________________

(c) $\lim_{x \to -\infty} f(x) =$ ____________

(d) $\lim_{x \to -1} f(x) =$ ____________

APPENDIX D

Quiz 1

Type 2 problem

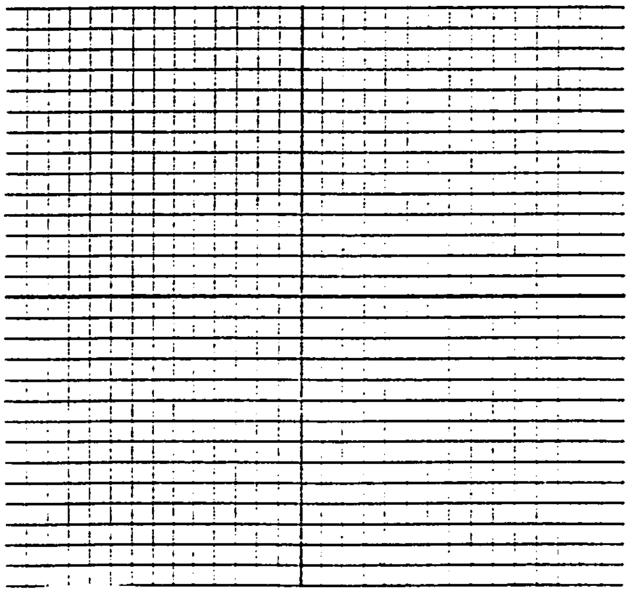
2. Sketch the graph of a function that satisfies *all* of the following conditions:

(a) $\lim_{x \to \infty} f(x) = \infty$

(b) f has a vertical asymptote at $x = 2$

(c) $\lim_{x \to -2} f(x)$ DNE

(d) $f(3) = -2$



APPENDIX D

Quiz 1

Type 2 problem

4. **Sketch the graph of a function that satisfies *all* of the following conditions:**

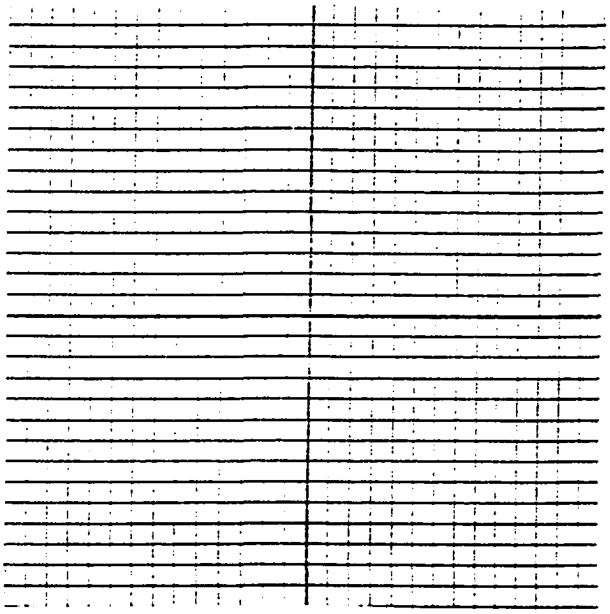
(a) f is defined for all real numbers except $x=-2$

(b) f is continuous everywhere ***except*** at $x=-4$, $x=-2$, $x=1$, and $x=3$.

(c) $\lim_{x\to\infty} f(x)=2$

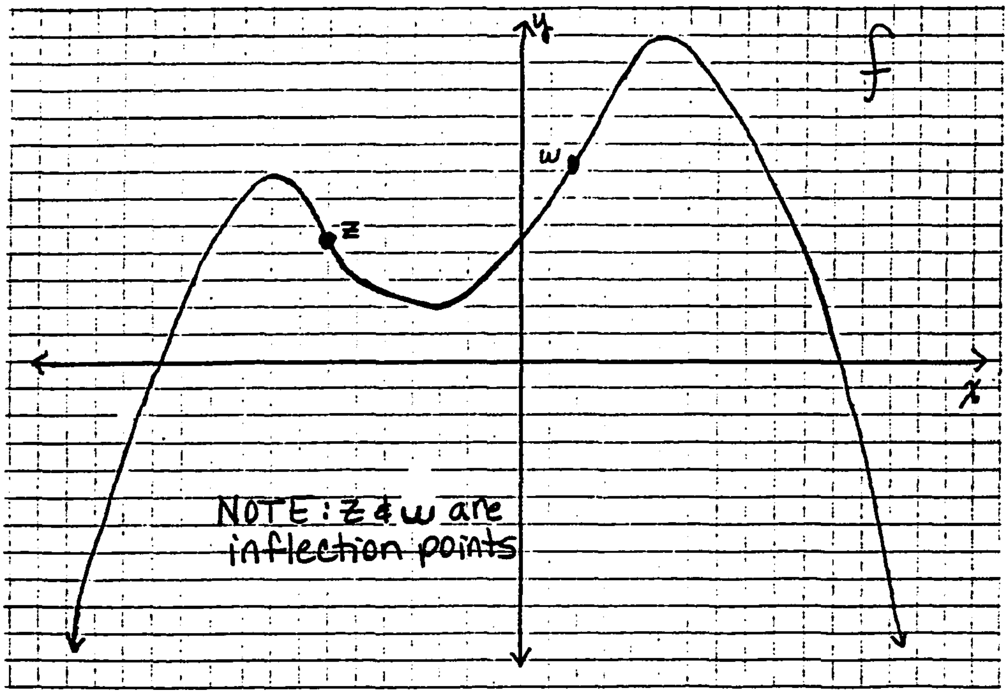
(d) $f(-2)$ is undefined

(e) $\lim_{x\to 5^-} f(x)=1$



APPENDIX D
Quiz 3
Type 1 problem

2.



(a) On what interval(s), if any, is f increasing?________________

(b) On $(-\infty,\infty)$, at what value(s), if any, does f have a local max?

(c) On what interval(s), if any, is f concave up?

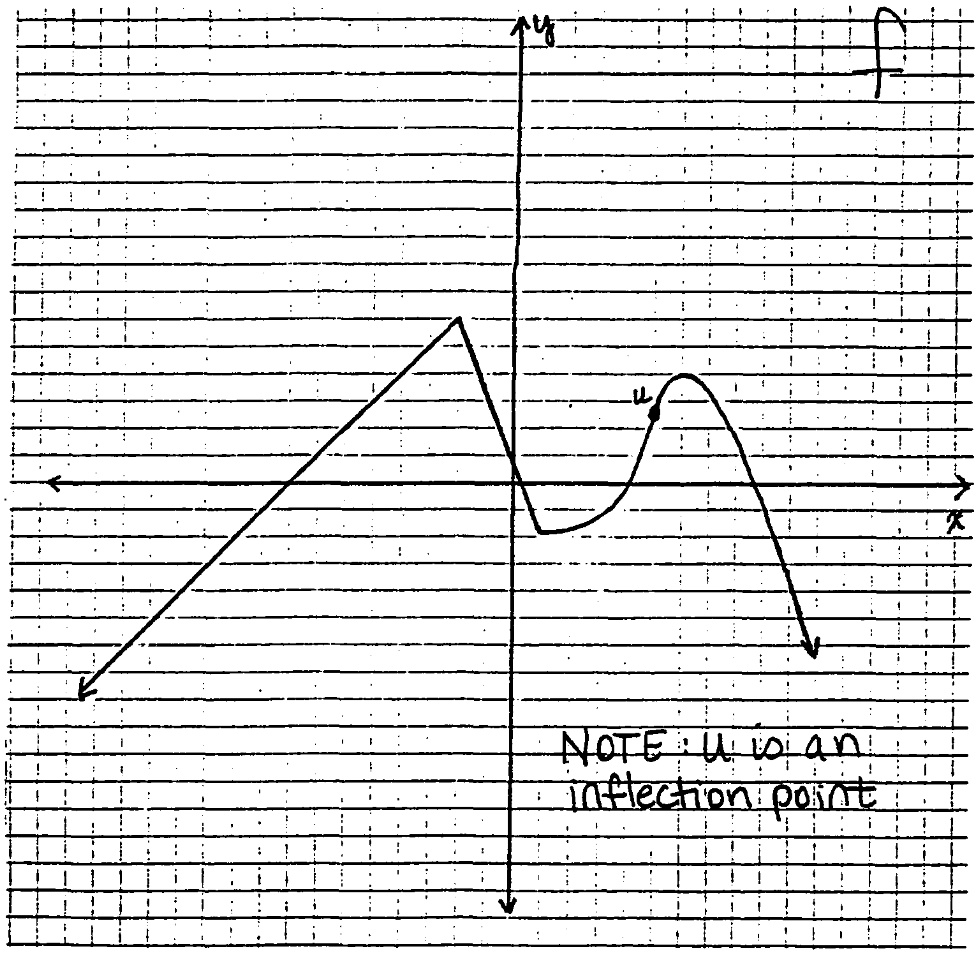
(d) On what interval(s), if any, is f decreasing?

(e) On what interval(s), if any, is f concave down?

(f) On $(-\infty,\infty)$, at what value(s), if any, does f have a local min?

APPENDIX D
Quiz 3
Type 1 problem

3.



(a) $f'(-2) =$ _______

(b) At what value(s), if any, does f have a local max?

____

(c) At what value(s), if any, does f have a local min?

____

(d) $f(4) =$ _______

(e) $f'(6) =$ _______

(f) Is $f'(5)$ negative or positive? ____________________

APPENDIX D
Quiz 3
Type 2 problem

1. Sketch the graph of a function f that satisfies all of the following conditions:

(a) f' DNE at $x = 4$ and $x = 2$

(b) $f(-1) = -2$

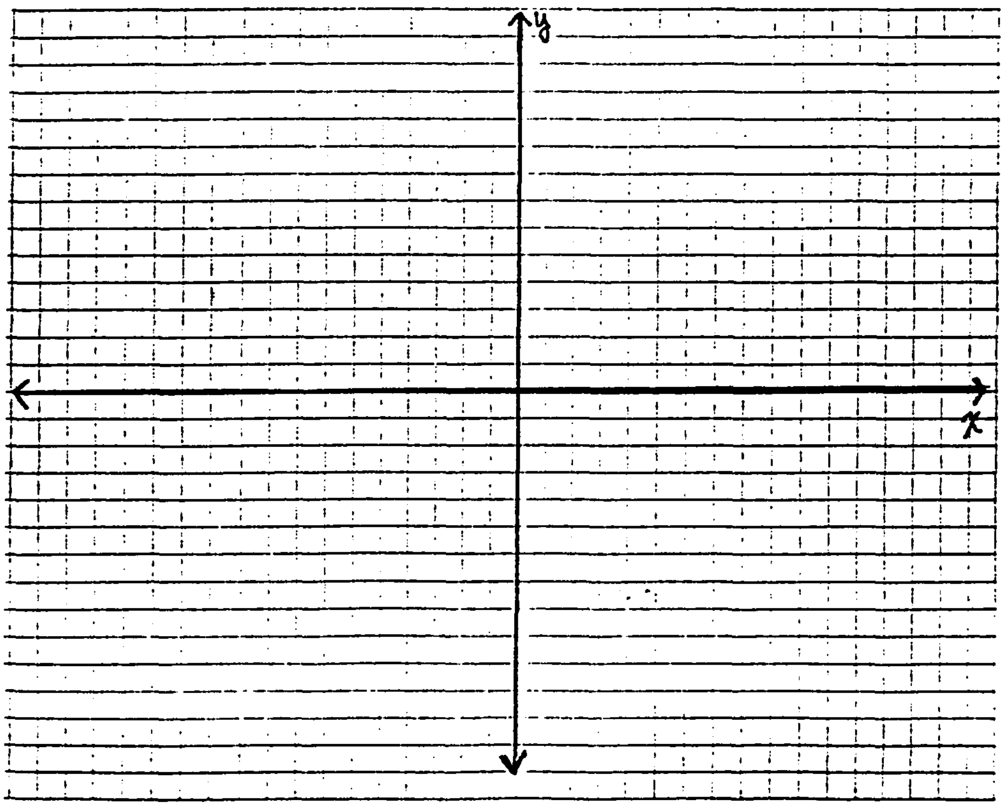
(c) $f' > 0$ on $(5, \infty)$

(d) $f''(-1) = -9$

(e) $f(2) = 0$

(f) $f'(-1) = 0$

IMPORTANT: Label any inflection points.



APPENDIX D
Quiz 3
Type 2 problem

4. Sketch the graph of a function that satisfies all of the following conditions:

(a) f **is concave up on (1,3)**

(b) f **is decreasing on** $(6,\infty)$

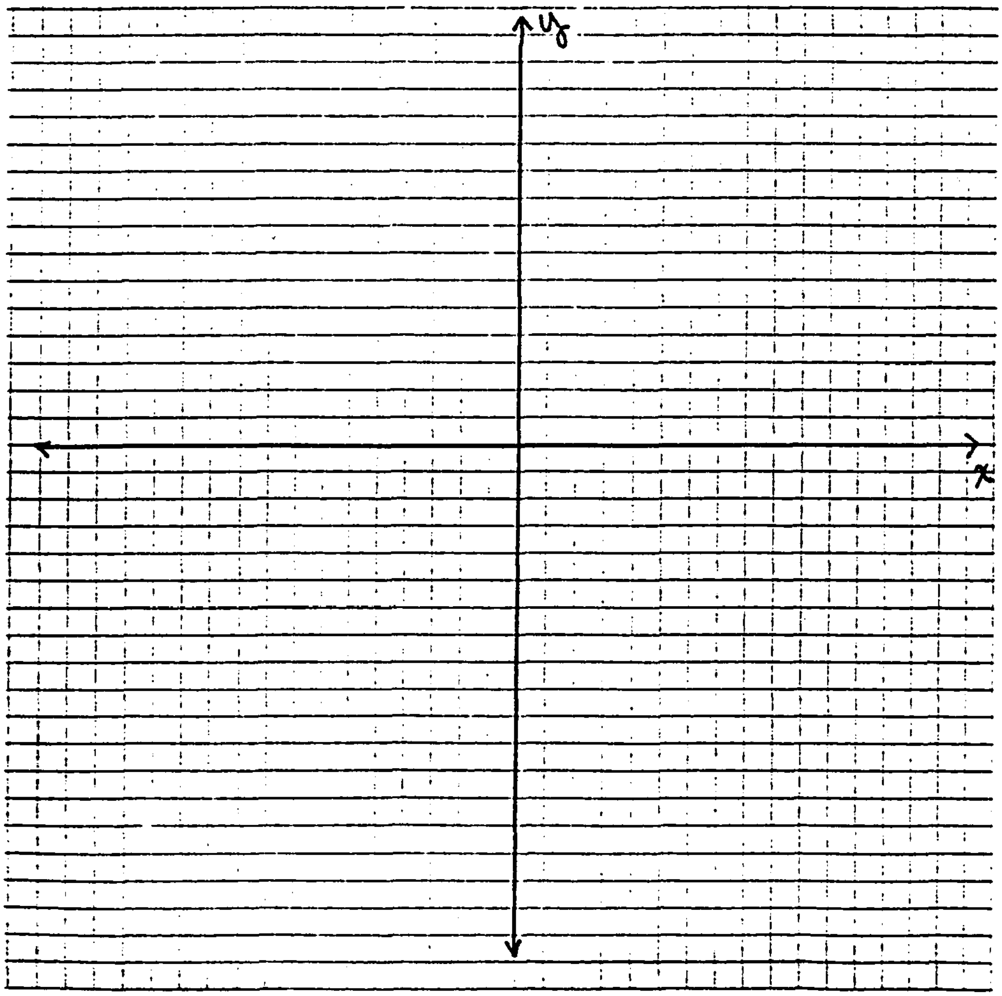
(c) f **has a local min at** $x = 2$

(d) f **is increasing on** $(-\infty,-1)$

(e) f **is concave down on** $(3,\infty)$

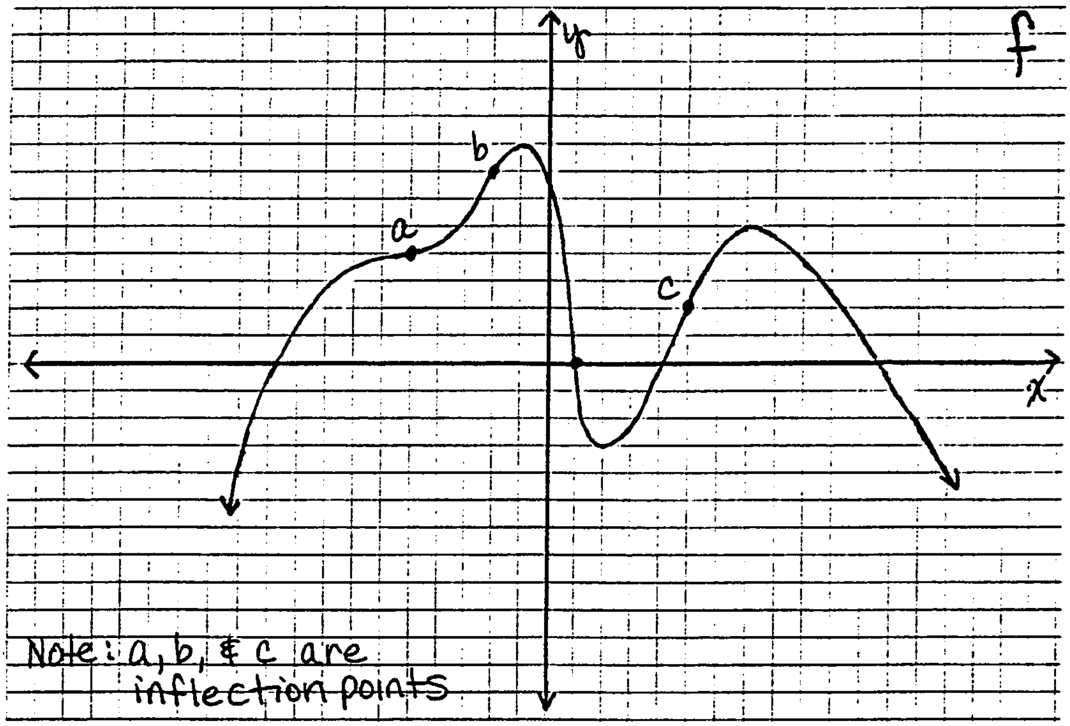
(f) f **has a local max at** $x = -1$

***IMPORTANT*: Label any inflection points.**



APPENDIX D
Quiz 4
Type 1 problem

1.



This is the graph of a function f.

(a) $f''(-5) =$ ___________

(b) On what intervals, if any, is $f' > 0$?___________________________

____

(c) $f''(7) =$ ___________

(d) $f'(7) =$ ___________

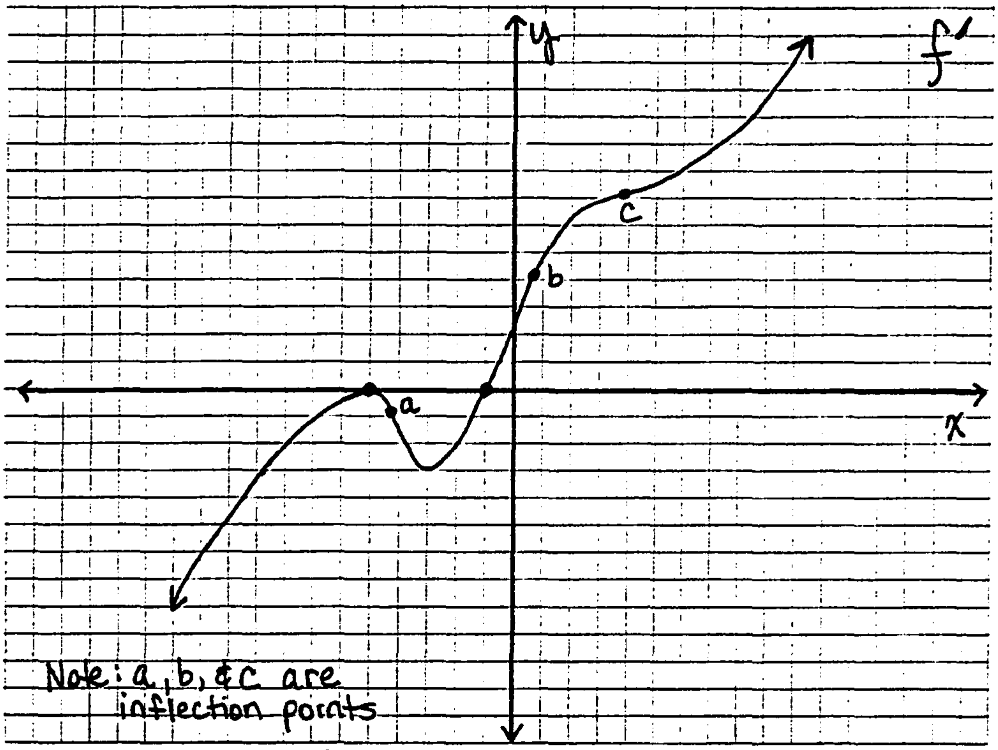
(e) On what intervals, if any, is $f'' < 0$?___________________________

____

(f) $f(1) =$ ___________

Quiz 4
Type 1 problem

3.



This is the graph of f' (the derivative of f).

(a) What values, if any, are critical values of f?____________________

(b) $f''(-5) =$______________

(c) At what values, if any, does f have a local max?__________________

(d) At what values, if any, does f have a local min?__________________

(e) $f'(-1) =$______________

(f) At what values, if any, does f have an inflection point?______________

APPENDIX D

Quiz 4

Type 2 problem

2. **Sketch the graph of a function that meets the following conditions:**

(a) $f'' > 0$ on $(7,9) \cup (9,\infty)$

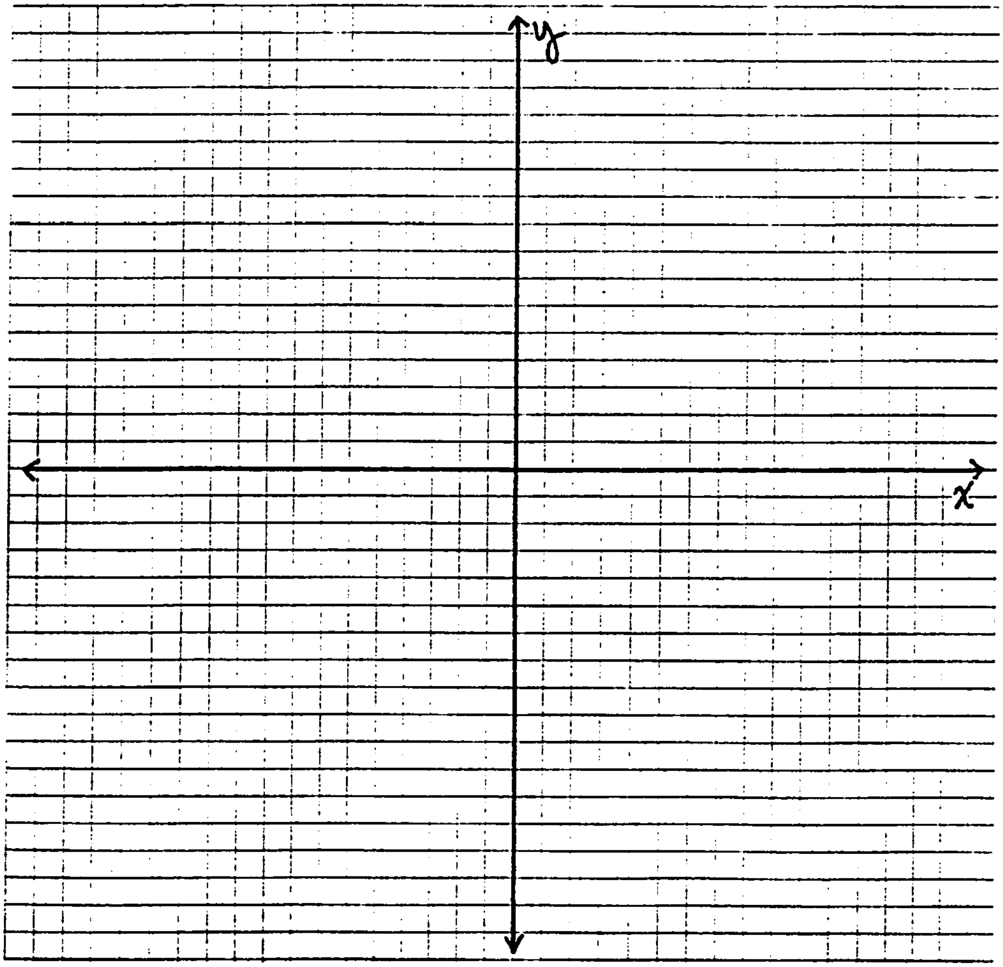
(b) f is defined and continuous on $(-\infty,\infty)$

(c) $f'' < 0$ on $(-10,-3)$

(d) $f' > 0$ on $(-\infty,-3) \cup (-3,4)$

(e) $f''(-3) = f''(9) = 0$

Important: Label any inflection points

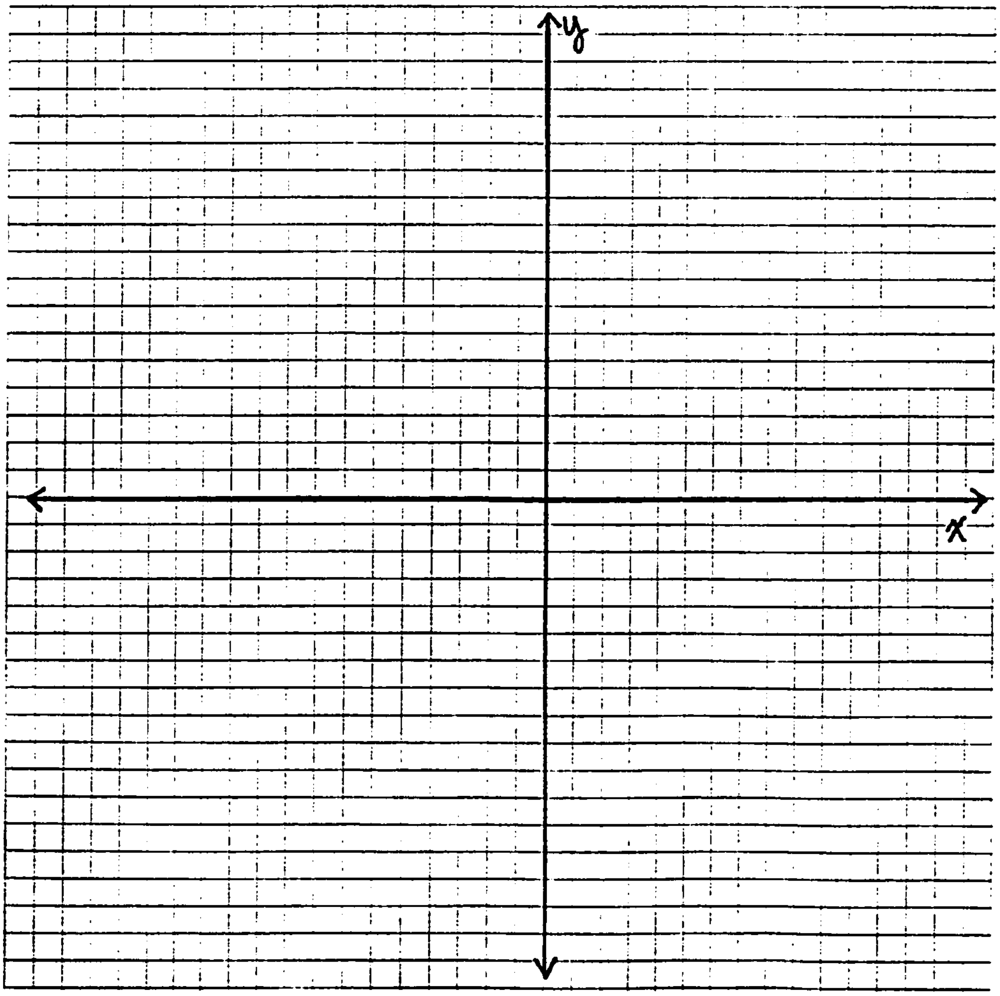


APPENDIX D
Quiz 4
Type 2 problem

4. Sketch the graph of a function that meets the following conditions:

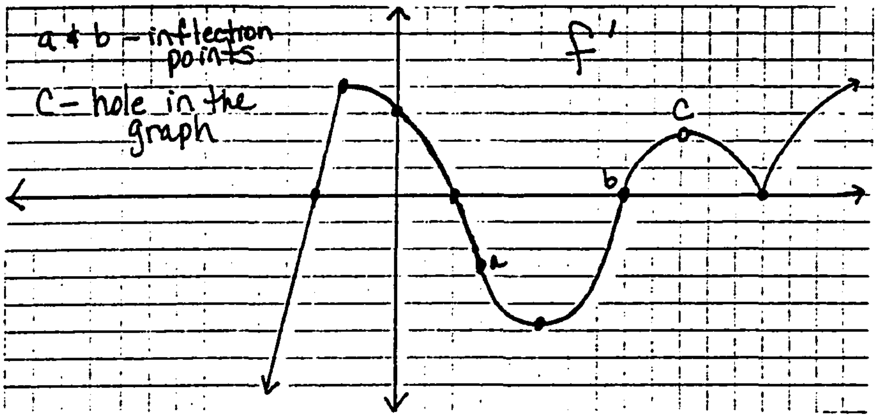
(a) $f'' > 0$ on $(-\infty,1)\cup(1,6)$
(b) $f'(11) = 0$
(c) $f'(6)$ and $f'(1)$ are undefined
(d) f is defined and continuous on $(-\infty,\infty)$
(e) $f'' < 0$ on $(6,8)$
(f) f has a local min at $x = 6$

Important: Label any inflection points



APPENDIX D
Quiz 5
Type 1 problem

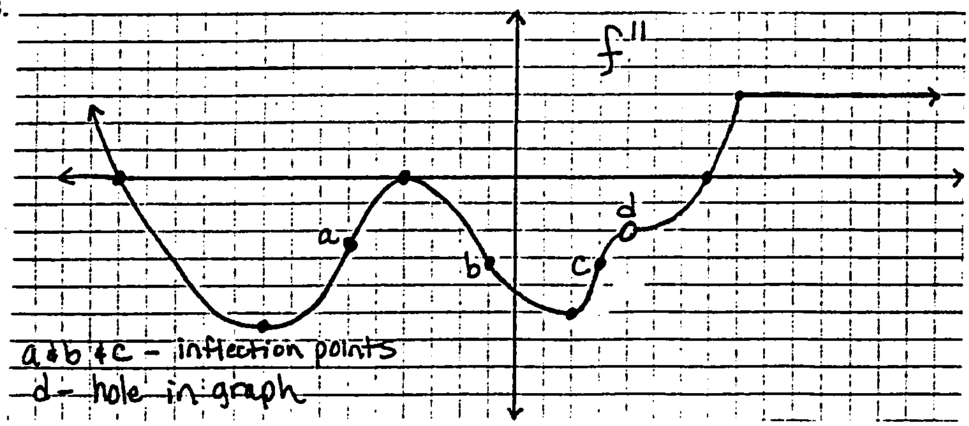
2.



(a) For what values, if any, is $f''(x) > 0$?____________________
(b) For what values, if any, does f have an inflection point?____________________
(c) For what values, if any, is $f'(x) = 0$?____________________
(d) For what values, if any, is f concave down?____________________
(e) $f''(-2)=$__________

APPENDIX D
Quiz 5
Type 1 problem

3.



(a) For what values, if any, is f' increasing?________________________

(b) For what values, if any, is $f'''(x) > 0$?________________________

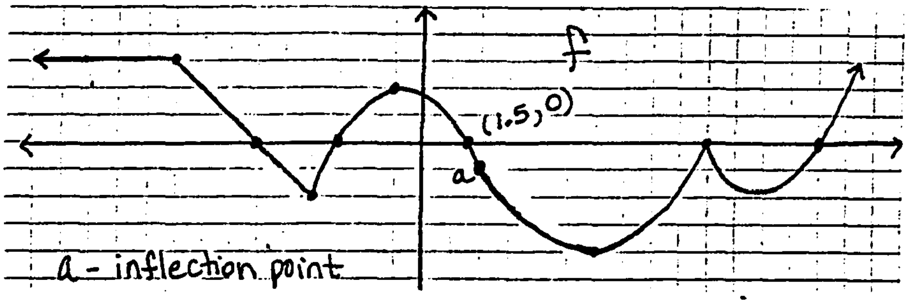
(c) For what values, if any, does f' have a local min?________________________

(d) For what values, if any, does f have an inflection point?________________________

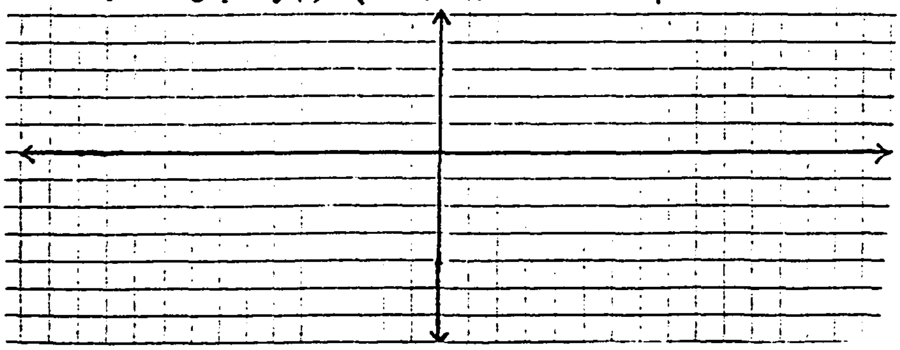
(e) What can you say about f' at $x = 4$?________________________

APPENDIX D
Quiz 5
Type 2 problems

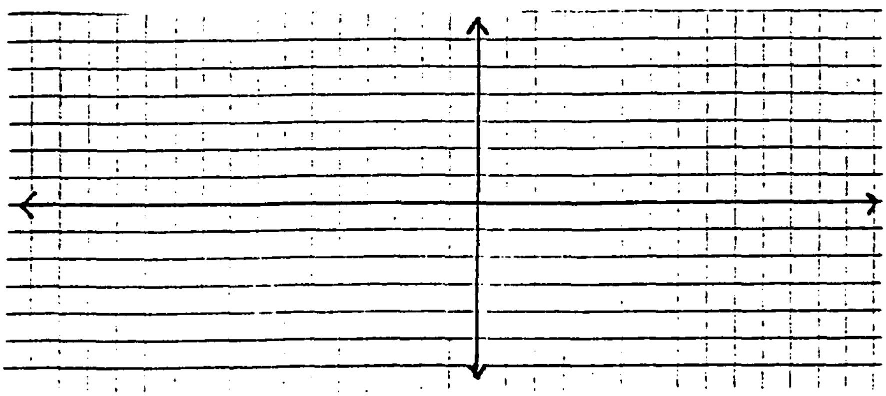
1. Given this graph of a function f,



(a) Sketch a possible graph of $f'(x)$: (label all inflection points)

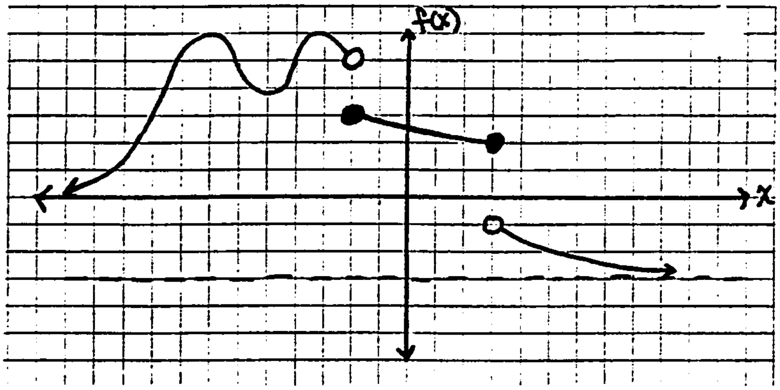


(b) Sketch a possible graph of $\int_2^x f(t)dt$: (label all inflection points)



APPENDIX D
Examination 2
Type 1 problem

6.



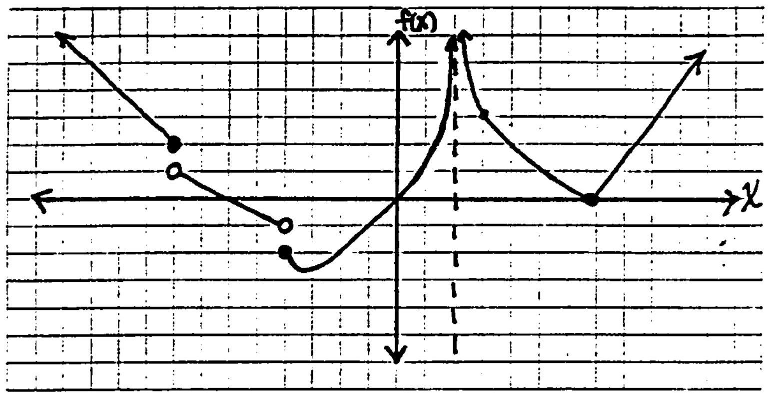
(a) $\lim_{x \to -2^-} f(x) =$ __________

(b) $f(3) =$ _______________

(c) At which points, if any, is f continuous? ____________________

(d) $\lim_{x \to \infty} f(x) =$ __________

APPENDIX D
Examination 2
Type 1 problem

10. 

(a) At which points, if any, is f discontinuous?____________________

(b) $f(2) =$ ________________

(c) $\lim_{x \to 3} f(x) =$ ________________

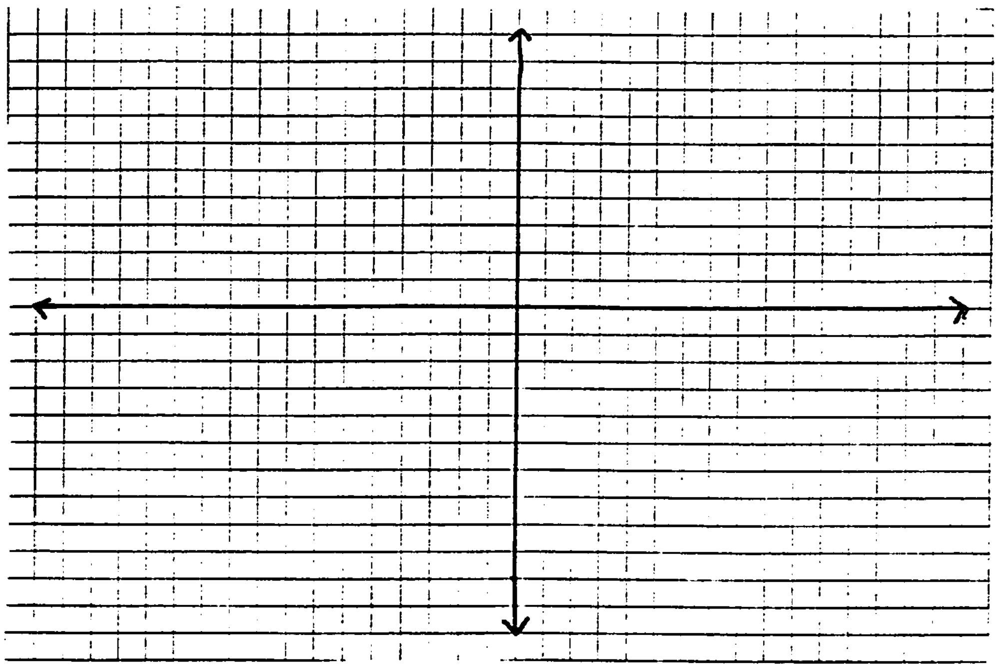
(d) $f(-4) =$ ________________

APPENDIX D
Examination 2
Type 2 problem

3. Sketch the graph of a function that meets the following conditions:

(a) $\lim_{x \to 4} f(x) = -3$ (b) f is discontinuous at $x = 2$

(c) $f(-1) = -2$ (d) $\lim_{x \to 1} f(x)$ DNE



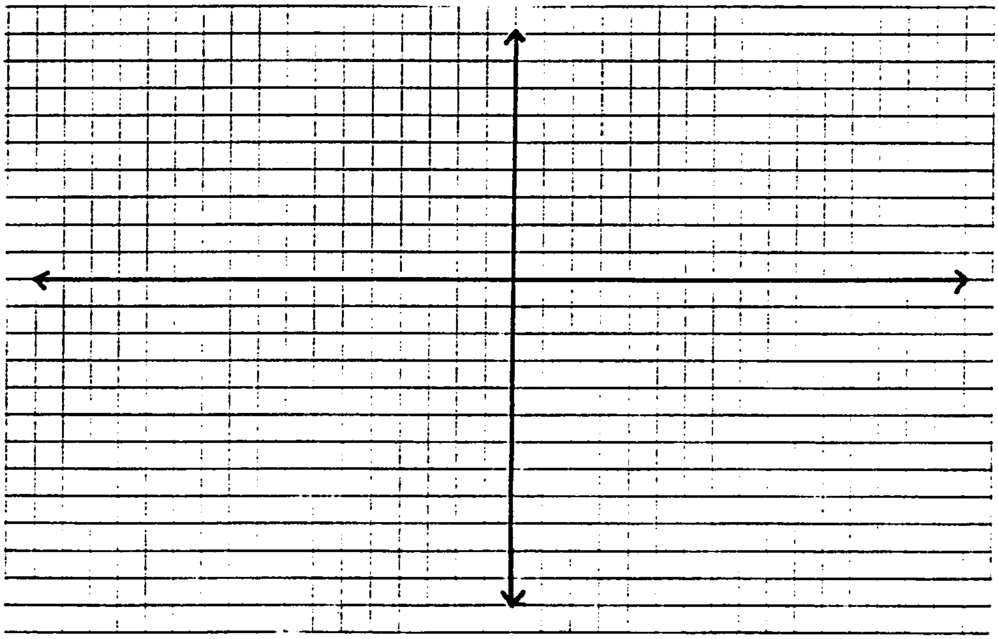
APPENDIX D
Examination 2
Type 2 problem

11. Sketch the graph of a function that meets the following conditions:

(a) $\lim_{x\to 3^+} f(x) = \infty$ (b) f is undefined at $x = 0$

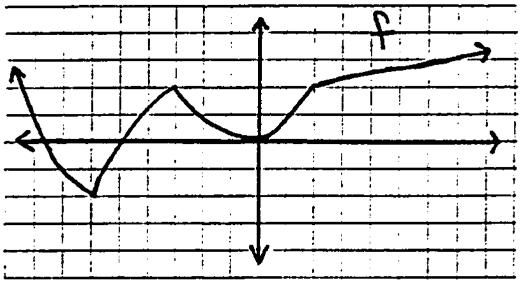
(c) f has a horizontal asymptote at $y = -3$

(d) $\lim_{x\to -\infty} f(x) = \infty$



APPENDIX D
Examination 3
Type 1 problem

1.



Note f has a local min at x=0

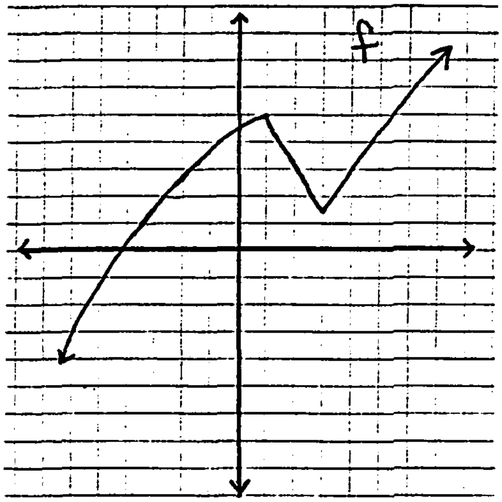
(a) At what values, if any, does f have an inflection point?________________

(b) At what values, if any, is $f'(x) > 0$?________________

(c) $f''(0) =$________

APPENDIX D
Examination 3
Type 1 problem

8.



(a) At what values, if any, is $f'(x) = 0$?____________________

(b) $f''(1) =$__________

(c) At what values, if any, does f have a local min?____________________

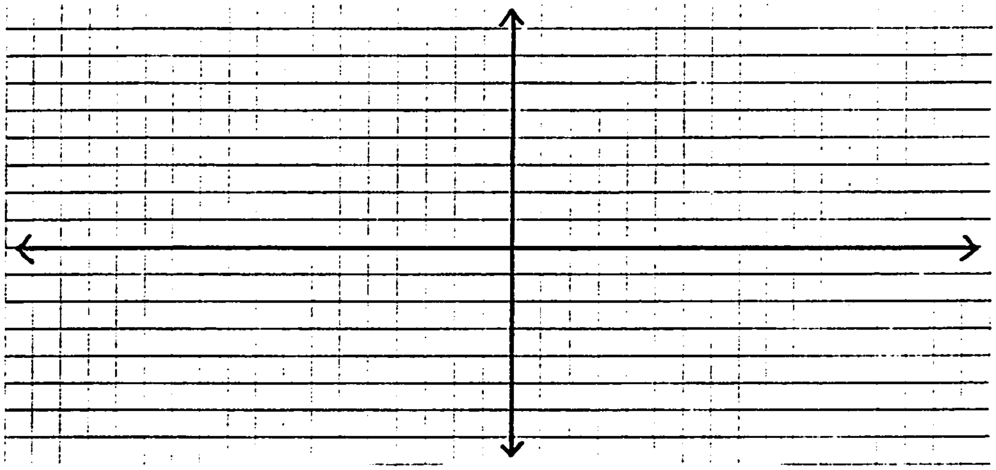
APPENDIX D
Examination 3
Type 2 problem

4. **Sketch the graph of a function that meets the following conditions:**

(a) $f'(x)=1$ on $(3,6)$ (b) $f''>0$ on $(7,\infty)$

(c) $f''(-2)=0$ (d) $f'(-4)=0$

Label any inflection points.



APPENDIX D
Examination 3
Type 2 problem

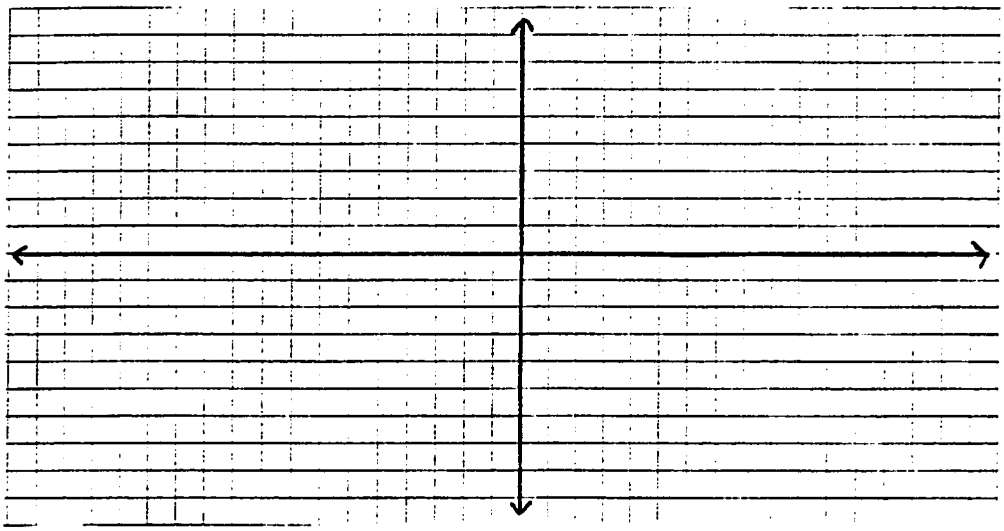
6. Sketch the graph of a function that meets the following conditions:

(a) $f''(x) = 0$ on $(4,\infty)$ (b) f is defined on $(-\infty,\infty)$

(c) $f' < 0$ on $(1,4) \cup (4,\infty)$ (d) $f'' > 0$ on $(0,4)$

(e) $f'(x) = -1$ on $(-8,-4)$

Label any inflection points.

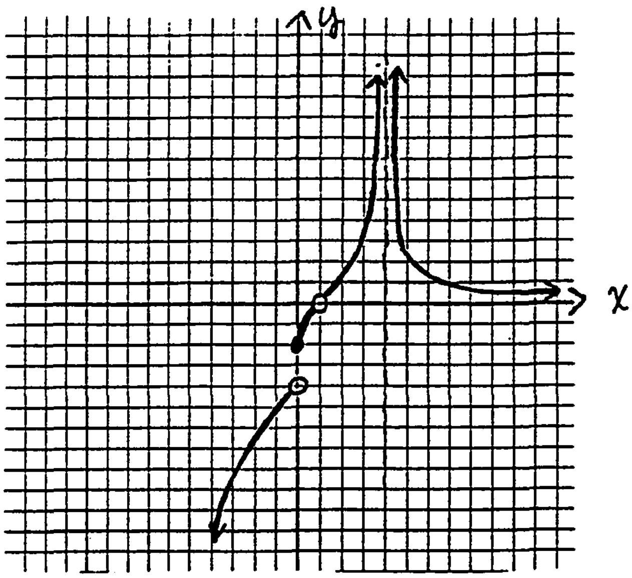


APPENDIX E

Examples of Student Solutions

APPENDIX E
Quiz 1
Type 1 problem

1.



(a) $\lim_{x \to \infty} f(x) =$ 0

(b) $f(1) =$ DNE

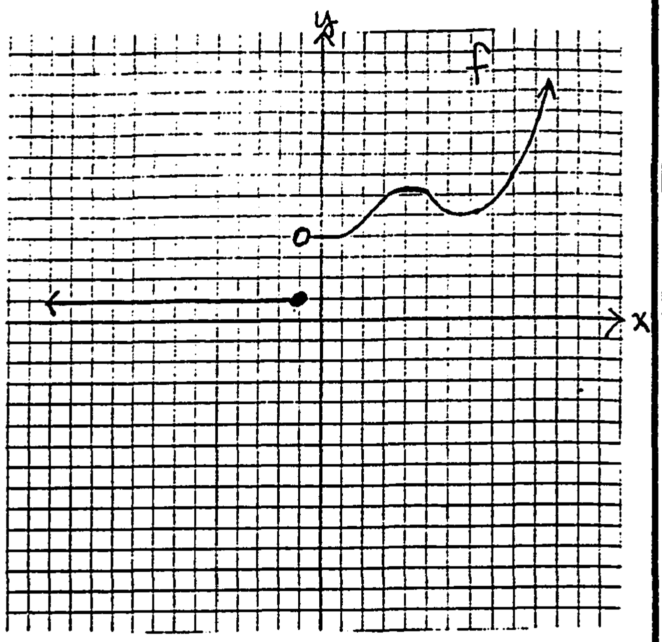
(c) $\lim_{x \to 0^-} f(x) =$ -4

(d) At what values of x, if any, is f discontinuous? 0, 1, 4

Quiz 1

Type 1 problem

3.



(a) $f(-1) =$ 1

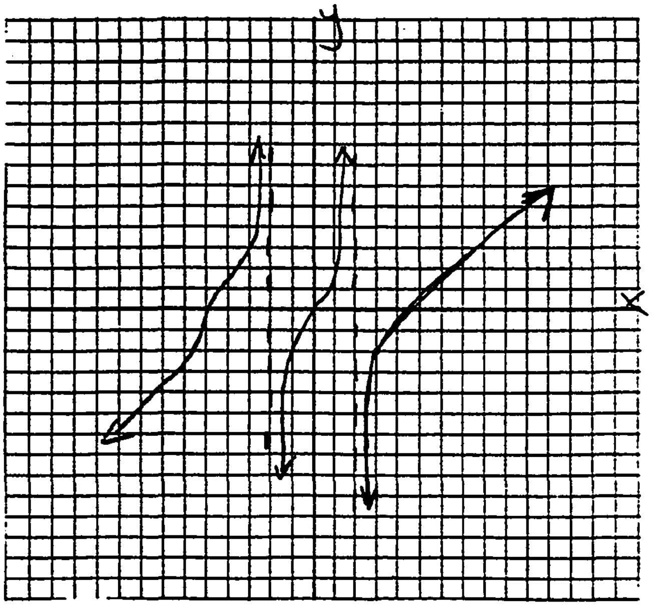
(b) For what values of x, if any, is f continuous? $(-\infty, -1) \cup (-1, \infty)$

(c) $\lim_{x \to -\infty} f(x) =$ 1

(d) $\lim_{x \to -1} f(x) =$ DNE

APPENDIX E
Quiz 1
Type 2 problem

2. Sketch the graph of a function that satisfies *all* of the following conditions:
(a) $\lim_{x\to\infty} f(x) = \infty$
(b) f has a vertical asymptote at $x = 2$
(c) $\lim_{x\to -2} f(x)$ DNE
(d) $f(3) = -2$

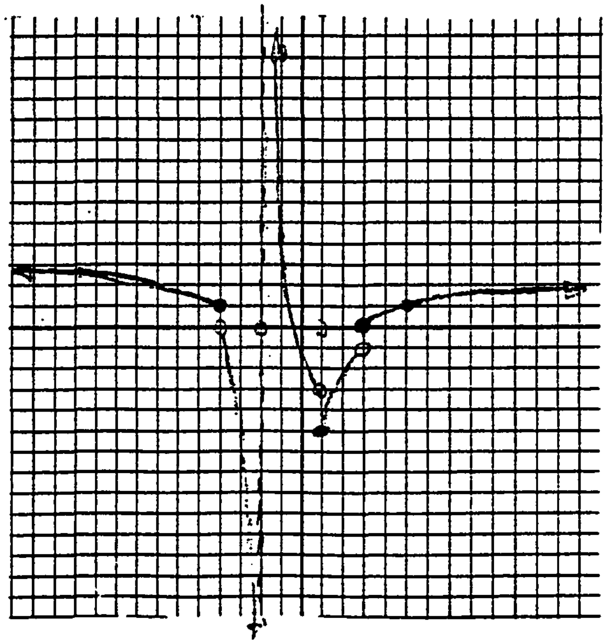


scale = 1

Quiz 1

Type 2 problem

4. **Sketch the graph of a function that satisfies *all* of the following conditions:**
   (a) f is defined for all real numbers except $x = -2$
   (b) f is continuous everywhere ***except*** at $x = -4$, $x = -2$, $x = 1$, and $x = 3$.
   (c) $\lim_{x \to \infty} f(x) = 2$
   (d) $f(-2)$ is undefined
   (e) $\lim_{x \to 5^-} f(x) = 1$

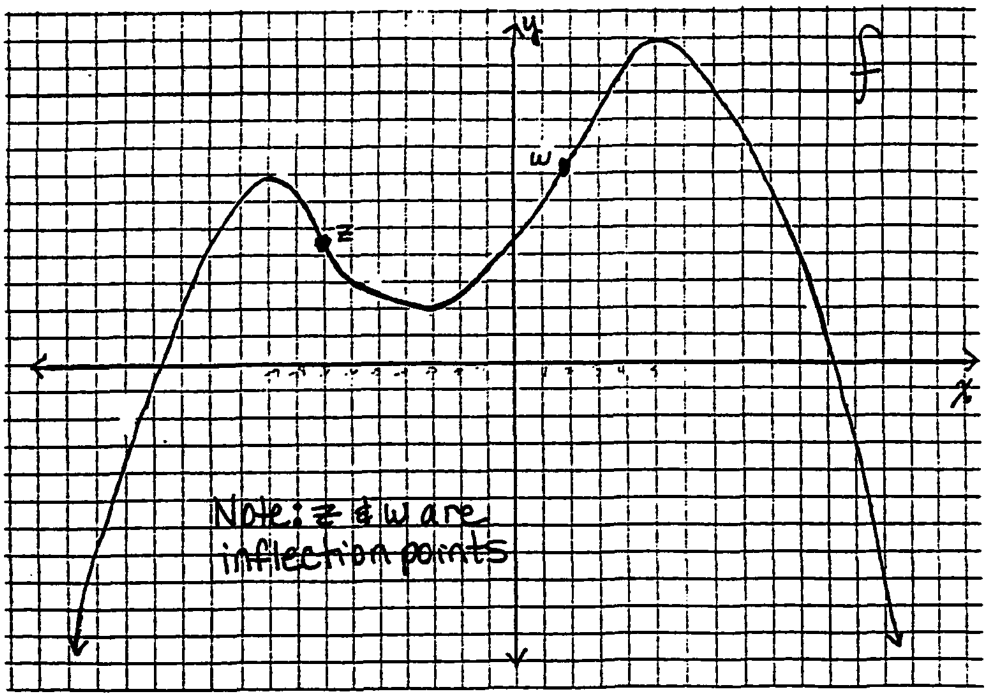


APPENDIX E

Quiz 3

Type 1 problem

2.



(a) On what interval(s), if any, is f increasing? $(-\infty, -9)$ and $(-3, 5)$

(b) On $(-\infty, \infty)$, at what value(s), if any, does f have a local max?

$x = -9$ and $x = 5$

(c) On what interval(s), if any, is f concave up?

$(-7, 2)$

(d) On what interval(s), if any, is f decreasing?

$(-9, -3)$ and $(5, \infty)$

(e) On what interval(s), if any, is f concave down?

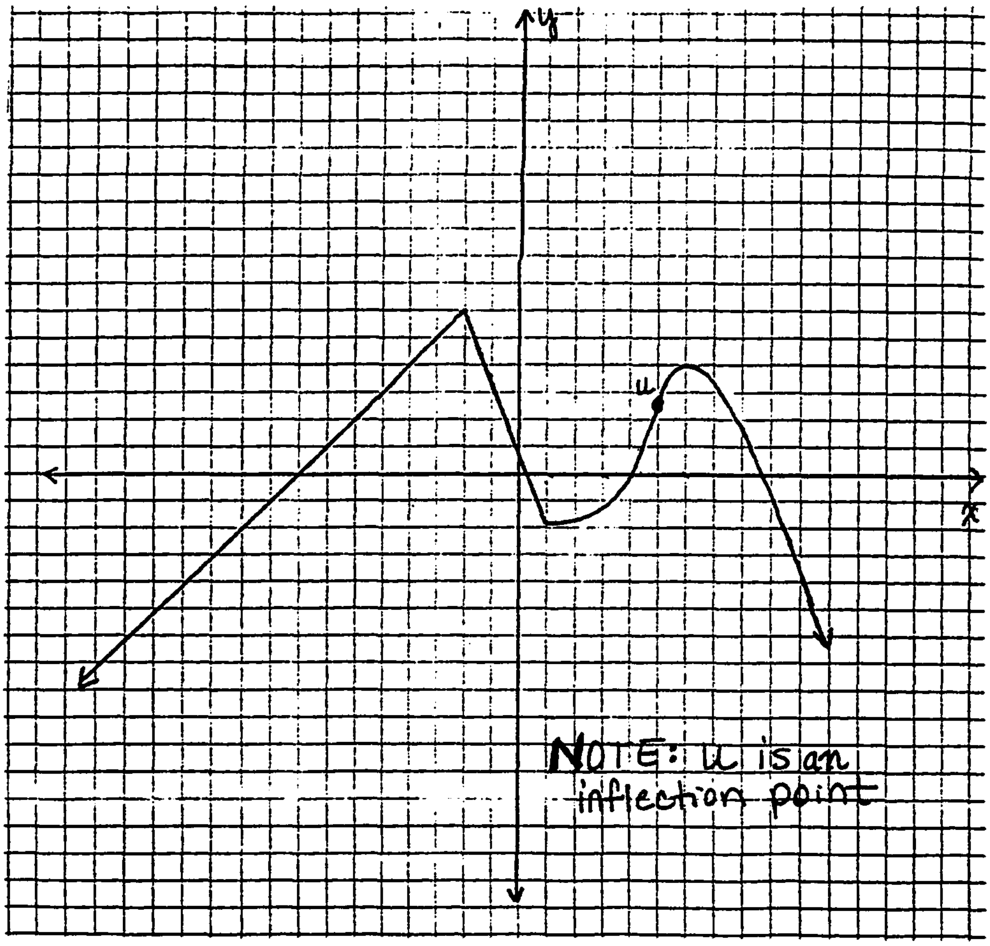
$(-\infty, -7)$ and $(2, \infty)$

(f) On $(-\infty, \infty)$, at what value(s), if any, does f have a local min?

$x = -3$

APPENDIX E
Quiz 3
Type 1 problem

3.



(a) $f'(-2) =$ DNE

(b) At what value(s), if any, does f have a local max?

x = -2, 6

(c) At what value(s), if any, does f have a local min?

x = 1

(d) $f(4) =$ 0

(e) $f'(6) =$ 0

(f) Is $f'(5)$ negative or positive? positive

APPENDIX E
Quiz 3
Type 2 problem

1. Sketch the graph of a function f that satisfies all of the following conditions:

(a) f' DNE at $x = 4$ and $x = 2$

(b) $f(-1) = -2$

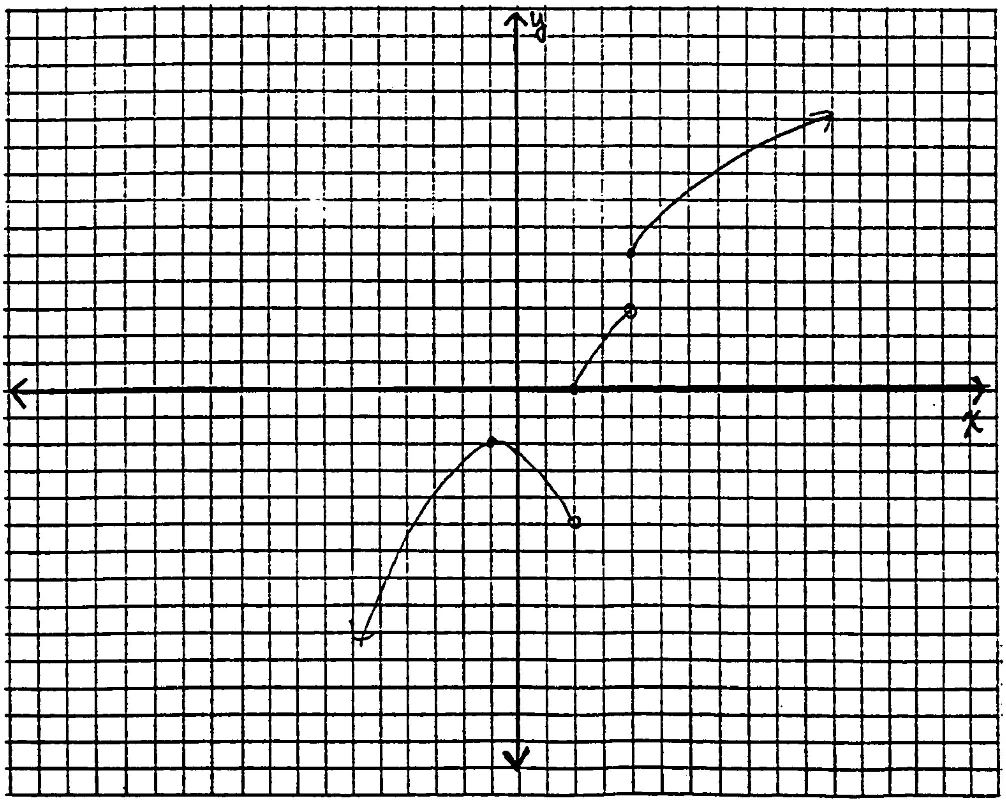
(c) $f' > 0$ on $(5, \infty)$

(d) $f''(-1) = -9$

(e) $f(2) = 0$

(f) $f'(-1) = 0$

IMPORTANT: Label any inflection points.



APPENDIX E
Quiz 3
Type 2 problem

4. Sketch the graph of a function that satisfies all of the following conditions:

(a) f is concave up on $(1,3)$

(b) f is decreasing on $(6,\infty)$

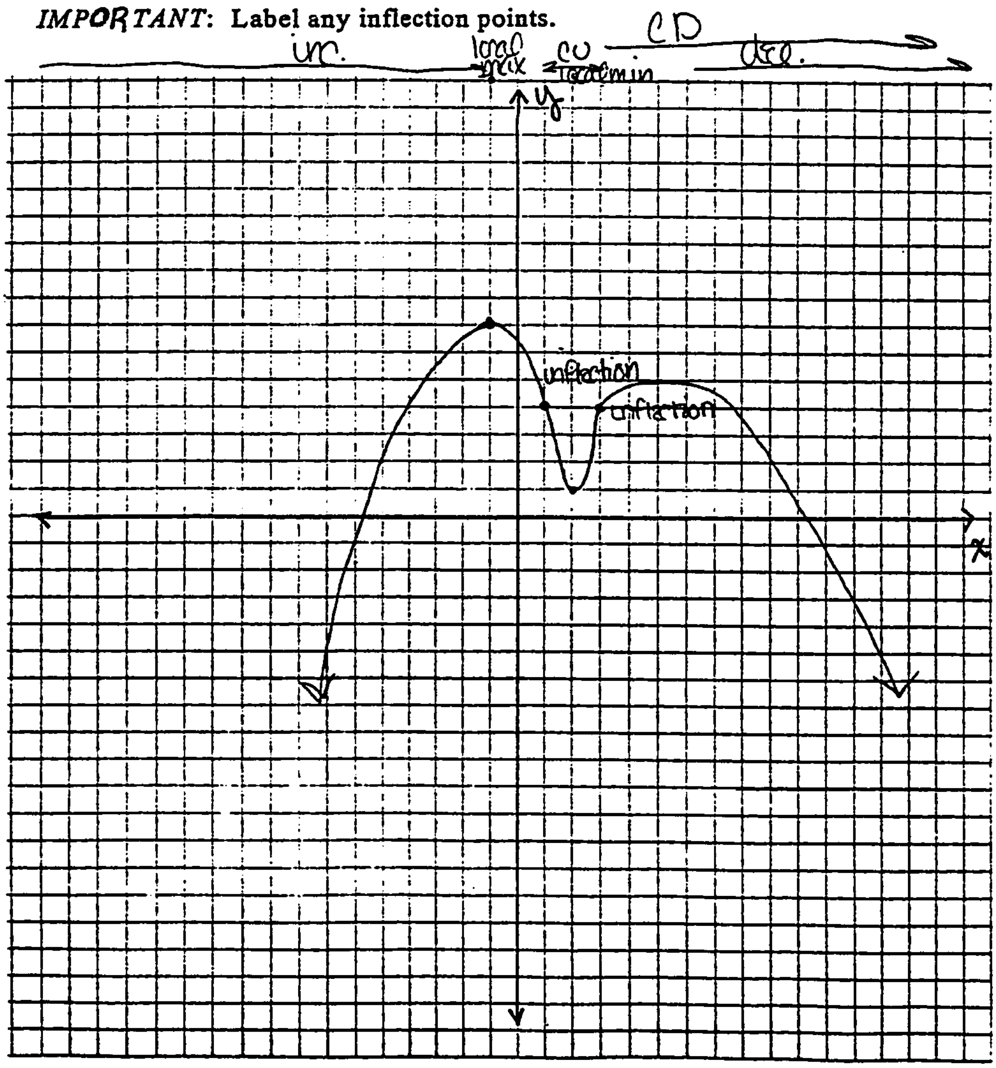
(c) f has a local min at $x = 2$

(d) f is increasing on $(-\infty,-1)$

(e) f is concave down on $(3,\infty)$

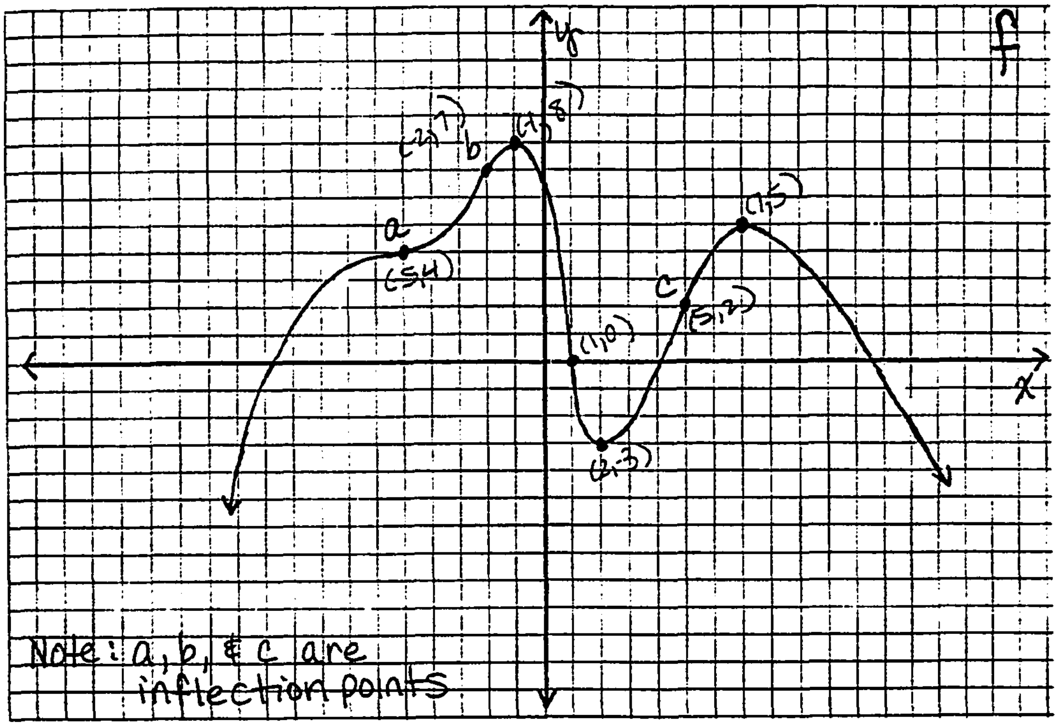
(f) f has a local max at $x = -1$

IMPORTANT: Label any inflection points.



APPENDIX E
Quiz 4
Type 1 problem

1.



This is the graph of a function f.

(a) $f''(-5) =$ 0

(b) On what intervals, if any, is $f' > 0$? $(-\infty, -1) \cup (2, 7)$

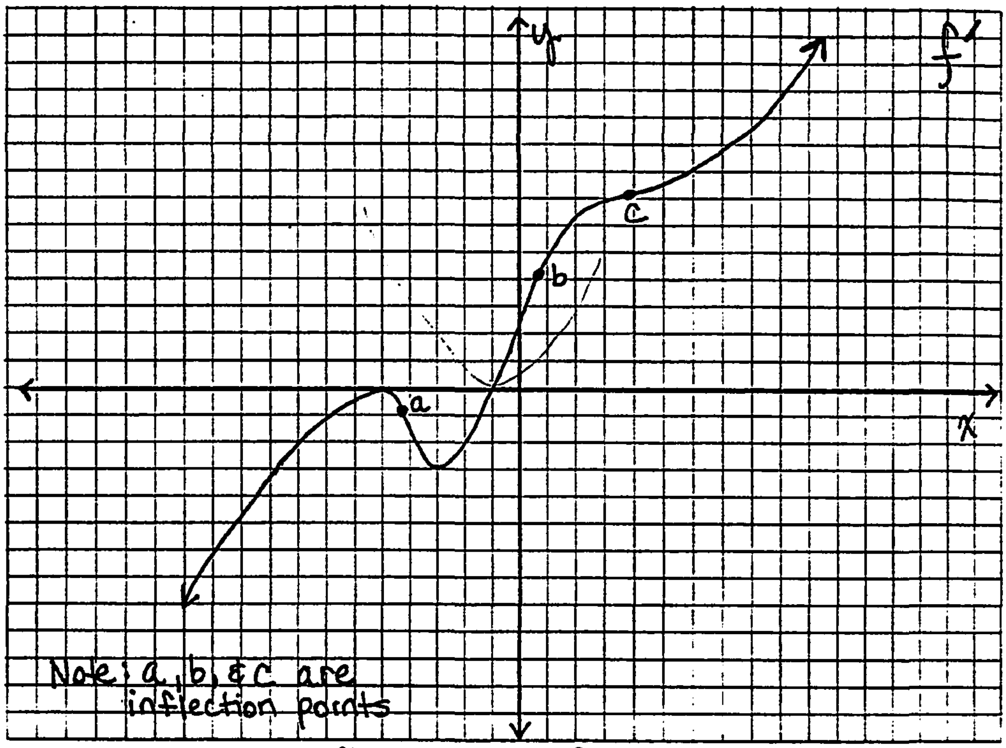
(c) $f''(7) =$ < 0

(d) $f'(7) =$ 0

(e) On what intervals, if any, is $f'' < 0$? $(-\infty, -5) \cup (-2, 1) \cup (5, +\infty)$

(f) $f(1) =$ 0

3.



This is the graph of f' (the derivative of f).

(a) What values, if any, are critical values of f? x = -1 and x = -5

(b) $f''(-5) =$ 0

(c) At what values, if any, does f have a local max? none

(d) At what values, if any, does f have a local min? x = -1

(e) $f'(-1) =$ 0

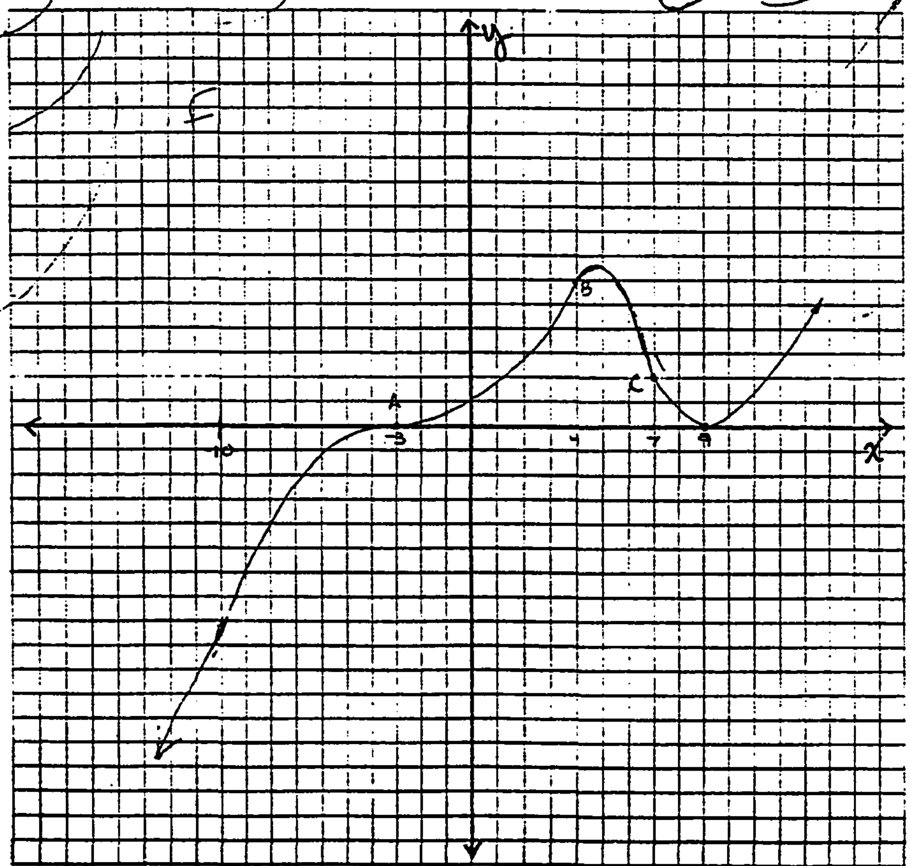
(f) At what values, if any, does f have an inflection point? x = -3 and x = -5

APPENDIX E
Quiz 4
Type 2 problem

2. Sketch the graph of a function that meets the following conditions:

(a) $f'' > 0$ on $(7,9) \cup (9,\infty)$ CU
(b) f is defined and continuous on $(-\infty,\infty)$
✓ CD (c) $f'' < 0$ on $(-10,-3)$
✓ ↑ easing (d) $f' > 0$ on $(-\infty,-3) \cup (-3,4)$
(e) $f''(-3) = f''(9) = 0$ Inf

Important: Label any inflection points



Inflections A, B, C

APPENDIX E

Quiz 4

Type 2 problem

4. Sketch the graph of a function that meets the following conditions:

(a) $f'' > 0$ on $(-\infty,1) \cup (1,6)$

(b) $f'(11) = 0$

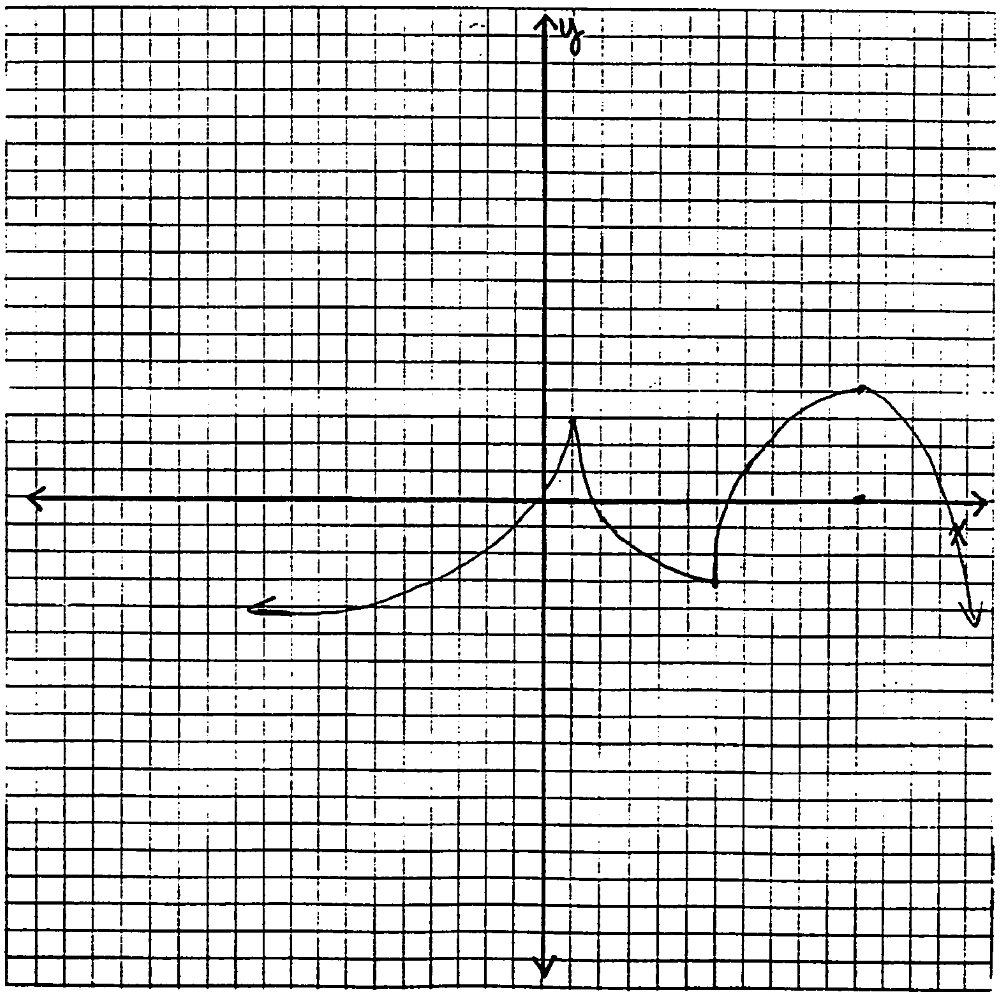
(c) $f'(6)$ and $f'(1)$ are undefined

(d) f is defined and continuous on $(-\infty,\infty)$

(e) $f'' < 0$ on $(6,8)$

(f) f has a local min at $x = 6$

Important: Label any inflection points

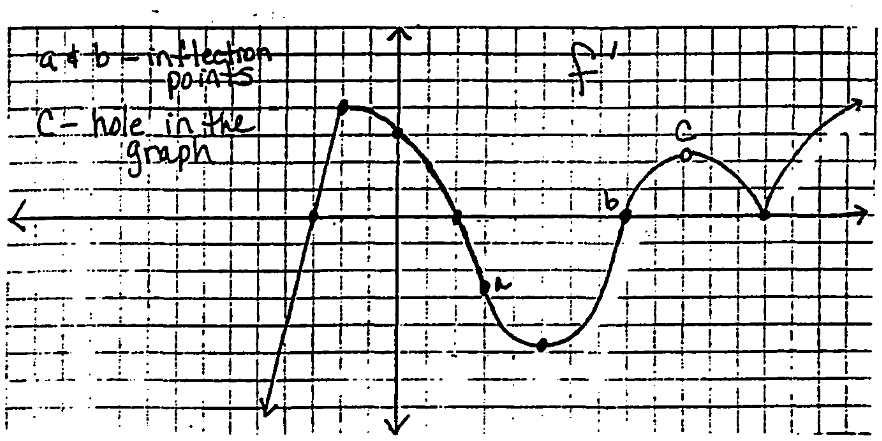


APPENDIX E

Quiz 5

Type 1 problem

2.



(a) For what values, if any, is $f''(x) > 0$? $(-\infty, -2)$ and $(5, 10)$ and $(13, \infty)$

(b) For what values, if any, does f have an inflection point? -2, 5

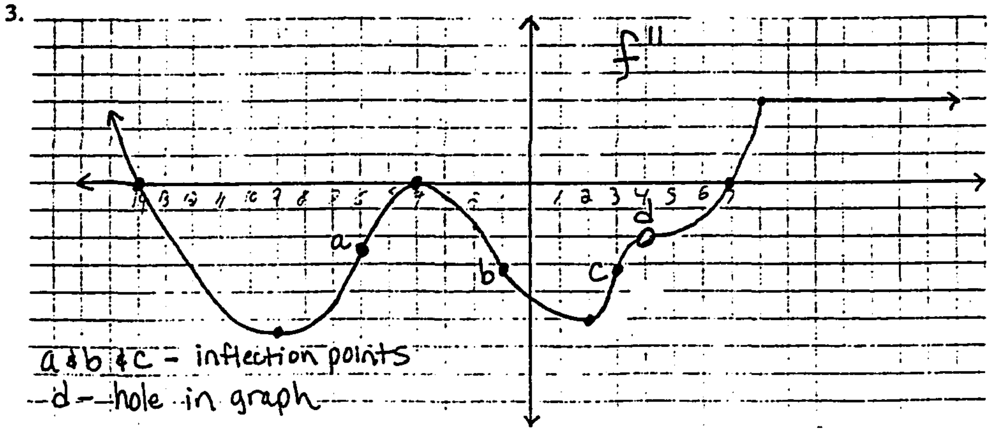
(c) For what values, if any, is $f'(x) = 0$? -3, 2, & 13

(d) For what values, if any, is f concave down? $(-2, 5)$ and $(10, 13)$

(e) $f''(-2) =$ undefined

APPENDIX E
Quiz 5
Type 1 problem

3.



(a) For what values, if any, is f' increasing? $(-\infty, -14)$ and $(7, \infty)$

(b) For what values, if any, is $f'''(x) > 0$? $(-9,-4)$, $(2, 4)$, and $(4, 8)$

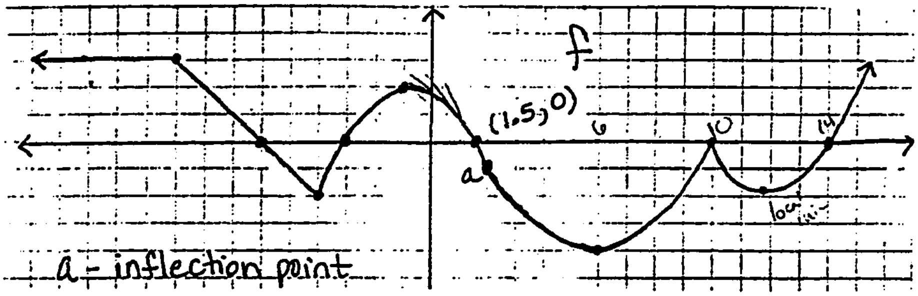
(c) For what values, if any, does f' have a local min? $x = 7$

(d) For what values, if any, does f have an inflection point? $x = -14$ and $x = 7$

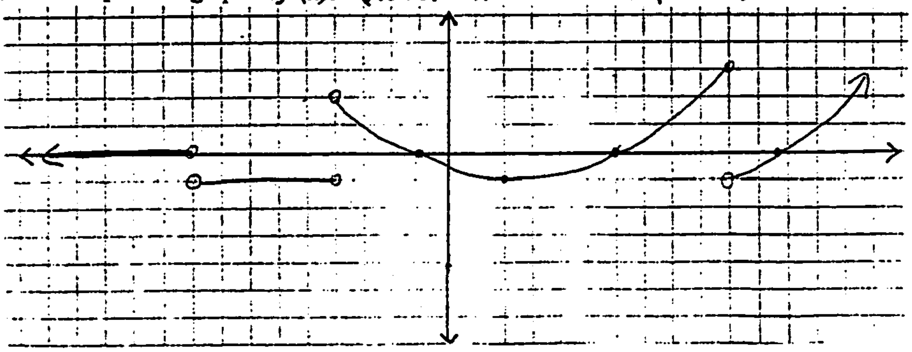
(e) What can you say about f' at $x = 4$? There is a hole or kink in the graph of f' at x=4.

APPENDIX E
Quiz 5
Type 2 problems

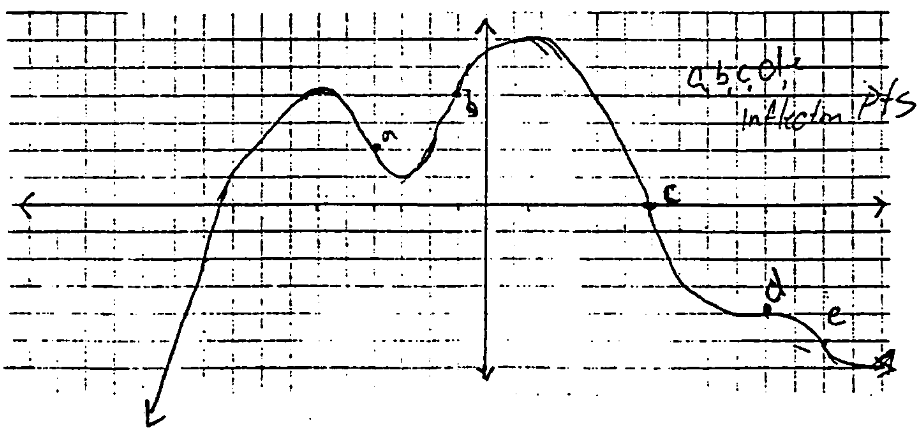
1. Given this graph of a function f,



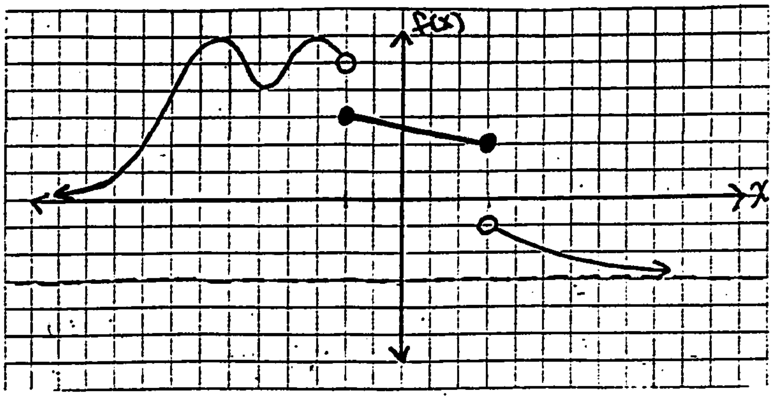
(a) Sketch a possible graph of $f'(x)$: (label all inflection points)



(b) Sketch a possible graph of $\int_2^x f(t)dt$: (label all inflection points)



APPENDIX E
Examination 2
Type 1 problem



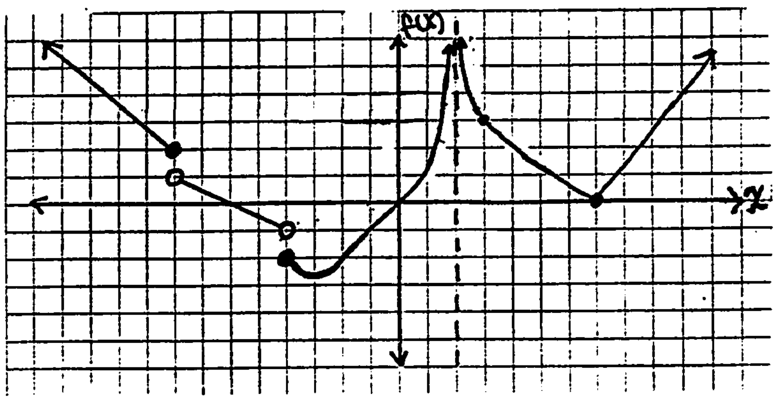
(a) $\lim_{x \to -2^-} f(x) =$ 5

(b) $f(3) =$ 2

(c) At which points, if any, is f continuous? All real numbers except at $x = -2$ and $x = 3$

(d) $\lim_{x \to \infty} f(x) =$ −3

APPENDIX E
Examination 2
Type 1 problem



(a) At which points, if any, is f discontinuous? -8, -4, 2.

(b) $f(2) =$ undefined

(c) $\lim_{x \to 3} f(x) =$ 3

(d) $f(-4) =$ -2

~~(-∞, -8] ∪ (-8, -4) ∪ [-4, 2) ∪ (2, ∞)~~

APPENDIX E
Examination 2
Type 2 problem

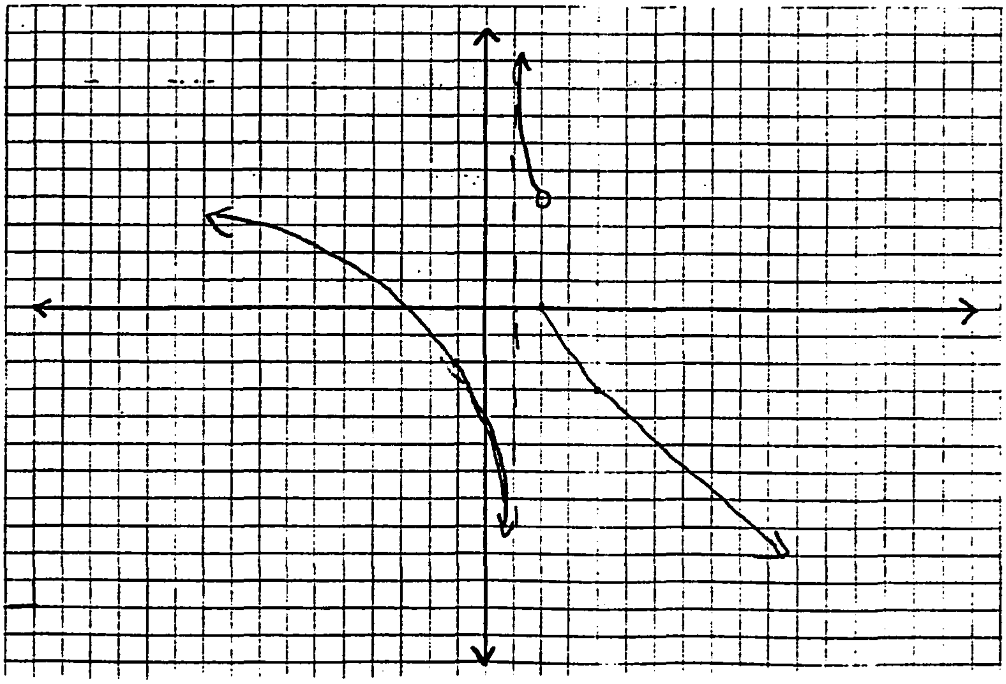
Sketch the graph of a function that meets the following conditions:

(a) $\lim_{x \to 4} f(x) = -3$

(b) f is discontinuous at $x = 2$

(c) $f(-1) = -2$

(d) $\lim_{x \to 1} f(x)$ DNE



APPENDIX E

Examination 2

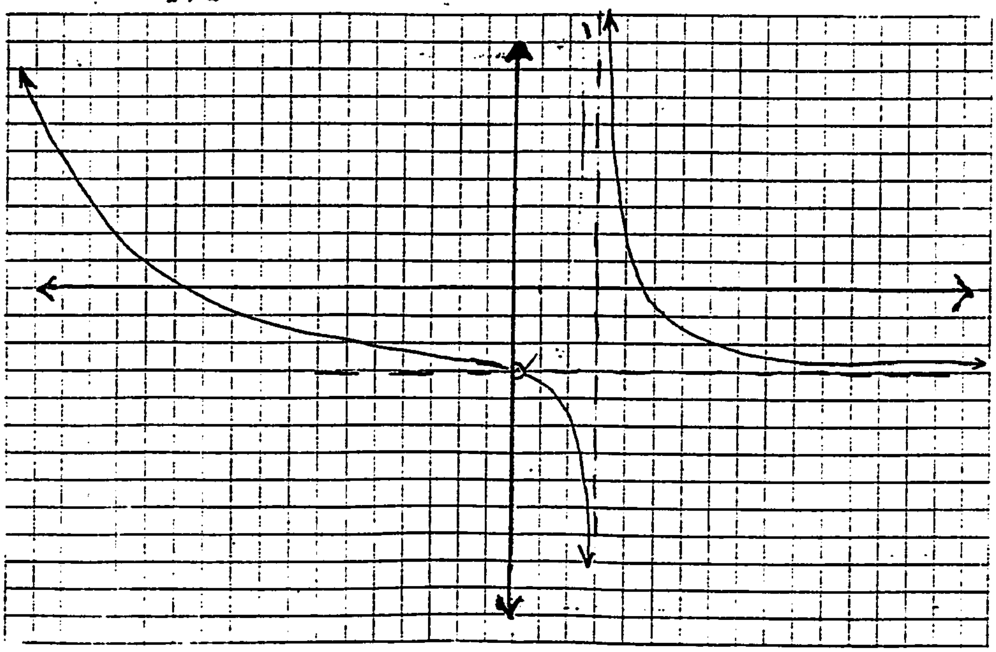
Type 2 problem

11. Sketch the graph of a function that meets the following conditions:

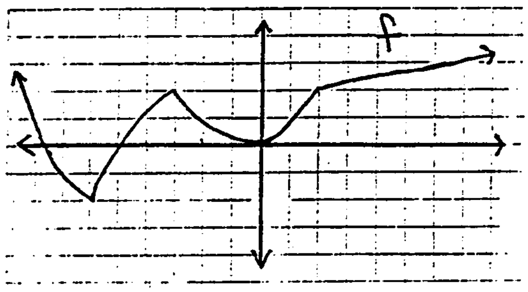
✓ (a) $\lim_{x \to 3^+} f(x) = \infty$ ✓ (b) f is undefined at $x = 0$

✓ (c) f has a horizontal asymptote at $y = -3$

✓ (d) $\lim_{x \to -\infty} f(x) = \infty$



APPENDIX E
Examination 3
Type 1 problem

1. 

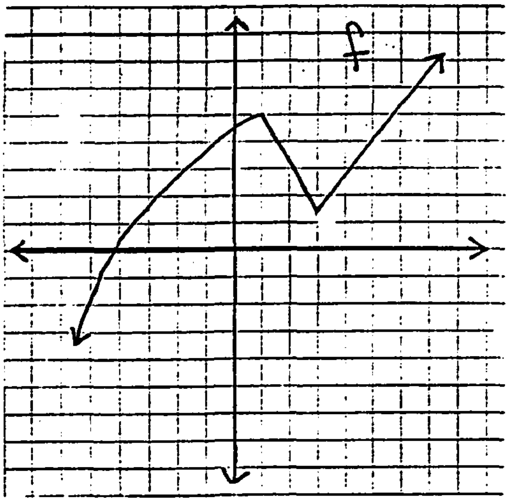
Note f has a local min at x=0

(a) At what values, if any, does f have an inflection point? No inflection

(b) At what values, if any, is $f'(x) > 0$? $(-6,-3), (0,2), (2,\infty)$

(c) $f''(0) =$ > 0

8.



= loc. max or m
f' = increasing or decreasing

f'' = CU
CD
inflection point

(a) At what values, if any, is $f'(x) = 0$? none

(b) $f''(1) =$ DNE

(c) At what values, if any, does f have a local min? $x = 3$

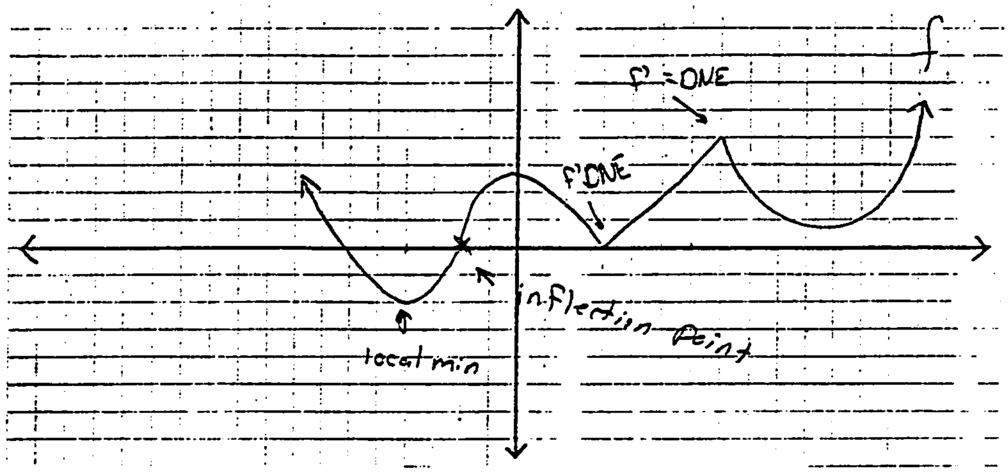
APPENDIX E
Examination 3
Type 2 problem

4. **Sketch the graph of a function that meets the following conditions:**

(a) $f'(x)=1$ on $(3,6)$ (b) $f''>0$ on $(7,\infty)$

(c) $f''(-2)=0$ (d) $f'(-4)=0$

Label any inflection points.



APPENDIX E
Examination 3
Type 2 problem

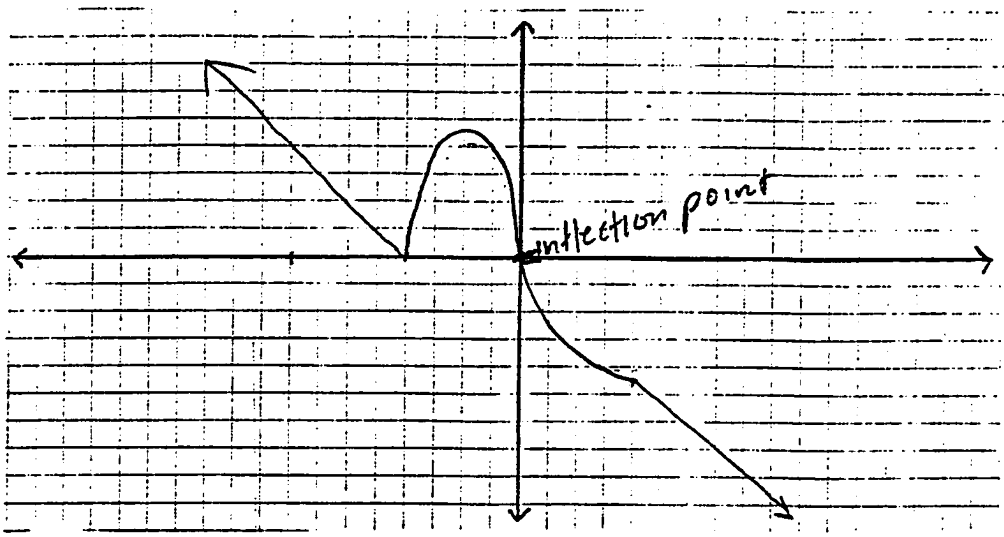
6. Sketch the graph of a function that meets the following conditions:

(a) $f''(x) = 0$ on $(4,\infty)$ (b) f is defined on $(-\infty,\infty)$

(c) $f' < 0$ on $(1,4) \cup (4,\infty)$ (d) $f'' > 0$ on $(0,4)$

(e) $f'(x) = -1$ on $(-8,-4)$

Label any inflection points.



APPENDIX F

Steve's Solution to Problem 4 on Quiz 3

APPENDIX F
Steve's Solution to Problem 4 on Quiz 3

4. Sketch the graph of a function that satisfies all of the following conditions:

~~(a) f~~ is concave up on (1,3)

~~(b) f~~ is decreasing on $(6,\infty)$

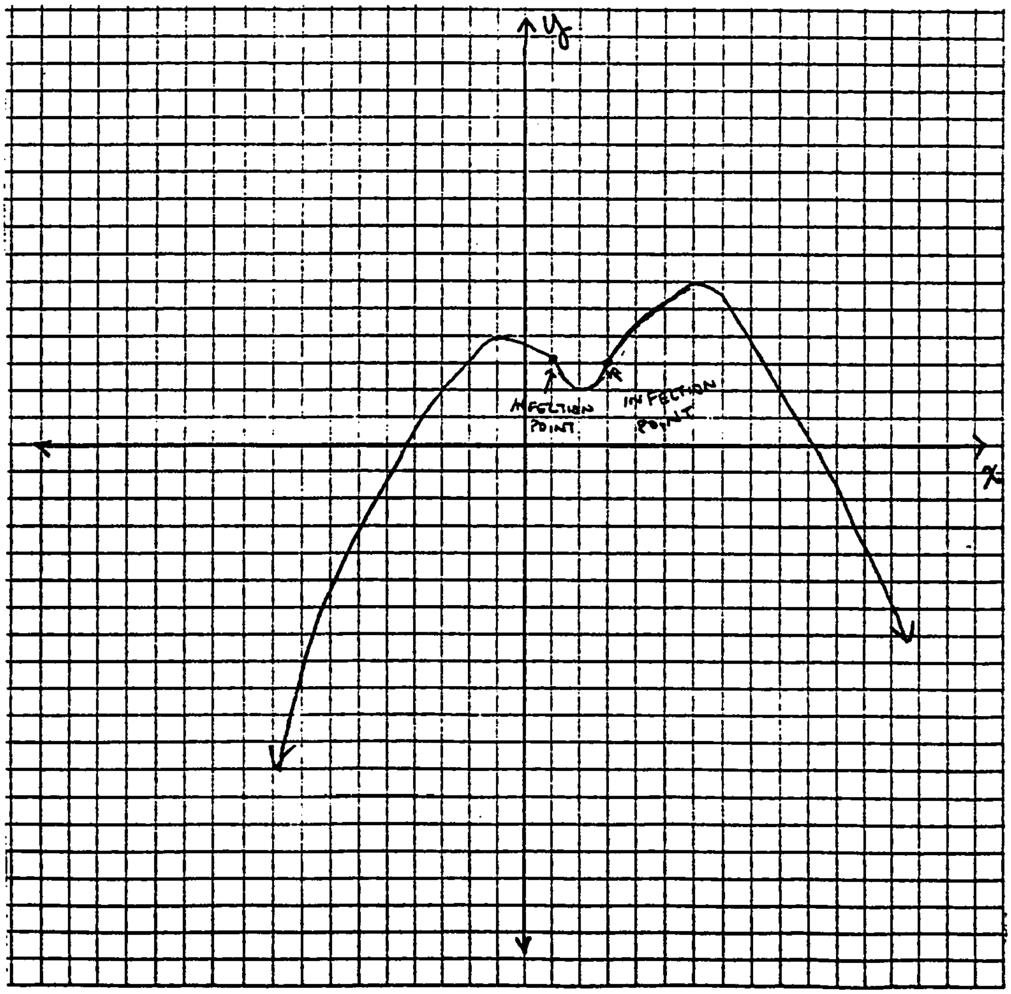
~~(c) f~~ has a local min at $x = 2$

~~(d) f~~ is increasing on $(-\infty,-1)$

~~(e) f~~ is concave down on $(3,\infty)$

~~(f) f~~ has a local max at $x = -1$

IMPORTANT: Label any inflection points.



APPENDIX G

Tom's Solution to Problem 4 on Quiz 3

APPENDIX G

Tom's Solution to Problem 4 on Quiz 3

4. **Sketch the graph of a function that satisfies all of the following conditions:**

(a) f is concave up on (1,3)

(b) f is decreasing on $(6,\infty)$

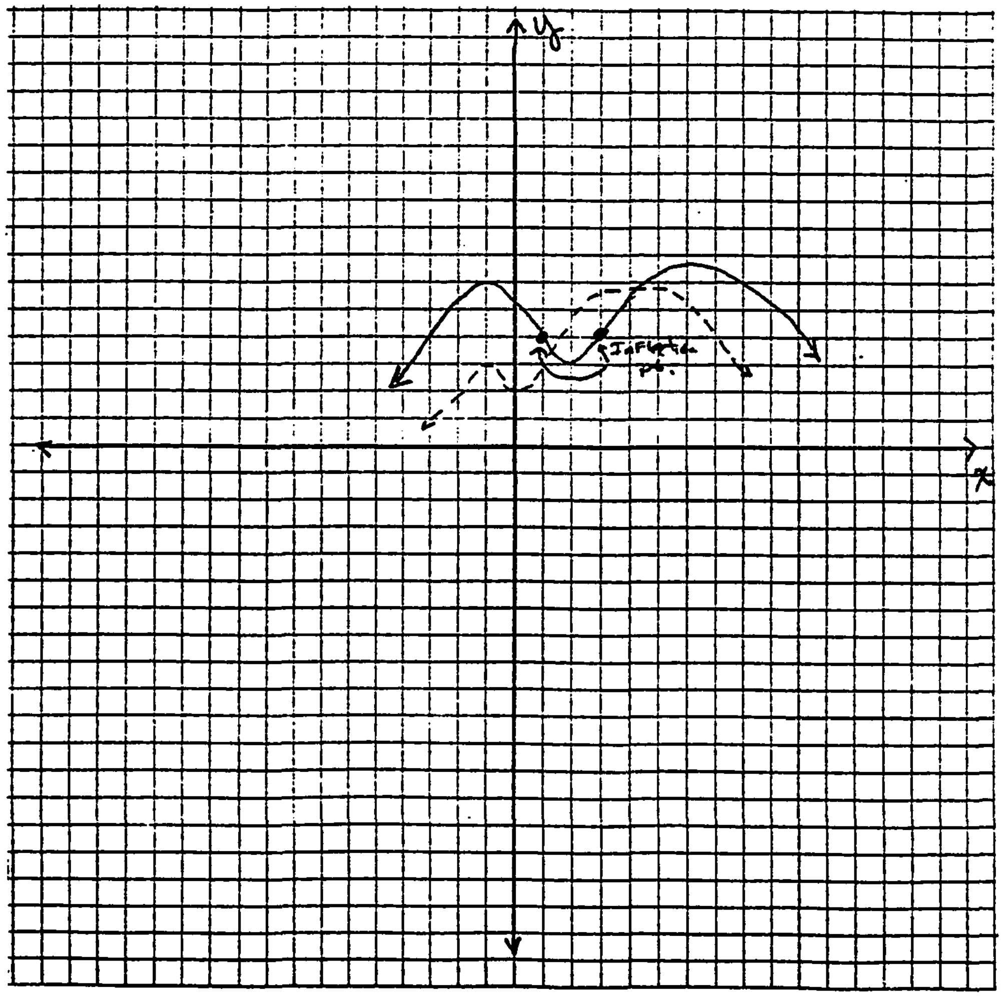
(c) f has a local min at $x = 2$

(d) f is increasing on $(-\infty,-1)$

(e) f is concave down on $(3,\infty)$

(f) f has a local max at $x = -1$

IMPORTANT: Label any inflection points.



APPENDIX H

Ben's Solution to Problem 4 on Quiz 3

APPENDIX H
Ben's Solution to Problem 4 on Quiz 3

4. Sketch the graph of a function that satisfies all of the following conditions:

(a) f is concave up on $(1,3)$

(b) f is decreasing on $(6,\infty)$

(c) f has a local min at $x = 2$

(d) f is increasing on $(-\infty,-1)$

(e) f is concave down on $(3,\infty)$

(f) f has a local max at $x = -1$

***IMPORTANT*: Label any inflection points.**