This is to certify that the

dissertation entitled

AN ANALYSIS OF STUDENTS' ERRORS IN MATHEMATICS
AT THE PRE-COLLEGE LEVEL

presented by

Allen Babugura

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Teacher Education

William M. Fitzgerald
Major professor

Date May 2, 1985

AN ANALYSIS OF STUDENTS' ERRORS IN MATHEMATICS AT THE PRE-COLLEGE LEVEL

By

Allen Babugura

A DISSERTATION

presented to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Teacher Education

1985

ABSTRACT

AN ANALYSIS OF STUDENTS' ERRORS IN MATHEMATICS AT THE PRE-COLLEGE LEVEL

By

Allen Babugura

The purpose of this study was to identify, classify, and analyze students' errors in mathematics at the pre-college level. The identification of errors was first done by administering an appropriate written mathematics test, composed of items designed to capture certain types of errors, to a sample of 146 students. Fifty-two different types of errors were identified, analyzed, and classified according to Radatz's (1979) error categorization. The methodology for this study was largely qualitative and exploratory, designed to enable the investigator to generate non-trivial assertions as well as to formulate some working hypotheses which subsequent studies may confirm or disconfirm. However, some quantitative analyses were applied to some of the research questions posed in the study.

The results of the study suggest that certain types of errors are committed by students markedly more frequently than others; that there is a non-trivial relationship between particular types of errors a student commits and his/her mathematical achievement; that students who score comparably on the test also exhibit comparable error patterns in mathematics; that the dependence of error patterns on gender, number of years of high school mathematics, and age is, respectively, not significant at the 0.05 level of significance; that individual

student error frequencies for algebra significantly correlate with those for arithmetic and geometry, but individual student error frequencies for arithmetic do <u>not</u> correlate significantly with those for geometry at the 0.01 level of significance; and that a non-trivial error pattern difference exists between the group that scored highest and the group that scored lowest on the test.

In addition, five major assertions and five working hypotheses were formulated, the former warranted by observable evidence from the data, and the latter motivated by some consistent observations from the data.

DEDICATION

To my dear parents,

ALFRED KIHUMA (deceased)

and

MELLANIA BARUNGI KIHUMA,

who taught me to wake up early and work

ACKNOWLEDGEMENTS

So many good friends have supported me as I labored with this study that an exhausting listing of their names would take prohibitively long. Consequently, I will only say to all those good friends: my thanks to you will forever know no bounds! My special gratitude must go to the members of my dissertation committee: Professors William Fitzgerald (chair), Perry Lanier, Donald Freeman, and Bruce Mitchell. To all of my committee members, I have only this to say: your remarkable abilities both to guide and to support me in my scholarly tasks defy the power of mere words for description. Thank you.

Last, but by no means least, I am dearly indebted to my wife Fidelis and our children Alex, Agnes, Doreen, Christine, and Celia, who bore with me in our physical separation as I pursued graduate studies.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

In the recent past, mathematics education all over the world has undergone some kind of curricular metamorphosis. To be sure, some countries have shown more radical changes than others in their school mathematics curriculum, but all countries have been motivated into these changes by a similar desire: to optimize the extent and use of mathematical knowledge the learner can acquire. That a pedagogically sound curriculum is a must prior to achieving this desire cannot be doubted. However, an effective use of a curriculum in the teaching-learning process is at least as important as the soundness of that curriculum in facilitating knowledge acquisition. Following are two narrative vignettes for two classroom instances the writer has recently witnessed as he taught. (These vignettes are intended to illustrate the pedagogical need for an effective use of curriculum.)

I. While teaching about solving quadratic equations by the method of "completing the square," I used the example $2x^2 + x - 6 = 0$ to illustrate the method we were going to use. We proceeded as follows:

$$
\begin{aligned}
& 2x^2 + x - 6 && = 0 && \\
\Longrightarrow\ & 2x^2 + x && = 6 && \text{(a)} \\
\Longrightarrow\ & x^2 + \frac{x}{2} && = 3 && \text{(b)} \\
\Longrightarrow\ & x^2 + \frac{x}{2} + (\tfrac{1}{4})^2 && = 3 + (\tfrac{1}{4})^2 && \text{(c)} \\
\Longrightarrow\ & (x + \tfrac{1}{4})^2 && = 3 + \frac{1}{16} = \frac{49}{16} && \text{(d)} \\
\Longrightarrow\ & x + \frac{1}{4} && = \pm\sqrt{\frac{49}{16}} = \pm\frac{7}{4} && \text{(e)} \\
\Longrightarrow\ & x && = -\frac{1}{4} \pm \frac{7}{4} && \\
\Longrightarrow\ & x = -2 \quad \text{or} \quad x = \frac{3}{2} && &&
\end{aligned}
$$

One student asked how we went from (b) to (c) above. Her question was, "Where did we get $\frac{1}{4}$ from?" I briefly tried to have the student recall the perfect square identity and hoped that her question would be answered once we saw the connection between the perfect square identity and our step (c) above. However, it was soon clear to me that I had to abandon that approach and adopt another one if I were to succeed in helping the student see "where we got $\frac{1}{4}$ from." I needed more time for this task than was available for that class, so I encouraged the student to meet with me at our convenience after the class. After class, I was able to help the student after getting at her real source of difficulty through a brief interview. At some point during the interview, she asked, "From step (a) to step (b), we divided through by 2. Why do we divide $\frac{1}{2}$ by 2 again to get $\frac{1}{4}$ in step (c)?" It was now clear to me that her difficulty had something to do with viewing one solution strategy as duplicating the role of a preceding solution strategy. In our case above, however, the step from (b) to (c) does not duplicate the step from (a) to (b), although the student thought it did. Just before I tried to help the student see that our two steps serve two distinct roles in the solution process, I gave her the example $3x^2 - 5x + 2 = 0$ to solve using the method of

completing the square. With this example, she had no difficulty in carrying out the solution process step by step up to the answer. This observation confirmed my earlier hunch that it was the 2 before the square term that caused her difficulty since it would force division by 2 from (a) to (b), and this strategy would appear to the student as duplicated by what goes on between (b) and (c) to obtain $\frac{1}{4}$ from $\frac{1}{2}$. It took me a very short time to help the student legitimize the source of $\frac{1}{4}$ in our original example after I had identified the real source of her difficulty.

II. While discussing "cash discounts," we used the following example:

> Alex purchases an item valued at $600.00 and makes a down payment of $300.00. Given that the cash discount rate for the item is 10%, how much does Alex still owe after making the down payment?

The class felt that there were at least two methods of solving this problem, each of which made sense to them; but each method yielded a different answer. We shall label these two methods A and B.

> Method A (the text method): Since there is a 10% discount, $ .90 will clear $1.00 of the bill. Thus, $300.00 cleared $\frac{\$300}{.90}$ of the bill, i.e., $333.33. Thus, Alex still owes $266.67.

> Method B (the class method): Alex should have been given 10% discount on the $300.00. But since he did not receive this $30.00 discount, he effectively cleared $330.00 of the bill. Thus, Alex still owes $270.00.

As a teacher of the class, it was not enough for me to stress the sense of Method A (which is the method that leads to the correct answer), but I also had to find where the flaw in Method B was and convince the students that a flaw really existed in the method.

(It may be in order to remark, here, that the cash discount terms that applied to our class problems had an important time element. The discount rate quoted applies only within a limited period, say 20 days or less after the date of completing the square. With this example, she had no difficulty in carrying out the solution process step by step up to the answer. This observation confirmed the order. Any amount of the bill still outstanding after the discount period will not be discounted on payment. Thus it would not be correct to say that the purchaser owes $540.00, i.e., 90% of $600.00 even before the down payment! In our example, this kind of reasoning would yield $240.00 as the answer. This is a wrong answer. Only cash paid within the discount period is thereby discounted, and the amount still owed is not discounted, since there is no proof that such an amount will be paid within the discount period.)

The two incidents above testify to the reality of some classroom situations, and the subject of this study is meant to stimulate some reflection on "how to make a teacher equal to such classroom tasks" as exemplified by, inter alia, those incidents.

Purpose of the Study

This study seeks to identify, classify, and analyze students' errors in mathematics at pre-college level. By examining student's work on mathematics items from an appropriate instrument, errors committed by the students (in the process of working out solutions to the items) will be identified. Each error observed will then be classified using error categories as enunciated by Radatz (1979). In case some errors are observed that do not belong to any of the categories proposed by Radatz, then more categories will be created to accommodate those errors. An analysis of the errors will then be performed in search of evidence that could confirm or disconfirm existing hypotheses about students' errors in mathematics. Any important observations that may be made during the analysis of errors, which may not appear to relate to any of the hypotheses that will have been found in the literature, will be duly noted.

Significance of the Study

Mathematics education has had a history fraught with concerns about diagnosis and remediation of mathematical learning difficulties. These difficulties have, for the most part, manifested themselves in mathematical errors committed by students. The study has its main significance in the realm of error diagnosis--a necessary prerequisite for error remediation. It must be emphasized, however, that the need for error remediation arises only in as far as it heals mathematical learning difficulties. Thus it becomes necessary to examine the errors carefully in order to establish the extent to which some of them cause difficulties in the learning of mathematics. Indeed, some research questions in this study are about the significance of effects upon mathematics achievement of given types of error. After the broad characterization of the significance of this study as given above, we find it convenient to discuss some specifics. We will treat these specifics under three subheadings: (a) knowledge about errors that cause difficulties in the learning of mathematics, (b) knowledge about errors that are useful in the development and the teaching-learning process of mathematics, and (c) the unique role of this study in mathematics education.

Knowledge about Errors That Cause Difficulties in the Learning of Mathematics

Teachers of mathematics are likely to find studies dealing with students' mathematical errors helpful in stimulating diagnostic abilities as well as thoughts about remediation. In this respect, West (1971) maintains that "there is hardly a skill in the teacher's repertoire that is more important than the ability to identify pupil errors and to prescribe appropriate remedial procedures." Thus teachers of mathematics need to involve themselves in serious exercises of error identification and remediation as they teach, and the success of these exercises

Significance of the Study

Mathematics education has had a history fraught with concerns about diagnosis and remediation of mathematical learning difficulties. These difficulties have, for the most part, manifested themselves in mathematical will be greatly assisted by research studies in these areas. The two classroom incidents I have quoted in the introduction to this chapter exemplify some of the errors that cause difficulties in the learning of mathematics. My initial response to the students' concern did miss the mark as far as helping the student was concerned, and this happened because my initial assumption about the nature of her difficulty was not correct. Regrettably, most classroom student-teacher interactions do terminate prematurely--often never giving the students the chance necessary to communicate their problems accurately to the teacher. Thus, very often, the teacher gives irrelevant responses to a great many of the students' questions if only because the teacher does not, in the first place, get at the students' real difficulties before responding. In my case, the interview I had with my student subsequent to the class session was both very revealing to me about what nature students' difficulties can assume and ultimately very helpful to the student as I was able to help her understand the solution to the problem.

The cash discount problem created a rewarding classroom experience both by capturing an error that most of the class committed and providing a lively classroom discussion of a meaningful application of arithmetic. Hence the error the class committed in this problem both caused learning difficulties and aided the learning of mathematics. Finding a flaw in Method B of solving the problem did not produce nearly as much difficulty for me as did convincing the students that such was a flaw. However, the whole exercise of identifying and verifying the flaw led us to modify Method B of solution to obtain the correct answer--an undertaking that enriched our understanding of cash discounts beyond a solution such as Method A yielded.

Lest an impression be created that errors that cause difficulties in learning originate from only student behavior, teachers would do well to examine the important ways in which students' learning difficulties are enhanced by what will be greatly assisted by classroom studies of these issues. The two classroom incidents I have recounted in the introduction to this chapter represent some of the errors that teachers fail to do properly. One question we can ask is, "In posing mathematical problems for students to solve, how clearly do teachers communicate?" While pondering this question, we may as well look at one of the exercises a teacher gave to her students as a "holiday cheer." Hall (1980) asked her students to sketch a graph for the set which she represented as $\{(x, y) \mid y = |x| \; ; \; x = 0, y \leq 0\}$. In 1985, I read this exercise and tried to sketch the graph for the set she described. After the description $y = |x|$, I was somewhat disappointed with the additional constraints of $x = 0$ and $y \leq 0$. Thus my graph could only consist of one point, namely (0 , 0) in the Cartesian plane. But what answer did the teacher expect from her students? Her answer was the 'Y' shape! I am not nearly so worried about what her students failed to learn by <u>not</u> getting her "correct" answer as I am worried about what the students who got her "correct" answer might have learned notationally. Incidentally, there were several exercises which were given to the students which, when done according to the set language description given, yielded different sketches from the answers the teacher gave. I believe that getting into the habit of error diagnosis has one important payoff--it helps one to examine one's own work for errors as well.

Knowledge about Errors That Are Useful in the Development and the Teaching-Learning Process of Mathematics

Errors committed by learners have been socially stigmatized to the extent that hardly any error can be perceived as useful by most students and teachers alike. However, a more rational view of errors can be afforded when they are carefully studied and characterized. Brownwell and Hendrickson (1950) had this to say, "Errors in the course of concept formation provide fruitful opportunities for constructive teaching." I have already narrated the fruitfulness of the class error on the cash discount problem in aiding better understanding of the cash teachers fail to do presently. One question we can ask is, "In posing mathematical problems for students to solve, how clearly do teachers communicate?"

discount concept. Teel (1978) remarked, "Throughout history, some of the world's greatest discoveries have been due to what was first thought to be an error. Our thanks go to all those who have not quit when they 'thought' they were wrong." (Teel was commenting on a discovery of a new theorem that some classroom teacher had reported as having arisen from what had appeared to be an error.) Studies such as this are likely to foster among all those interested in teaching and learning an understanding that not all student errors are detrimental to the learning process. It will not be difficult to agree with this view once we realize how often we have effectively learned through some mistakes.

The Unique Role of This Study in Mathematics Education

As one reviews the literature, it becomes evident that most of the studies that have so far been done in the areas of error diagnosis and remediation have dealt with elementary school pupils on the one hand and with arithmetic on the other. However, substantial research has indicated that students form very important attitudes about mathematics during high school. Also, it is common knowledge that geometry and algebra are introduced to pupils as early as elementary school and that interesting errors occur in geometry and algebra as they occur in arithmetic. This study addresses errors in mathematics, not limited to arithmetic only, committed by some eleventh grade students and some college students who have yet to take college level mathematics. The uniqueness of this study is due to the facts that error diagnosis has not been limited to arithmetic only and that older students have been used as subjects in the study.

discount concept. Teel (1978) remarked, "Throughout history, some of the world's greatest discoveries have been due to what was first thought to be mistakes. Our thanks go

Summary

During the teaching-learning process of mathematics, situations arise where the teacher must get at the real source of the students' learning difficulty before the latter can learn from the former. Thus it becomes necessary for teachers to have the ability of properly diagnosing students' difficulties for effective teaching to occur. This study addresses the identification, classification, and analysis of students' errors in mathematics at the pre-college level for the main purpose of assisting mathematics' learning difficulty remediation. The major significance of this study is, therefore, the diagnosis of students' errors in order to inform appropriate remediation strategies. (This study does not address the remediation process itself, but only deals with a necessary prerequisite to remediation--error diagnosis.) More specifically, the study set out to

1. contribute to knowledge about errors that cause difficulties in the learning of mathematics,

2. contribute to knowledge about errors that are useful in the development and the teaching-learning process of mathematics, and

3. serve a unique role in mathematics education by dealing with students' errors in pre-college mathematics instead of just dealing with arithmetic errors at the elementary school level as the majority of previous studies in this area have done.

Summary

During the teaching-learning process, mathematical situations arise where the teacher must get at the real source of the student's [illegible] before the latter can learn from the former. Thus it becomes necessary for teachers to have the ability of [illegible] diagnosing students' difficulties [illegible] effectively [illegible]

CHAPTER II

LITERATURE REVIEW

Introduction

A review of relevant literature reveals that error analysis in mathematics education has been of much interest internationally for over 80 years now. Many studies have been done on error diagnosis and remediation; some of the studies have focused on error classification and possible causes while other studies have focused on some particular characteristics of errors--the persistence of errors, for example, is one of the most quoted characteristic of errors in the literature. Emerging from the current state of knowledge about students' errors in mathematics and their implications for the learning of mathematics is a strong need to learn more about the nature of these errors in order to better determine the extent to which they affect the acquisition of mathematical knowledge.

Historical Survey

Radatz (1979) gives a brief historical survey of error analysis in mathematics education. He gives an account of studies by German, Russian, and United States' scholars about errors in mathematics. His "Overview of Accessible Publications on Errors in Mathematical Education" shows German and Anglo-American publications spanning the period between 1904 and 1979. From his analysis of the accessible literature, Radatz maintains that "errors in the non-arithmetical content areas of mathematical education" form one of the several research levels that have not yet received any substantial attention. It is

hoped that this study will contribute to alleviating his concern about research deficits he outlines in his account.

Buswell and Judd (1925) report 31 diagnostic studies in arithmetic between 1909 and 1924. Of these, 20 are devoted to an analysis of errors in arithmetic. The subjects in these studies are pupils from elementary schools. Qualitative information was largely the yield from the research--potential error techniques were listed, the frequency distribution of these error techniques across age groups was determined, and the persistence of individual error techniques was determined. In one of the earliest studies on sex differences, Smith (1895) observed that sex differences in mathematics achievement is negligible, but boys seem to make more gains as time goes on. Although his study does not specifically address errors in mathematics, one finds some revealing observations about difficulties encountered in arithmetic by pupils as viewed across gender.

Error Persistence

Any teacher of mathematics will most likely be amazed by how often an error is repeated by a subject even after numerous interventions by the teacher to correct the error. That errors persist has been found by many researchers (for example, Myeres, 1924, and Radatz, 1979). We hold the view that the persistence of errors in mathematics makes an analysis of errors a natural curiosity since one is bound to wonder why an error is often so hard to eliminate. An attempt to gain insight into this phenomenon has resulted in deeper studies about the nature of errors as well as their possible causes. For example, Habel (1958) had this to say:

> Intensive studies of errors which students make and why they make them show that, year in and year out, there are different students passing through classes making the same types of errors and often the same errors are made in different grades by the same students. The students may or may not change. The errors change not. (p. 81).

hoped that this study will contribute to alleviating his concern about research deficits he outlines in his account.

Buswell and Judd (1925) report 31 diagnostic studies in arithmetic between 1909 and 1924. Of these, 20 are devoted to an analysis of errors in arithmetic. The subjects in these studies are pupils from elementary schools. Qualitative information [illegible]

Some reflection on the quotation above is in order. The quote contains two important assertions, namely that, across the years, (1) different students pass through classes making the same types of errors and (2) the same students often make the same errors in different grades. Reflecting on the implications of the first assertion, one gets the impression that there exists a pool of errors in the educational path of students and that students acquire errors from this common pool as they progress through school. If this impression is true, then the diagnostic studies should be designed to discover the location of that pool of errors and remediation efforts should be aimed at draining that pool of those errors that cause mathematical learning difficulties. The implications of the second assertion point to the persistence of errors within an individual student across years and grades. The truth of the second assertion suggests that studies should be designed to accurately identify those errors that persist and remediation strategies should be developed to be used against the persistent errors that hinder the learning of mathematics until those errors are eradicated. For the purposes of research design, it is important to note that the two assertions above are not mutually exclusive.

While still on the persistence of errors, the writer wish to mention that his own interest in this study derives from his observation that an error like $\sqrt{a^2 + b^2} = a + b$ has been notoriously persistent among many of his pre-college algebra students. This same error is listed by Laursen (1978) among other errors in first year algebra.

Error Stigmatization

Not all mathematical errors are detrimental to the positive development and learning of mathematics. Yet nearly every error has been conceived by our larger society as undesirable in the learning process. To counter this stigma

Some reflection on the quotation above is in order. The quote contains two important assertions. [illegible] across the years; (b) [illegible] students pass through [illegible] against all errors, various positive views have been contributed to literature. Schwarzenberger (1984) stressed the importance of mistakes in these mathematical activities.

1. In the historical development of mathematics, mathematical mistakes that have been particularly useful are quoted.

2. In the learning of mathematics, mathematical mistakes that serve as an aid to understanding the subject are documented.

3. In the teaching of mathematics, students' mistakes that serve as a diagnostic tool are mentioned.

Schwarzenberger believes that "the important thing is not to eradicate mistakes but to learn from them, not to avoid making them but to admit them" (p. 169).

Brownwell and Hendrickson (1950), West (1971), and Bainbridge (1981) are some of the scholars who express the view that some errors are useful. The phrase "trial and error" is all too familiar to warrant much explanation when communicating with a learner. Numerous instances can be recounted in any branch of mathematics where trial and error techniques form the only natural way to begin exploring subsequent results. The well-known procedures of "proof by induction" and "polynomial factoring" will serve as examples where trial and error techniques are almost indispensable. In spite of the natural appeal trial and error affords many thinkers, many mathematicians have been known to show dissatisfaction with any method that relied substantially on trial and error techniques to derive a mathematical result. It is well known that many mathematicians would not appeal to proof by induction if another method of proof existed for the same theorem. Even when no alternative method has been found, some mathematicians are known to have very little respect for results that follow by induction. Of course, there are bound to be other reasons why many mathematicians hold the method of proof by induction in contempt, but it is not unreasonable to assume that one of these reasons may often be the trial

and error aspect of the first induction step. This may be yet another instance where the stigma on all errors makes the "trial and error" procedure unattractive to some.

Dramatic Uses of Errors

Even more dramatic uses for errors in mathematics show up every day, both in classrooms and in private study rooms. For example, Rostad (1971) communicated an interesting classroom incident where an error in the method of solving a problem led to a discovery of a mathematics theorem. This discovery may not have been mathematically profound, but it certainly was an opportunity to get the class involved in some exciting mathematical activity. Then Tubridy (1978) communicated another result his class discovered after what appeared to be a student's error of solving a problem in analytic geometry. Students' "errors" of this type provide a unique learning experience whereby the teacher joins the students to form a group of mathematical explorers.

Implications for Mathematics Teaching

There are some examples in the literature that serve to illustrate how careful mathematics teachers need to be when responding to students' answers which appear to be in error. Some of them are indeed in error; some of them are not in error though they appear to be. Some of them are correct but follow from erroneous operations. Yet others are incorrect but follow from methods which make perfect sense to the students at their cognitive level.

A common error in the simplification of algebraic expressions in a first algebra course may be illustrated thus: when asked to simplify an expression like $\frac{x + xy}{x}$, many students will give answers like $1 + xy$ or $x + y$. A remediation strategy that a teacher would adopt to correct this error will, of course, depend on what is causing this error in the work of a particular student. I find, almost and error aspect of the first induction step. This may be yet another instance where the stigma on all errors makes the "trial and error" procedure unattractive to some.

invariably, that an interview with the student is necessary for the teacher to get at the real source of the student's difficulty in these kinds of problems. Only then can the teacher hope to help the student out of the difficulty. Hart (1978) provided another example of a consistent and common error in ratio problems. She called it an "addition strategy." When students were asked to find the altitude of triangle PQR whose base QR was, say, 12 and which was similar to triangle ABC with base BC, say, 10 and altitude 14, many of them obtained the answer 16. Their addition strategy worked like this: in similar triangles PQR and ABC, base BC has "grown" into base QR from 10 to 12--an increase of two units. Therefore, the altitude will "grow" from 14 to 16--the same two units' increase. This kind of strategy makes sense to students until they are carried to a level deeper in the similarity concept. Hart tried to account for the reasons for errors of this nature in the following:

> The message one receives from these errors and their high incidence is that there is a considerable amount of confusion in the children's minds as to what they are doing in mathematics. Some errors will be symptomatic of the child being unable to grasp the level of abstraction being presented; others might arise because we never consolidated the teaching. (p. 39)

Concerning children's strategies, Booth (1981) maintained that teachers "must understand children's strategies in order to aid their replacement with sophisticated ones" that will lead to correct solutions. Kent (1978, 1979) gave a very elaborate descriptive account of a process by which he gained deep insights into the nature of his students' mathematical mistakes. The following is the way he saw mathematical mistakes:

> There has been a tradition in mathematics teaching to regard mistakes simply as things to be corrected. In this paper the viewpoint is replaced by one which suggests that mistakes are a source of learning about the thought process of others. (p. 27)

In order to learn about the thought processes of students, one practical method the teacher can use is to interview the students. Clearly, this points to the invariably, that an interview with the student is necessary for the teacher to get at the real source of the student's difficulty in these kinds of problems. Only then can the teacher hope to help the student out of the difficulty. Hart (1978) provided another example of a consistent and common error in ratio problems. She called it [illegible] strategy. When students were asked to find the [illegible] incorporation of individualized instruction into any other instructional techniques in the teaching of mathematics.

What Carman (1971) called "mathematical misteaks" are not only intellectually ingtriguing, but they also serve to illustrate how one can unexpectedly go wrong in error remediation by counter example. We will illustrate. Suppose the following happens in an algebra class: The teacher writes the expression $\frac{x^3 + y^3}{x^3 + z^3}$ and asks the class to simplify it. One student responds with $\frac{x^3 + y^3}{x^3 + z^3} = \frac{x + y}{x + z}$. The teacher then asks the student to work out $\frac{37^3 + 13^3}{37^3 + 24^3}$ as well as $\frac{37 + 13}{37 + 24}$ and compare the results. If the teacher's intention is to show the students that the response to $\frac{x^3 + y^3}{x^3 + z^3} = \frac{x + y}{x + z}$ is wrong, then the numerical example the teacher has chosen is unfortunate since, in fact, $\frac{37^3 + 13^3}{37^3 + 24^3} = \frac{37 + 13}{37 + 24}$. To be sure, happenings of this nature are rare in a classroom; but when they occasion, a great learning advantage could be taken of them. From the example above, the teacher and the class may be motivated to discover the result that $\frac{x^3 + y^3}{x^3 + z^3} = \frac{x + y}{x + z}$ if and only if $z = x - y$.

The above numerical example is one of the many rather surprising results that follow from an incorrect operation but are, themselves, correct. Carman called them "mathematical misteaks" and distinguished them from "mathematical mistakes."

How might many teachers familiar with mathematical misteaks react to the following student's work? $\frac{14}{45} \div \frac{7}{15} = \frac{14 \div 7}{45 \div 15} = \frac{2}{3}$. It is not unreasonable to expect that many teachers familiar with the notion of mathematical misteaks would describe the student's work in the example above as containing a mathematical misteak. This is because an unusual division procedure for fractions has been used by the student; and although it is a correct procedure incorporation of individualized instruction into any other instructional techniques in the teaching of mathematics.

What Carman (1971) called "mathematical mistakes" are not only intellectually intriguing, but they also serve to illustrate how one can unexpectedly go wrong in error remediation by counter example. We will leading to a correct answer, many teachers may say that the student got the correct answer only for a special case. O'Donnel (1980) communicated this idea on fraction division and pointed out that $\frac{a}{b} \div \frac{c}{c} = \frac{a \div c}{b \div d}$ for all fractions $\frac{a}{b}$ and $\frac{c}{d}$. The proof for this follows after very straight forward fraction manipulations.

Sex-Related Differences

There exists a considerable amount of literature about sex-related differences in mathematics achievement. Since mathematics achievement by students has very often been measured in terms of performance on mathematics tests, observations of sex-related differences in mathematics achievement have often incorporated observations of sex-related difficulties in the learning of mathematics. Benbow and Stanley (1980) reported that they observed "a large sex difference in mathematical ability in favor of boys." However, Fennema and Sherman (1977) reported that their data did not support "either the expectations that males are invariably superior in mathematics achievement and spatial visualization or the idea that differences in sexes increase with age and/or mathematics difficulty." But Hilton and Berglund (1974) did report that the differences (in mathematics achievement) between the sexes increased with age in favor of males, although they failed to observe any sex differences in mathematics at grade five level. Then Dallas and Alexander (1983) maintained that differences between sexes in mathematics achievement largely emerge during high school. And Marshall (1981) maintained the following:

> Sex differences are present in children's selection of multiple choice responses, and the types of errors committed can be identified. It has been demonstrated that the distributions of boys' and girls' errors differ. (p.)

It is evident from the literature that research findings about sex differences in mathematics achievement conflict. But there seems to be some consistency across studies that sex differences are not significant until after

grade six. Moreover, many of the sex-related differences reported in the studies seem to be consistent across countries. Badger (1981) talks about the report by the International Association for the Evaluation of Educational Achievement (IEA) which surveyed mathematical attainment in 12 countries in 1964. This report bears some similar findings across the countries. For example, within each country, girls were found to have more difficulties than boys on both computation and verbal problems.

Summary

Results from many studies on students' errors in mathematics suggest a strong need to learn more about the nature of these errors in order to better understand their effect upon the acquisition of mathematical knowledge. A brief historical survey of error analysis in mathematics education shows that many scholars from countries such as Germany, Russia, and the United States have carried out studies on students' errors for more than 80 years to date. However diverse the research interests of these scholars appear to be, their findings play a unified role of informing mathematical pedagogy.

Perhaps the one characteristic of errors that is most consistently reported across studies is error persistence. Some important longitudinal studies have been done to study the persistence of errors, and the types of errors that persist have been documented. Many teachers and students have come to view every error in mathematics as detrimental to mathematics learning. However, a substantial amount of literature exists which counters the stigma associated with errors that are, in fact, useful. Some errors can be used in class very effectively to lead to discoveries of interesting mathematical results. Reports also exist in the literature about some mathematically intriguing error techniques that lead to a correct result. These error techniques are described as "mathematical grade six. Moreover, many of the sex-related differences reported in the studies seem to be consistent [illegible] (1981) talks about the report by the Internal [illegible] misteaks" by at least one scholar. Since many errors students commit follow from strategies that make sense to the children at their cognitive level, it is important that teachers understand their student's strategies before the former can hope to effectively help the latter learn. One instructional technique that could be employed to help teachers understand their students' strategies is individualization.

Many studies in mathematics education have been done on sex-related differences in mathematics achievement. Findings from these studies conflict, but many of them report students' difficulties which seem to show sex-related differences. Moreover, some of these sex-related differences seem not to be country-specific since they are reported in studies on subjects from different countries.

mistakes" by at least one scholar. Since many errors students commit follow from strategies that make sense to the children at their cognitive level, it is important that teachers understand their students' strategies before the former can hope to effectively help the latter learn. One instructional technique that could be employed to help teachers understand their students' strategies is individualizat

CHAPTER III

RESEARCH QUESTIONS AND METHODOLOGY

Introduction

The last two chapters have focused on the purpose and significance of this study as well as the current state of knowledge about students' errors in mathematics as found in the literature. This chapter will address the research questions and methodology under the following subheadings:

1. specific delineation of research questions;
2. discussion of the variables involved;
3. design and proposed method of analysis;
4. population definition and sample specification;
5. instrument, delimitation, and limitations of the study.

Specific Delineation of Research Questions

From reading the literature about students' errors in mathematics as well as the writer's experience, as a teacher of mathematics, concerning students' mathematics learning behavior, the following research questions arise.

1. What error types seem to be most frequent among students?
2. Is there a relationship between particular types of errors a student commits and his/her mathematical achievement as measured by test scores?
3. Do students who score comparably on the test also exhibit comparable error patterns in mathematics?
4. Do error patterns vary significantly across (a) gender? (b) number of years of high school mathematics? (c) age?

5. Do individual student error frequencies correlate across the three content areas of mathematics: arithmetic, algebra, and geometry?

6. Is there a significant error patterns difference between the group that scores highest and the group that scores lowest?

Following is a brief discussion of the importance of each question.

Question one asks about error types that seem to show up most frequently among students. Knowledge of such errors, if they exist, will enable one to narrow down one's field of investigation to those errors which most students seem to commit in order to find out why such errors are widespread. Strategies of remediation that are most effective are most likely to be devised once a sound knowledge base has been acquired about such target errors.

Question two asks whether a relationship exists between particular types of errors and the mathematical achievement, as measured by test scores, of the students who commit them. Research attempts to answer this question will most likely inform the decision making process of mathematics educators about priority areas for remediation efforts. Also, should the results of the study indicate a substantive relationship, then such knowledge could be put to use in predicting students' mathematics achievement potentials from the kinds of errors they commit.

Question three seeks to unravel those error patterns, if they exist, that are shared by students in the same achievement groups. The knowledge gained from research attempts to answer this question will most likely be useful in determining degrees to which certain error patterns are detrimental to mathematical achievement.

Question four is trying to get at information that will confirm or disconfirm existing theories about students' mathematical errors. Scholars like Marshall (1981) have reported sex-related differences in error patterns. Radatz

(1979) has documented the persistence of errors across years of mathematics learning. And Kent (1978, 1979) has reported no substantive differences in error patterns across age groups.

Question five invites a search for individual student error frequency correlations across the three content areas of mathematics, namely arithmetic, algebra, and geometry. Findings from research attempts to answer this question are likely to aid our understanding of how difficulties in one content area in mathematics are predictive of difficulties in another content area. This understanding is essential for mathematics educators in order to carry out appropriate curriculum planning.

Question six asks for an error pattern comparison for two extreme groups on the achievement scale. If significant differences are found, then one may tentatively hypothesize that error patterns displayed by the lowest achievement group are more detrimental to mathematical achievement than error patterns exhibited by the highest achievement group. Then studies may be designed to confirm or disconfirm such a hypothesis. If errors that are detrimental to mathematical achievement could be validly identified, then mathematics education would greatly benefit from such information since it would then be possible to effectively prioritize remediation efforts.

Discussion of the Variables Involved

The principal task in this study is an analysis of students' errors at pre-college level. For the purposes of handling this task meaningfully, several important research questions have been formulated and an instrument has been duly designed for use in collecting data pertinent to the questions. The data collected consist of measurements and other kinds of observations which have been deemed appropriate for providing answers to the research questions posed.

(1979) has documented the persistence of errors across years of mathematics learning. And Kent (1978, 1979) has reported no substantive differences in error patterns across age groups.

Question five invites a search for individual student error frequency correlations across the three content areas of mathematics, namely arithmetic, algebra, and geometry [illegible]

We will now list the variables that have been measured or otherwise observed for data necessary for this study.

Type of Error

Each subject in the study has responded to mathematical items on the instrument. As one would expect, the kind of error committed during the process of solving each of the problems on the instrument is bound to vary from subject to subject. It is this variation of type of error from subject to subject that provides most of the substance for the error analysis of this study.

Mathematical Achievement

The work of each subject in the study has been scored, and a total percent score for each subject has been taken as a measure of his/her mathematical achievement. In this study, mathematical achievement is an important variable as it applies to research questions two, three, and six directly.

Error Pattern

A group of subjects is said to exhibit a similar error pattern if it is evident they have committed similar types of errors on the instrument. Error pattern is also an important variable in this study as it applies to research questions three, four, and six directly.

Individual Student Error Frequency

Each subject has been associated with a whole number which stands for the total number of errors that subject has committed on a particular content area of mathematics on the instrument. Since there are three content areas represented on the instrument (arithmetic, algebra, and geometry), three whole numbers will, accordingly, be associated with each subject. This trio of whole numbers will be called an error frequency vector. Thus, with each individual

We will now list the variables that have been measured or otherwise [illegible] the data necessary for [illegible] study. [illegible]

Type of Error

Each subject in the study has [illegible] to mathematical [illegible] of the instrument. [illegible]

process [illegible]

subjects [illegible]

[illegible]

subject will be associated a three-tuple of whole numbers, each entry of which will represent the total number of errors that subject has committed on a particular content area of mathematics. For example, a subject who has committed five errors in arithmetic, four errors in algebra, and six errors in geometry will be associated with the error frequency vector (5,4,6). Individual student error frequency is also an important variable for this study since research question five asks about the correlation of individual error frequencies across arithmetic, algebra, and geometry.

Gender

The subjects have been classified as male or female. Since sex-related differences in error patterns are to be examined under question four of this study, the variable gender is pertinent to the study.

Number of Years of High School Mathematics

Information about how much mathematics each subject has had at high school level is needed so we can examine a relationship, if it exists, between error patterns and the number of years of high school mathematics that have been taken. Research question four asks about this relationship.

Age

The age of each subject has been recorded so that use can be made of this information in determining whether a relationship exists between how old one is and the kinds of mathematical errors one commits. Research question four asks about this relationship.

Arithmetic

Most research currently available about students' errors in mathematics education has been done on arithmetical errors. In this study arithmetic is one of the mathematics content areas we are examining for students' errors.

Algebra

Some studies, comparatively fewer than those concerned with arithmetic, have been done about students' errors in algebra. This study has incorporated algebra as one of the three content areas of mathematics upon which students' errors are to be analyzed.

Geometry

Of the three mathematics content areas in this study, geometry seems to be the least frequently studied for students' errors. This study seeks to examine some students' work on geometric items for errors in this area of mathematics.

The 10 variables listed above just happen to be the main ones featured in the research questions for this study. They are, by no means, the only ones that could be of educational interest for this kind of study. The limitation for scope of any manageable study has applied to this study, and the above 10 variables have been of main consideration.

Design and Proposed Method of Analysis

This study is exploratory and largely descriptive in design. Following is the rationale behind this methodology. Any research effort going into this study is geared to obtaining the best possible answer to any of the research questions posed at the beginning of this chapter. Research question one, for example, appears to be asking for the most basic information any teacher of mathematics will easily observe over several years of teaching. However, once one reflects

Arithmetic

Most research currently available about students' errors in mathematics education has

of the

upon the best one could do to obtain a good answer to the question, one realizes that merely having taught mathematics for years is not enough to supply one with data needed to answer the question satisfactorily. In order to single out, with some confidence, a particular type of error in mathematics as the one most frequently committed by eleventh grade students, for example, one needs to design a study to explore the types of mathematical errors exhibited by some eleventh grade students as well as the frequency of occurrence for each type of error. A good way to begin this exploration would be to administer a mathematical test to some sample of eleventh grade students and examine the students' work on this test for errors committed as well as the frequency of occurrence for each type of error. The value one attaches to the information obtained from such a study will, of course, depend on the validity of the findings as measured against the original objectives for the study. In our case, we are after the kind of information about students' errors in mathematics that will help the development of testable hypotheses. To this end, only questions have been posed and no hypotheses for testing have been formulated. An exploratory and largely descriptive study has been considered adequate for this purpose. Subsequent studies can later be designed to test any hypotheses that seem to be implied by findings from studies such as ours. What now follows is a question by question proposed method of analysis for data pertinent to that question.

Question One

What error types seem to be most frequent among students? In order to answer this question, we propose to have a sample of students taking mathematics at the pre-college level respond to mathematical items on an instrument designed to capture certain error types. Each script will then be examined critically for errors committed. Each error observed will be recorded upon the best one could do to obtain a good answer to the question, one realizes that merely having taught mathematics for years is not enough to supply one with data needed to answer the question satisfactorily. In order to single out, with some confidence, a particular type of error in mathematics as the one most frequently committed by eleventh grade students, for example, one needs to as soon after the observation as possible. This precaution is necessary mainly to avoid the possible loss of observable errors due to forgetfulness. A tally sheet will be developed showing error type codes as well as their frequency of occurrence taken over the number of subjects for this study. For the purposes of data for other questions, each examined script will be marked with error type codes as and where the errors occur. Although the instrument has been designed to capture some expected types of errors, there is plenty of room for the occurrence of interesting errors that may not have been originally expected. Thus, no original listing of error types will be done on the tally sheet--rather, the tally sheet entries together with their codes will be developed during the process of examining individual scripts. The following diagram illustrates how the tally sheet might look after examining, say, 30 scripts.

Error Type	Tally	Frequency
A	~~IIII~~ I	
B	III	
C	~~IIII~~ II	
D	II	
E	~~IIII~~	
.		
.		
.		
.		
.		

Since more scripts are yet to be examined, no entry has appeared in the frequency column yet. Complete information about any entry in the frequency column will be available only after every script has been examined. Also the error type column is still developing from top to bottom. Since more error types may show up as the examining process proceeds, there may be more entries yet

to go into this column. After the script examination is completed, the tally sheet will then be used to make a summary table where only the error type column and the frequency column will be used, together with their entries, to present an answer to question one.

Question Two

Is there a relationship between particular types of errors a student commits and his/her mathematical achievement as measured by test scores?

This research question will be answered by first identifying several errors that are most frequently committed, then examining the frequency distributions of these errors across quartiles on the achievement scale. For example, suppose error type B is found to be one of the most frequently occurring. The subjects will have been divided into four equal groups, each group representing a quartile on the achievement scale. (At most three subjects chosen at random will be dropped from the total number of subjects in this analysis in case to the total number is not divisible by four.) Then the number of subjects who commit error type B in each quartile will be recorded. This process will be repeated for other error types, and a table such as the one below will thus be generated.

Table 3.1
Error Type Frequency Versus Achievement Quartiles

	Frequencies				
Error Type	In First Quartile	In Second Quartile	In Third Quartile	In Fourth Quartile	Total
A	35	30	31	20	116
B	43	41	12	4	100
C	37	23	16	14	90
D	21	18	17	6	62
E	18	18	12	10	58
F	20	12	2	0	34

to go into this column. After the script examination is completed, the tally sheet will then be used to make a summary table where only the error type column and the frequency column will be used, together with their entries, to present an answer to question one.

<u>Question Two</u>

Is there [illegible]

con[illegible]

The first quartile will contain the bottom 25% on the achievement scale. In the above table, one can observe a lot of information relating achievement with error type. For example, error types B and F dramatically distinguish achievement groups, while error types A and E are not so distinguishing of achievement groups.

Question Three

Do students who score comparably on the test also exhibit comparable error patterns in mathematics? In answering this question, use will be made of error categories such as those described below. Radatz (1979) proposed a five-category classification of student errors in mathematics as follows:

1. errors in semantics (meaning, translation),
2. errors in spatial visualization (scale, shape confusion),
3. errors in mastery (operations),
4. errors in association (number of patterns, key words, formulas), and
5. errors characterized by use of irrelevant rules.

It is anticipated that more categories will be created to accommodate those errors that cannot meaningfully fit any of Radatz's five categories.

For question three, comparability in achievement will be measured by score classes of width not more than five. A group of subjects falling in the same class will be said to be comparable in achievement. Comparability in error types will be qualitatively assessed by looking at how errors committed by subjects in the same achievement class are distributed across error categories.

Question Four

Do error patterns vary significantly across (a) gender? (b) number of years of high school mathematics? (c) age? For this question, error category by gender, error category by number of high school mathematics years, and error

category by age contingency tables will be made and three chi-square tests run to determine the dependence of error patterns on gender, number of years of high school mathematics, and age.

Question Five

Do individual error frequencies correlate across the three content areas of mathematics, i.e., arithmetic, algebra, and geometry? This question will be answered by correlating individual error frequencies in arithmetic with those in algebra, individual error frequencies in arithmetic with those in geometry and individual error frequencies in algebra with those in geometry. The product-moment correlation coefficient $\underline{r}$ will be used for each of the three comparisons. Significance tests for correlations will then be run.

Question Six

Is there a significant error pattern difference between the group that scores highest and the group that scores lowest? In answering this question, the first quartile (on the achievement scale) will be compared with the fourth quartile for differences in error patterns. Qualitative error pattern assessment will be done in order to determine error pattern differences between the two groups of subjects.

Level of Significance

Statistical significance tests will be used in search of answers to questions four and five. For question four, a level of significance $\alpha = 0.05$ will be used. However, for question five, a higher level of significance will be required since, with the number of subjects as large as is expected (between 100 and 160), problems whereby statistical significance may not correspond with practical (meaningful) significance might arise. Burroughs (1975) discussed this kind of

problem in statistical significance tests for correlations. For this study, we will use a level of significance $\alpha = 0.01$ in question five.

Use of Quantitative and Qualitative Methods of Analysis

The research methodology employed in this study is mainly qualitative, although various aspects of quantitative statistical analysis are used to answer some of the research questions. Although a variety of comparisons among variables are made, the nominal nature of such a variable as "type of error" will render inappropriate most of the parametric statistical methods of analysis. However, where quantitative statistical analysis seems to be the most appropriate for valid conclusions, such statistical analysis is done.

Population Definition and Sample Specification

The population for this study is eleventh grade students in Michigan and Michigan State University freshmen taking pre-calculus mathematics or a first course in calculus during the spring term of 1985. The sample consists of 95 eleventh grade students from two Michigan high schools and 51 Michigan State University freshmen taking a pre-calculus mathematics course or a first course in calculus during the spring term of 1985. The rationale behind selection of this sample follows.

Most eleventh grade students in Michigan and Michigan State University freshmen have learned all the material on the instrument. The often observed persistence of errors in mathematics among students justifies our assumption that most errors committed by the subjects of this study are the same kinds of errors these subjects were committing in their earlier years of high school. So, with respect to these subjects, their being in the eleventh grade or in college has not much changed the types of errors they have been committing in high school.

Moreover, according to Habel (1958), the same kinds of errors are committed by high school students year after year. This stability of error types, irrespective of the fact that different students commit them, gives us some confidence that a sample of 146 students from two large Michigan high schools and from Michigan State University will exhibit appropriate data for the type of exploratory research questions we have concerned ourselves with in this study.

Instrument, Delimitation, and Limitations of the Study

Instrument

The instrument consists of 15 items, five on each of the three mathematics content areas, to be answered in 50 minutes. The considerations made in designing this instrument were as follows:

1. errors were sought from each of the three mathematical content areas of arithmetic, algebra, and geometry;
2. certain error categories were anticipated;
3. students at pre-college levels were the subjects of the study;
4. not more than one hour was anticipated as the time the participants in the study could afford for the study; and
5. certain demographic data about the subjects were necessary for important information required by the study.

Delimitation of the Study

As indicated by the nature of research questions one through six, this study is delimited to those variables that are appropriate for exploring the nature of errors committed and the possible effects of those errors on mathematics achievement. This study attempts nothing at learning about error remediation since such a task would require more time than is available for this study. This study does not attempt to test any hypothesis about students' errors in

Moreover, according to Fisher (1952), the same kinds of errors are committed by high school students year after year. This stability of error types, irrespective of the fact that different students commit them, gives us some confidence that a sample of 146 students from two large Michigan high schools and from Michigan State University will [illegible] for the type of exploratory research [illegible] mathematics. Instead, the study is aimed at contributing to the development of testable hypotheses about the errors.

Limitations of the Study

During the assignment of errors to error categories, it will be difficult to decide valid assignments for some errors since some cases will require interviews of the subjects in order to learn more about the nature of the errors. The time constraint both with respect to the time each subject will afford the study and the time available for the study as a whole will limit the scope of any findings from the study. For example, five items per content area of mathematics are too few to afford a near complete error classification in that area. The complexity of the constructs being compared in some of the research questions will mask the clarity of the implications that seem to emerge from the analysis of data pertinent to those questions. For example, in order to answer research question three clearly, clear ways to measure "comparability in achievement" and "comparability of error patterns in mathematics" should avail themselves to the researcher. In this study, "comparability in achievement" is less difficult to describe than "comparability of error patterns in mathematics"; whereas the former can be delineated by appropriate intervals on the achievement scale, the best one can do for the latter appears to be some qualitative characterization.

In spite of all the limitations of this study--those mentioned above and those not mentioned--some light will be shed upon the nature and consequences of students' errors in mathematics at the pre-college level. Some of the benefits to mathematics education of knowledge gleaned from this study, however limited it may be, have been discussed in the "significance of the study" section of this report.

mathematics. Instead, the study is aimed at contributing to the development of testable hypotheses about the errors.

Limitations of the Study

During the assignment of errors to error categories, it will be difficult to decide valid assign[illegible] of the s[illegible]

Summary

Chapter III has addressed several issues about research questions and methodology. Six research questions have been posed as follows:

1. What error types seem to be most frequent among students?

2. Is there a relationship between particular types of errors a student commits and his/her mathematical achievement as measured by test scores?

3. Do students who score comparably on the test also exhibit comparable error patterns in mathematics?

4. Do error patterns vary significantly across (a) gender? (b) number of years of high school mathematics? (c) age?

5. Do individual student error frequencies correlate across the three content areas of mathematics: arithmetic, algebra, and geometry?

6. Is there a significant error patterns difference between the group that scores highest and the group that scores lowest?

Research efforts into finding answers to these questions will be rewarded by knowledge that will aid the development of remediation strategies, on the one hand, and knowledge that will be used to form testable hypotheses about students' errors in mathematics, on the other.

Ten main variables are involved in this study. These are (a) types of error, (b) mathematical achievement, (c) error pattern, (d) individual student error frequency, (e) gender, (f) number of years of high school mathematics, (g) age, (h) arithmetic, (i) algebra, and (j) geometry. Each of these variables is to be measured or otherwise observed for data that are required to answer at least one of the six research questions.

This study is exploratory and largely descriptive, although some quantitative analysis will be applied to questions four and five. For questions three and four, we will use Radatz's error categories to determine the similarity

Summary

Chapter III has addressed several issues about research questions and methodology. Six research questions have been posed as follows:

1. What error types seem to be most frequent among students?

2. Is there a relationship between particular types of errors a student commits and [illegible] mathematics achievement as [illegible]

of error patterns. Other error categories are anticipated in case some errors show up which do not fit into Radatz's five categories.

The level of significance to be used for question four is $\alpha = 0.05$. However, due to our need to avoid statistical significance where no meaningful significance exists, we will use $\alpha = 0.01$ in testing for statistical significance for question five. The population for this study is eleventh grade students in Michigan and Michigan State University freshmen taking pre-calculus mathematics or a first course in calculus during the spring term of 1985. Ninety-five eleventh grade students from two Michigan high schools and 51 Michigan State University freshmen taking a pre-calculus mathematics course or a first course in calculus during the spring term of 1985 will constitute the sample for this study. Each of the subjects in the study will respond to test items on an instrument designed to capture students' errors in mathematics. The instrument includes 15 mathematics items to be answered in 50 minutes.

This study is delimited to those variables that are appropriate for exploring the nature of errors committed and their possible effects on mathematics education. The study attempts virtually nothing at learning about error remediation since such a task would require more time than is available for this study. Several limitations of this study include the difficulty in assigning errors to error categories validly, the time constraints which will limit the scope of the findings, and the complexity of the constructs being compared which will mask the clarity of analysis. However, some light will be shed upon the nature and consequences of students' errors in mathematics as a result of this study, in spite of its limitations.

of error patterns. Other error categories are anticipated in case some errors show up which do not fit into Radatz's categories.

The

CHAPTER IV

DATA ANALYSIS AND RESULTS

Introduction

The first three chapters dealt with the purpose and significance of this study, the current state of knowledge about students' errors in mathematics as found in the literature, and the research questions and methodology of the study, in that order. This chapter sets out to analyze data and to state and discuss the results emerging from the data analysis. The chapter is organized under the following nine sections:

1. discussion of performance and achievement on the test;
2. description of errors committed;
3. general description of emerging error/achievement patterns;
4. error types most frequently observed among the subjects of the study;
5. relationship between particular types of errors committed and mathematical achievement;
6. error patterns as observed across different achievement groups;
7. error categories as observed across gender, number of years of high school mathematics, and age;
8. individual student error frequencies as observed across arithmetic, algebra, and geometry; and
9. comparison of error patterns exhibited by the fourth quartile on the achievement scale with those exhibited by the first quartile on the achievement scale.

Discussion of Performance and Achievement on the Test

One hundred, forty-six students took the 15-question, 45-minute test and scored in the range of 90% to 01%. The mean score was 44.45%, and the median score was 45%. (The test items used in this study are found on the instrument shown in Appendix A). Table 4.1 shows how many respondents answered an item correctly for each item. Items 12, 2, 8 and 13 were the top four, in that order, in being correctly responded to, while items 7 and 15 shared the last position, each having been correctly responded to by only three of the 146 students. Although item 12 was the most often correctly responded to, it was also the sole item that captured the most frequently committed error. There are two explanations for this apparently contradictory state of affairs. First, since no error type was associated with an omission of an item by a respondent, those items that were not as often attempted tended to show comparatively fewer errors associated with them than did those items that were more often attempted. As Table 4.2 shows, item 12 was one of the most often attempted items. Second, since the requirement for accurate positioning of R relative to P as well as data to enable one to acquire that accuracy were only implicitly given in the question, the error of placing R apparently due east of P did not constitute an incorrect response to item 12 for grading purposes. Thus, it was possible for someone to commit the most frequently observed error in item 12 and still be classified as having responded correctly to that item.

Looking at the information within Tables 4.1 and 4.2, we find examples that would disconfirm the assertion that "the more often an item is responded to, the greater the number of correct responses to that item." For example, every member of the sample responded to item five and only 36% of those respondents gave correct responses to it, while four respondents omitted item one and 56%

Discussion of Performance and Achievement on the Test

One-hundred, forty-six students took the 15-question, 45-minute test and scored in the range of 90% to 0%. The mean score was 44.45%, and the median score was 45%. (The test items used in this study are found on the instrument shown in Appendix

Table 4.1
Item by Item Numbering of Completely Correct Responses

Item Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Number of Correct Responses	76	106	71	54	52	16	3	94	30	37	28	119	86	32	3
Accuracy Coefficient	0.52	0.73	0.49	0.37	0.36	0.11	0.02	0.64	0.21	0.25	0.19	0.81	0.59	0.22	0.02

Notes: (a) We have defined "accuracy coefficient" for each item as follows:

$$\text{Accuracy Coefficient} = \frac{\text{Number of Correct Responses}}{\text{Total Number of Respondents}}$$

In our case, the total number of respondents was 146.

(b) The accuracy coefficient for each item was rounded up to two places after the decimal point.

Table 4.2
Item by Item Omissions

Item Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Number of Respondents Who Omitted It	4	1	0	2	0	72	58	2	10	15	22	2	4	14	18

of all respondents gave correct responses to it. Items 6 and 15 show an even better example. The number of omissions for item six is 72, compared to 18 for item 15. Yet only two percent of all respondents gave correct responses to item 15, compared to 11% for item six. To be sure, there seems to be more information in the tables that confirms the above assertion than disconfirms it, but we must reflect some more on the implications of examples such as the one with items 6 and 15. When 72 respondents omitted item six, the maximum possible number of correct responses, after the fact, could only be 74. And 16 of those 74 respondents attempting item six gave a correct response to the item. But when 18 respondents omitted item 15, there were then 128 respondents who attempted it, and only three of those 128 gave a correct response to the item. The question now arises: Does the popularity of a test item always imply that the competencies the item seeks to test have been achieved by most respondents? The performance of members of our sample on items 6 and 15 would discourage an affirmative answer to this question. We are now in position to say that there may well be other important factors that contribute to the popularity of a test item besides the competence of the respondents to answer that test item correctly.

Description of Errors Committed

From all responses the subjects of this study gave to test items, the writer was able to identify and describe 52 different types of errors. Responses to each item were examined for errors, starting with item 1 through item 15, and each error was described as well as coded. What follows is a listing of error codes and corresponding error descriptions for the 52 different types of errors that were observed.

Code	Error Description
SG	Got the sign wrong; switched - with +.
IN	Incomplete argument.
TS	Tried to solve mere expression.
SD	Subtracted denominator from numerator.
FF	Did not add four in item one.
FD	Did not include denominator in result.
WO	Incorrect/unintelligible operation.
CF	Cancelled out X-2 in item two.
DT	Divided through by RHS--division by possible zero in item two.
FS	Did not solve--treated equation as mere expression.
ST	Switched terms in item three.
AD	Added terms instead of subtracting one term from the other in item three.
NR	Said $\sqrt{\frac{1}{4}} = \pm \frac{1}{2}$ in item four.
DR	Said $\sqrt{a^2 - b^2} = a - b$ (a, b as in item four).
OR	Gave final answer as $\frac{1}{4}$ in item four.
LS	Did $\frac{a}{b} - \frac{c}{d} = \frac{a-c}{b-d}$ equivalent for number four case.
SM	Misapplied $\sqrt{\ }$ operation; e.g., "squared" when should be taking square root.

Code	Error Description
SO	Subtracted three from six outside brackets in item five.
AI	Took -3^2 as nine in item five.
MZ	Took 2(-2) as 2-2 = 0 in item five simplification process.
TP	Tried pythagoras theorem in item six.
IT	Interchanged roles of two and three in solving correct equation in item six.
PC	Solved problem for special case-regular pentagon in item seven without an argument why this special case will do.
NI	Thought the net percentage increase in price was zero in item eight.
RA	Got correct answer through correct solution process, but resisted it! Stated answer should be $50 in item eight.
PB	Proceeded as though ant P starts from B in item nine.
AT	In equation, added 10 seconds instead of subtracting them in item nine.
AM	Took arithmetic mean of eight and ten to imply that the two pumps take nine hours when working together in item 10.
AS	Simply added 4.44 to 7.00 and got 11.44 a.m. which really, now, means eleven forty-four a.m. Did not change .44 of hour into minutes.
TF	Worked out $(\frac{2}{5})^3 \cdot 50$ in item 11.
SA	Thought same amount, i.e., 20 litres, is squeezed out each time and gave 60 as answer.
SQ	Thought same quantity is squeezed out each time and that all juice in apples is squeezed out at third press. This got 10 litres as answer.
SF	Literally subtracted $\frac{2}{5}$ from 50 (three times)
AA	Took amount squeezed out altogether as of end of third press.
EE	Put R due east of P in item 12.
SW	Put R southwest of Q.
NE	Put R northeast of P.
PQ	Switched positions of P and Q.
I	Appeared unable to understand bearings.

Code	Error Description
MP	Did not think D could be uniquely positioned in item 13.
AN	Represented angle of 90° as acute.
CD	Use of given data appeared confused.
SS	Separated the triangles in item 13.
ET	Not clear that student understands what an equilateral triangle is.
DE	Not clear that student understands fully the definition of an even number. For instance, says 0 is not even.
ES	Thinks that to square is to double. Thus, thinks every square number is even.
PN	Either thinks the issue in this question is the fact that the expression is always non-negative or equates "even" with non-negativeness.
TE	Thinks that an even number of terms must necessarily give an even number as their sum.
PH	Thinks pentagon is six-sided.
CA	Uses the converse of regular--equiangular to say "true" to item 15.
GT	Says that "there is some geometry theorem" that attests to this statement and says "true" to item 15.
DP	Appears to have a limited notion of "pentagon"; e.g., thinks pentagon is  necessarily.

With respect to how widespread across items these errors were, IN, WO, and SG were the most spread out (see Table 4.3). IN was observed in six different items, namely items 1, 2, 6, 9, 10, and 14; WO was observed in four items, namely 1, 2, 3, and 5; while SG was observed in four different items, namely 1, 2, 3, and 5. The fact that these three errors were the most spread out among different items should surprise no one, since every test item can be responded to incompletely, many test items involved arithmetical manipulations which thus became prone to sign errors, and every item can be responded to with

an incorrect or unintelligible operation. What may be interesting about these errors, though, would be to relate the extent to which the same error is widespread across respondents. Put another way, one wonders whether an error that is potentially observable among most test items is necessarily an error most respondents commit. The answer to research question one later in this paper will shed some light on the subject of this query. Some errors such as DR, SO, NI, and CA, to mention only a few, are individual item-specific. We will now give a brief mention of aspects that make some of the errors we observed appear very interesting. We will do this for errors NI, RA, TS, FS, SD, and MZ.

Table 4.3
Types of Errors by Item

Item Number	Error Types	Number of Error Types
1	SG, IN, TS, SD, FF, FD, WO	7
2	CF, SG, WO, DT, IN, FS	6
3	ST, TS, AD, WO, SG	5
4	NR, DR, OR, LS, SM	5
5	SO, SG, AI, WO, MZ	5
6	TP, IN, IT	3
7	PC	1
8	NI, RA	2
9	PB, IN, AT	3
10	IN, AM, AS	3
11	TF, AA, SQ, SA, SF	5
12	EE, SW, NE, PQ, ET	5
13	MP, AN, CD, SS, ET	5
14	PC, IN, PN, ES, DE, TE	6
15	PH, CA, GT, DP	4

an incorrect or unintelligible operation. What may be interesting about these errors, though, would be to relate the extent to which the same error is widespread across respondents. Put another way, one wonders whether an error that is potentially observable among most test items is necessarily an error most respondents commit. The answer to research question one later in this paper will shed some light [illegible] Some errors such as PR, SO, NI, and FA [illegible]

The error NI was observed only in item eight and was characterized by the respondent's thinking that the net percent increase in the price was zero. What makes this error very interesting is that although it was very frequently observed among the respondents as a whole group, no respondent on the fourth quartile of achievement committed the error. For our sample, we can make the assertion that the top 25% of the achievement scale was NI-free. The error RA occurred only once and only in item eight. RA was characterized by a respondent's using a correct method of solution for item eight, obtaining the correct response to the item, and then rejecting this correct response for an incorrect one which appeared to have made more sense to him/her. This respondent, having stated that the answer was $50, commented thus: "actually $49.5, according to my calculator, but I think something is fundamentally wrong." The right answer to this question is $49.50, the very answer this respondent resisted by insisting that the correct one is $50.00. It might be of interest to note that this respondent scored in the top 25% on the achievement scale. The error TS was observed in items one and three and was characterized by the respondent's trying to solve some equation where only an expression to simplify had been given. In most cases, the respondent committing this error simply equated the given expression (or its equivalent) to zero and then went ahead to solve. Though interesting in itself, we would probably not have singled this error out for comment at this point had we not observed its opposite! The error FS was observed several times in item two and was characterized by the respondent's not solving the equation but ignoring the equality sign and simplifying as though the equation had been a mere expression. We can see that TS and FS are opposites. What we found interesting about them is that those who committed TS almost invariably scored better on the achievement scale than those who committed FS.

The error NI was observed only in item eight and was characterized by the respondent's thinking that the net percent increase in the price was zero. What makes this error very interesting is that although it was very frequently observed among the respondents as a whole group, no respondent on the fourth quartile of achievement committed the error. For our sample, we can make the assertion that the top [illegible]

only [illegible]

The error SD occurred only in item one and was characterized by the respondent's subtracting the denominator from the numerator for the term $\frac{3x}{-x}$ to obtain 2x. What makes this error interesting to the writer is the fact that this error occurred often enough to convince one that it was an important observation for this study, in spite of the fact that the writer had least expected to observe that kind of error. The error MZ occurred only in item five and was characterized by the respondent's taking 2(-2) as 2-2 = 0. The same can be said for MZ as was said for SD as far as our interest in MZ goes.

We will now comment briefly on why we had least expected errors SD and MZ to occur as frequently a they did. (SD occurred 18 times, and MZ occurred 12 times). Our classroom experience has shown that a problem like $\frac{21}{-7}$ has rarely been responded to by 14 when pupils have been asked to "reduce to simplest terms." Yet committing SD on $\frac{21}{-7}$ results in 21 - 7 = 14. A student who, for example, would successfully simplify $\frac{21}{-7}$ to obtain -3 and still commit SD for $\frac{3x}{-x}$ would appear to have difficulties with the variable x in the expression. Our surprise with the frequency that MZ occurred is rooted in our prior assumption that almost every student who can multiply numbers at the high school level will always read a(b) as "a multiplied by b" and not "a added to b." Hence, a student who can correctly state that 2 x -2 = -4 and still commit MZ for 2(-2) would appear to have difficulties only with brackets ( ).

General Description of Emerging Error/Achievement Pattern

So far in this chapter, we have reported the types of errors that were identified from responses by members of our sample, described and coded these errors, and commented on what appeared to us to be the most interesting aspects of some of the errors. This section will present a unified, albeit broad, appraisal of a picture that seems to be emerging as we view errors with

The error SD occurred only in Item one and was characterized by the respondents subtracting the denominator from the numerator for the term [illegible] to obtain 2x. What makes this error interesting to the writer is the fact that this error occurred often enough to convince one that it was an important observation for this study, in spite of the fact that the writer had least expected to obs[illegible]

charac[illegible]

achievement. We find it convenient at this point to make some assertions that can be warranted by our data.

> Assertion 1: High achievers appear to exhibit similar kinds of misconceptions, if any, on the same item.

There is evidence in our data to show that this assertion is tenable since a typical reason given by six of eight respondents who say that the statement of item 15 is true is as follows, "If straight lines be drawn from the center point to the vertices of the pentagon, then the triangles so formed are isosceles and congruent whose bases are the sides of the pentagon. These sides must be equal." The eight respondents above scored between 80 and 90%--they are the top eight on the achievement scale. (Paradoxically, all three respondents who got the right answer to item 15 scored lower than the third quartile of the achievement scale.)

> Assertion 2: Some errors seem to be inaccessible to those who score lowest on the achievement scale.

The errors NR, TP, and CA were conspicuously rare among those respondents in the bottom 25% on the achievement scale. It was generally observed that the lower a respondent scored, the more likely it was that one found it difficult to characterize what the respondent often did on those items that were missed. It may thus be possible that the very low achievers may be committing errors that they do not succeed to communicate to an observer reading their work.

> Assertion 3: Girls appear to suffer more from lack of partial credit than boys since the former more often tend to refrain from showing work whose accuracy they are not sure of.

The total number of respondents in this study was 146, of which 76 were girls and 70 were boys. The writer observed a total of eight girls who erased work which would otherwise have earned them substantial partial credit. (No other work was substituted for the erased work.) Only one boy was observed to

have erased some work worth partial credit. Erasing work aside, most girls who scored in the top 25% did so with scores that were contributed to by perfectly correct responses about 90% of the time. Unlike the boys in the same achievement group (the top 25%), these girls tended to omit items they were not sure they could get perfectly correct.

Assertion 4: A large number of low achievers tend to perform simple arithmetic manipulations with numerical data in story problems even when these manipulations are totally unrelated to procedures that would lead to the correct answer.

In item 10, for example, many low achievers simply added eight hours to ten hours and concluded that the two pumps working together would take 18 hours. However, a more common error for the whole group tended to be characterized by the respondent's saying that the two pumps working together would take $\frac{8 + 10}{2}$ hours; i.e., nine hours. In item 11, for example, most low achievers committed errors like subtracting $\frac{2}{5}$ thrice from 50 and concluding that the result was the amount of juice squeezed out of the apples during the third press of the piston. Many in the low achievement group argued, for item 11, that $\frac{6}{5}$ of the original juice would be squeezed out of the apples during the third press of the piston! However, a more common error, typical of respondents from higher scoring groups, was to compute $(\frac{2}{5})^3 \cdot 50$ and conclude that $(\frac{2}{5})^3 \cdot 50$ pints is the amount of juice squeezed out during the third press of the piston. Now, reflecting some more about the (10 + 8) hours error, one gets the impression that the respondent did not check whether the result made ordinary sense. We hold the view that hardly anyone in the sample could fail to realize that two pumps working together must take a shorter time to empty the pool than either pump takes working alone. Thus, to obtain 18 hours as the time both pumps would take to empty the pool when their individual times to empty the pool are eight hours and ten hours, it is most likely that the respondent simply added the numerical have erased some work worth partial credit. Erasing work aside, poor girls who scored in the top 25% did so with scores that were contributed to by perfectly correct responses about 90% of the time. Unlike the boys in the same achievement group (the top 25%), these girls tended to omit items they were not sure they could get perfectly correct.

data eight and ten, if only because those data were given in the question. It is very unlikely, however, that such a respondent tried to check whether 18 hours makes ordinary sense as the time the two pumps need to empty the pool.

For item 11, to subtract $\frac{2}{5}$ thrice from 50 and claim that $(50 - \frac{6}{5})$ pints is the amount of juice that is squeezed out during the third press of the piston, the respondent most likely simply manipulated the numerical data $\frac{2}{5}$ and 50 in the question. Even if such a respondent viewed the question as saying that every press of the piston squeezed out $\frac{2}{5}$ of a pint, then the most likely answer, to follow logically from that viewpoint, would surely be $\frac{2}{5}$ of a pint! $(50 - \frac{6}{5})$ pints would then be the amount left in the apples after the third press of the piston, according to such a mistaken view of the item. As for saying that $\frac{6}{5}$ of the original juice was squeezed out, it would appear that it would be possible to squeeze out more juice than is conceivably present in the apples at any time of the process Our assertion number four above was motivated by a realization that to sustain those responses we have quoted is almost certainly a consequence of not checking whether such responses make ordinary sense. To conclude this section, we now state together the four assertions we have made and found evidence in our data to warrant them. These assertions constitute a broad picture that is beginning to emerge from our data about the errors we have observed and the achievement of our sample members on the instrument test. Following is the list of our assertions:

1. High achievers appear to exhibit similar kinds of misconceptions, if any, on the same item.

2. Some errors seem to be inaccessible to those who score lowest on the achievement scale.

3. Girls appear to suffer more from lack of partial credit than boys since the former more often tend to refrain from showing work whose accuracy they are not sure of.

data eight and ten, if only because these data were given in the question. It is very unlikely, however, that such a respondent [illegible] whether 18 hours makes ordinary sense as the time the two pumps need to empty the pool.

For item 11, to subtract $\frac{7}{y}$ thrice from 30 and claim that $(30 - \frac{21}{y})$ pints is the amount of juice that [illegible] respondent [illegible] quest[illegible] pro[illegible]

4. A large number of low achievers tend to perform simple arithmetic manipulations with numerical data in story problems even when these manipulations are totally unrelated to procedures that would lead to the correct answer.

Error Types Most Frequently Observed Among the Subjects of the Study

This section responds to our research question one which states, "What error types seem to be most frequent among students?" As can be seen in Table 4.4, the 12 most frequently committed errors by our respondents are listed together with the frequency with which each error occurred. Thus, the error EE was the most frequently observed, and it showed up among 57 of the 144 respondents who attempted item 12. Thus 39% of those who responded to our instrument items committed EE on item 12. Very closely following EE in frequency of occurrence were the errors SO and DR which both showed up among 51 of our respondents. Thus both DR and SO were committed by 35% of the total number in our sample. Thus, the best our research efforts could do to answer research question one of this study is summarized in Table 4.4.

Table 4.4
Errors Most Frequently Committed

Error	Frequency of Occurrence
EE	57
SO	51
DR	51
NI	35
ST	34
CA	34
IN	33
SG	31
SS	28
AI	25
TP	24
TS	23

NOTE: Refer to error code/description (page 39) to interpret the first column.

We are now in position to address the relationship between the extent to which an error is widespread across test items and the extent to which the same error is widespread across respondents. In the second section of this chapter, we saw that IN was the most widespread error across items while errors WO and SG followed IN closely in being widespread across items. Errors such as EE, SO, DR, IN, ST, and CA include some that were cited only for their being least widespread across items since they were specific only to individual items. Yet IN, the most widespread error across items, ranks only seventh after EE, SO, DR, NI, ST, and CA in being widespread across respondents. SG, the second most widespread error across items, ranks number eight in being widespread across respondents; and WO, the second most widespread error across items, ranks below the 12 most widespread errors across respondents. The nature of our survey is such that we can only justifiably state the following from the above observations. Being widespread across items, for a given error, does not necessarily make that error nearly as widespread across respondents. From the above observations, the following working hypothesis might be formulated to be tested by studies subsequent to this one, "Errors that are specific to individual items on the test are significantly more widespread across respondents than errors that can be committed on more than one test item."

Relationship Between Particular Types of Errors Committed and Mathematical Achievement

Through this section, we respond to our research question two which states, "Is there a relationship between particular types of errors a student commits and his/her mathematical achievement as measured by test scores?"

In section three of this chapter, we made four assertions that tended to relate, in very broad terms, achievement with type of error. We will now draw on the results of our research efforts to answer the above question. Table 4.5

gives a listing of 12 errors that occurred more frequently than others across achievement quartiles.

We will first define our column headings in this table. By first quartile, we mean those respondents who scored in the bottom 25% on the achievement scale. By second quartile, we mean those respondents who scored higher than the bottom 25%, but not higher than the top 50%. Third quartile means those who scored not lower than the bottom 50%, but not higher than the top 75%. The fourth quartile consists of those respondents who scored in the top 25%.

Table 4.5
Twelve Errors Most Frequently Occurring Across Achievement Quartiles

	Frequency					
Error	1st*	2nd*	3rd*	4th*	Total	Relative Frequency
EE	11	17	13	16	57	0.133
SO	23	19	6	3	51	0.120
DR	20	13	10	8	51	0.120
NI	18	11	6	0	35	0.082
ST	18	9	4	3	34	0.080
CA	4	7	17	6	34	0.080
IN	11	13	5	4	33	0.077
SG	12	8	6	5	31	0.073
SS	14	9	4	1	28	0.066
AI	6	8	9	2	25	0.059
TP	3	6	6	9	24	0.056
TS	6	6	7	4	23	0.054
TOTALS:	146	126	93	61	426	
PRCNTG:	34.2	29.8	21.8	14.5	100.0	1.000

* = quartile
NOTE: Refer to error code/description (page 39) to interpret the first column.

Figure 4.1 shows a graph of total error percentage against achievement quartiles. As expected, there is a downward trend of total error percentage

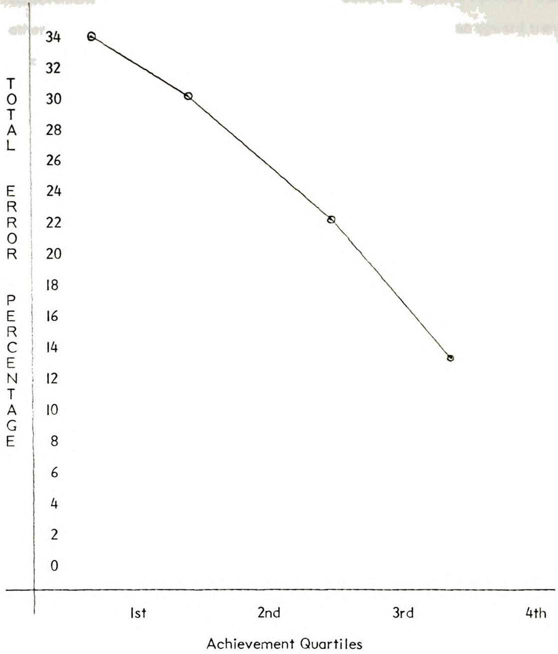


Figure 4.1. Total error percentage versus achievement quartiles.

Figures 4.2, 4.3, 4.4, and 4.5 show graphs of error frequencies against achievement quartiles for the 12 errors that occurred more frequently than others. As can be seen from Figure 4.2, error EE tends to show an upward trend of error frequency from the low achievers to the high achievers--a trend that is reversed only from the second quartile to the third quartile. Thus, it seems to be the case that error EE was less accessible to those scoring low on the achievement scale than it was to those scoring high. With respect to ordinary expectation, it would appear that EE has exhibited some anomalous characteristic. Errors SO and DR, however, show a downward trend of frequencies from low achievers to high achievers--a state of affairs that is commensurate with expectation. Figures 4.3 shows errors NI and ST with frequency trends similar to those shown by SO and DR across achievement quartiles. Error CA, however, shows an upward frequency trend from the first quartile through the third quartile, reversing this trend only between the third and fourth quartiles. Hence CA is another error exhibiting some anomalous characteristic. Figure 4.4 shows errors IN, SG, and SS with frequency trends essentially downward from the low achievers to high achievers. This is as expected since a high incidence of error tends to counter the acquisition of full credit where these errors occur. Figure 4.5 shows errors AI, TP, and TS, all showing upward frequency trends from the first through the third quartiles, with TP showing the same upward trend all the way, while TS and AI reverse their trends between the first and fourth quartiles. We will now discuss errors that are clearly distinguishing of achievement groups.

Errors SO, DR, NI, ST, SG, and SS do clearly distinguish achievement groups in that the higher a student scores on the achievement scale, the more likely it is that that student did not commit an error from the six listed. Of the six, errors NI and SS do show some profound distinguishing characteristics: they

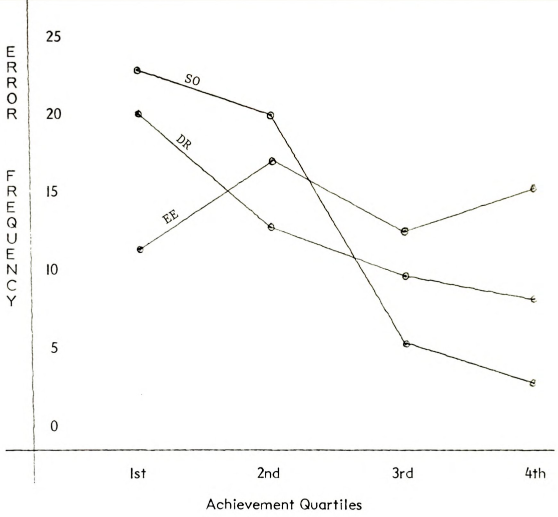


Figure 4.2. Error frequency versus achievement qualities for EE, SO, and DR.

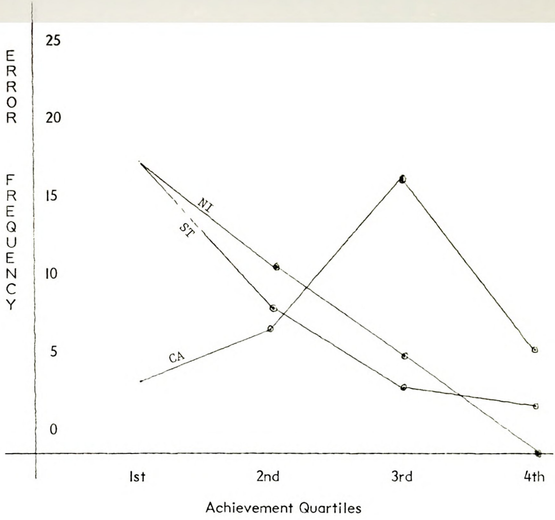


Figure 4.3. Error frequency versus achievement qualities for NI, ST, and CA.

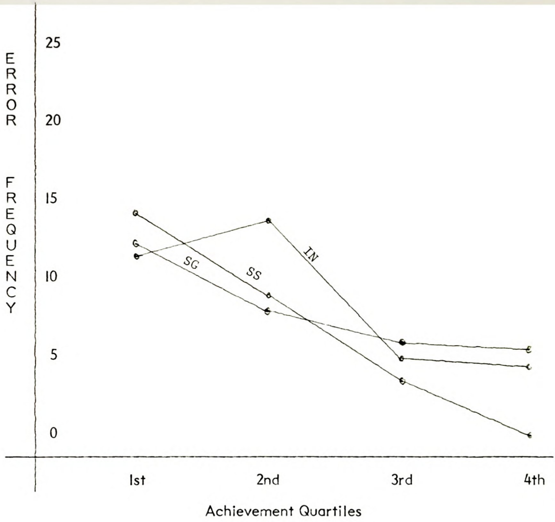


Figure 4.4. Error frequency versus achievement qualities for IN, SG, and SS.

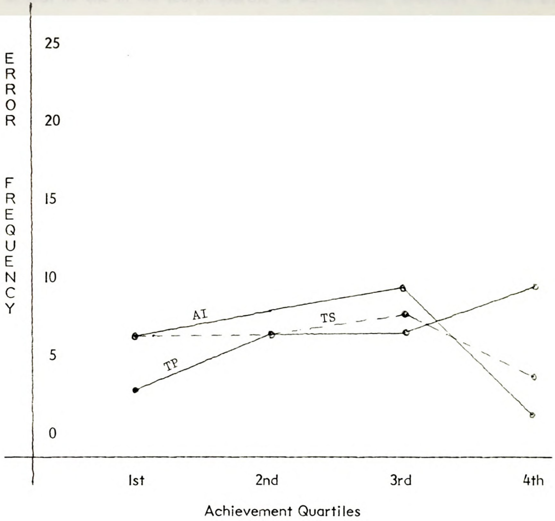


Figure 4.5. Error frequency versus achievement qualities for AI, TP, and TS.

consistently, almost linearly, decline in their frequency of occurrence from low achievers to high achievers. Error NI shows some additional unique phenomenon in that no one in the fourth quartile of achievement committed NI! We are literally saying that no one from the top 36 students on the achievement scale, out of a total of 146 students, committed NI. We hold the view that this error has excelled in distinguishing achievement groups. Recall that NI is the error characterized by the respondent who thinks that increasing 50 by 10% and then reducing the result by 10% gives 50 as the final answer. Recall also that SS is the error characterized by the respondent who separates the triangles described in item 13 in spite of the information that both triangles have side BC in common. Although we were unable to interview any respondent who committed either NI or SS, we may still observe that both these errors could arise from ignoring some crucial detail in the data of the relevant items. In item eight, a respondent who overlooked the fact that the 10% reduction was reckoned on the result of raising 50 by 10% and not on just the original 50 was most likely to commit NI. In item 13, a respondent who overlooked the information that both triangles had side BC in common was most likely to commit SS. We thus observe that both NI and SS very likely have similar sources--overlooking some crucial detail in the data. We now venture a working hypothesis, "High achievers are less likely to ignore details in data than are low achievers."

Among the 12 errors found to occur most frequently, TS seems to be the only one that fails to distinguish achievement groups substantially. This has the implication that a respondent who commits TS is nearly as likely to be a high achiever as a low or average achiever.

Errors EE, CA, AI, and TP seem to share an overall upward trend in frequency of occurrence from low achievers to high achievers, this trend being consistent for at least two of three inter-group frequency changes. Of the four

listed, TP does not show any frequency trend reversal. The implication for this characteristic of TP is that a respondent who commits TP is more likely to be a high achiever than a low achiever. Now this state of affairs would appear to run counter to ordinary sense until we note that the error TP was committed while responding to item six, an item that was not popular with most low achievers who simply omitted it. The anomalous characteristics of EE and CA have already been discussed: these two are fairly distinguishing of achievement groups. The error AI clearly distinguishes the top 25% of achievers from the rest.

The 12 errors listed in Table 4.5 can be divided into two groups. Errors EE, CA, and TP form the first group, characterized by their tendency to show up more frequently in the higher achieving respondents than in the lower. Errors SO, DR, NI, ST, IN, SG, SS, AI, and TS form the second group, characterized by their tendency to show up less frequently in the higher achieving respondents than in the lower. Each error in the first group is distributed across achievement quartiles in a manner contrary to our ordinary expectation, since one would expect a downward trend of error frequency across achievement quartiles from the lowest achievers to the highest achievers. Each error in the second group, however, is distributed across achievement groups as one might expect. Considering the anomalous characteristic exhibited by each of the three errors in the first group, one might ask, "What do these errors have in common that is not shared by all errors in the second group?" An answer to this question might suggest some explanation as to the anomaly each of the three errors appears to show.

Each of the errors EE, CA, and TP is specific to an item of geometry: error EE was committed on item 12 only, error CA was committed on item 15 only, and error TP was committed on item 6 only. And items 12, 15, and 6 are all

geometry items. However, the errors in the second group were committed on items distributed across all three areas--arithmetic, algebra, and geometry. Errors SO, DR, IN, and AI were committed on items 5, 4, 8, and 5, respectively; and items 5, 4, and 8 are on arithmetic. Error ST was committed on item 3 which is an item on algebra. Error SS was committed on item 13, an item on geometry. Errors IN and SG were not specific to any single item (see Table 4.3). Thus, it would appear that the one thing errors in the first group have in common that is not shared by all errors in the second group is the geometry factor. It is the case, perhaps, that committing each of these three errors in geometry results from an exposure to some geometric notions which the low achievers have not seen? Though this exploratory study cannot adequately answer this question, subsequent studies could be designed to investigate this geometry factor further.

Error TP, however, can be isolated from EE and CA and its apparent anomaly accounted for. As can be seen from Table 4.2, as many as 72 respondents from a total of 146 omitted item 6 to which error TP was specific. Moreover, those who omitted item 6 were largely from the category of low achievers. Thus low achievers committed error TP less frequently than high achievers simply because the former attempted item 6 less frequently than the latter--and error TP was specific to item 6. However, the anomalous characteristics of errors EE and CA cannot be explained as the apparent anomaly for error TP has been explained, since items 12 and 15 were attempted almost equally frequently by low achievers as by high achievers.

Concerning research question two, there is evidence from our data that indicates a relationship between particular types of errors a student commits and his/her mathematical achievement on this instrument. With respect to error NI, we have seen that a student committing that error is more likely to be a low geometry items. However, the errors in the second group were committed on items distributed across all three areas--arithmetic, algebra, and geometry. Errors SO, DR, IM, and AI were committed on items 3, 4, 8, and 5, respectively; and items 3, 4, and 8 are on arithmetic. Error ST was committed on item 1 which is an item on algebra. [illegible] geometry. Error [illegible]

Thus, it would [illegible]

that [illegible]

the [illegible]

achiever than a high achiever. We have observed the same for errors SS, SO, DR, ST, and SG. Errors EE, CA, AI, and TP, however, are more likely to be committed by high achievers than low achievers according to our data. Error TS has not been found to distinguish achievement groups.

Error Patterns as Observed Across Different Achievement Groups

This section will address issues pertinent to research question three which states, "Do students who score comparably on the test also exhibit comparable error patterns?" For the purposes of answering this question now and questions four and six later, we now describe a method that we have used to classify each of the errors in one of the five categories which we have named A, B, C, D, and E. (All 52 different types of errors observed were placed into these categories except BI which we were unable to classify.) The five categories used are from Radatz (1979), and the criteria used to assign errors to categories were strongly influenced by the way Marshall (1981) interpreted these categories. Following is a listing of Radatz's error categories, with Marshall's interpretive descriptions appearing in parentheses after each of the first four categories:

1. errors in semantics (meaning, translation),
2. errors in spatial visualization (scale, shape confusion),
3. errors in mastery (operations),
4. errors in association (number patterns, keywords, formulas), and
5. errors characterized by use of irrelevant rules.

To categories 1, 2, 3, 4, and 5 above, we have associated letters A, B, C, D, and E, respectively. Table 4.6 shows how each error (except BI) has been assigned to some category. Table 4.7 shows the frequency of error types across categories where one observes that most error types received category A classification achiever than a high achiever. We have observed the same for errors SS, SA, DR, ST, and SC. Errors EE, CA, AL, and TR, however, are more likely to be committed by high achievers than low achievers according to our data. Error TA has not been found to distinguish achievement groups.

(41.2% of the total error types classified) and category B received the least number of error types (6.0% of the total error types classified).

Table 4.6
Error Categories

Error	SG	IN	TS	SD	FF	FD	WO	CF	DT	FS	ST	AD
Category	C	B	A	A	A	C	C	E	E	E	A	A
Error	NR	DR	OR	LS	SM	SO	AI	MZ	TP	IT	PC	NI
Category	E	A	E	E	A	A	A	A	D	C	A	D
Error	RA	PB	AT	AM	AS	TF	SA	SQ	SF	AA	EE	SW
Category	A	A	A	E	D	D	E	A	E	A	B	A
Error	NE	PQ	BI	MP	AN	CD	SS	ET	DE	ES	PN	TE
Category	D	A	--	D	B	D	A	D	D	A	D	D
Error	PH	CA	GT	DP								
Category	D	D	E	D								

Table 4.7
Frequency of Error Types Across Categories

Category	A	B	C	D	E
Number of Errors	21	3	4	14	9
Percentage	41.2	6.0	8.0	27.4	17.4

In order to determine comparability in achievement, achievement groups were formed whereby each group was characterized by a score class of class width five. (The bottom class contained only one member who scored 01% and was thus merged with the next bottom class.) In order to determine comparability in error patterns, each respondent was assigned an error category as follows. Each error type a respondent committed was associated with a

category, and the category that received the biggest number of error types was the category assigned to that respondent. For example, a respondent who committed error types DR, OR, SG, TS, FD, and IT would be assigned error category C since the categories E, A, C, A, C, and C, respectively, were assigned to those error types and category C received the largest number of error types. In the case of multiple error categories tying for largest number of error types, the category chosen was the one least assigned prior to the tie. For example, assuming the example above to be that of the first respondent, suppose the second respondent showed error types SG, TS, FF, and FD. These error types were associated with error categories C, A, A, and C, respectively. A and C are tied for the largest number of error types, but category A will be assigned to the second respondent since category C has already been assigned once and A has not been assigned at all prior to the tie. The above analysis has resulted in information summarized in Tables 4.8 and 4.9. These tables will be followed by data addressed directly to research question three.

Table 4.8
Error Categories Across Achievement Groups

Score Class	90 - 86	85 - 81	80 - 76	75 - 71	70 - 66
# of Respondents	2	4	4	2	7
Error Categories	A,C	D,D,B,D	A,D,B,D	B,E	B,B,A,A, A,D,C
Score Class	65 - 61	60 - 56	55 - 51	50 - 46	45 - 41
# of Respondents	11	15	9	16	15
Error Categories	D,B,E,C, E,E,D,D D,E,A	A,A,D,D, D,E,A,D A,D,B,E, D,A	A,D,A,B C,D,D,B, A	A,A,A,A A,A,E,E, D,B,E,A, D,A,A,E	D,E,E,E B,E,A,A, A,B,A,D, C,D,B
Score Class	40 - 36	35 - 31	30 - 26	25 - 21	20 - 16
# of Respondents	15	8	13	6	8
Error Categories	A,A,A,C B,A,A,A, A,C,A,A, D,A,A	E,D,B,D D,A,A,A	A,A,E,A A,A,D,A C,A,B,A, A	E,D,A,A A,A	C,D,A,D D,C,C,B
Score Class	15 - 11	10 - 1			
# of Respondents	6	5			
Error Categories	A,E,B,A, A,A	E,E,D,D, A			

Table 4.9
Number of Respondents Across Error Categories

Category	A	B	C	D	E	TOTAL
# Rspnds.	60	18	11	36	21	146
Prcntg.	41.1	12.3	7.5	24.6	14.5	100

Looking at the top three achievement groups, it becomes apparent that error category A was grossly under-represented, while error category D was clearly over-represented. The extent to which error category A was under-represented in the top three achievement groups can be appreciated when it is noted that only 20% of the respondents in those groups belonged to error

category A. And the extent to which error category D was over-represented within the top three achievement groups can be appreciated when one notes that as many as 50% of the respondents in those groups belonged to error category D, while only 24.6% of the total number of respondents belonged to category D. From this observation, one might formulate a working hypothesis: very high achievers are less likely to exhibit errors in semantics than they are likely to exhibit errors in association.

Looking at the nine respondents who scored between 75 and 66%, one finds that error category D was under-represented while error category B was over-represented. We notice that 33% of the respondents who scored in the range of 75-66% were from error category B, while the percentage of all error category B respondents reckoned on the total number of respondents was only 12.3%. On the other hand, only 11.1% of the total respondents who scored between 75 and 66% belonged to error category D, while as many as 24.6% of the total respondents belonged to error category D. It was important to note that in this same scoring range, category A was still under-represented, error category E was under-represented, and error category C was slightly over-represented.

Eleven respondents scored within the range of 65-61%. In this range error category E was the most over-represented compared to the rest since 36.4% of respondents in this scoring range were from error category E which represents only 14.5% of the total respondents. Error category D was also over-represented in this scoring range since it accounted for 36.4% of the respondents in this achievement group, while the same error category accounted for only 24.6% of the total respondents. Error category C was slightly over-represented in this achievement group, while error categories B and A were under-represented, with A being grossly under-represented.

Fifteen respondents scored within the range 60-56%. Within this scoring range, error category D was the most overrepresented of all categories since it accounted for 46.7% of the respondents in this achievement group while accounting for only 24.6% of the total respondents. All other error categories were under-represented in this scoring group.

Nine respondents scored within the range of 55-51%. Error category B was the most over-represented in this scoring range because it accounted for 22.2% of the respondents in this achievement group compared to its having only 12.3% of the total respondents. Error category E was the most under-represented in this achievement group as it accounted for none of the respondents in this scoring range while accounting for as many as 14.5% of the total respondents. In this scoring range, error categories C and D were slightly over-represented while error category A was still under-represented. It is remarkable to observe that in spite of the fact that error category A dominated other error categories very strongly overall, we are now close to the median score of the respondents down the achievement scale without having found a scoring range where error category A was over-represented even once! This might suggest rather strongly that high achievers are less likely to suffer from errors due to semantics than otherwise. Looked at another way, one may venture the following working hypothesis: Errors due to semantics affect achievement significantly more adversely than do any of the other four error categories proposed by Radatz.

Now we look at the modal achievement group, namely the group scoring in the 50-46% range. There were 16 respondents in this scoring range, and this time error category A was over-represented, since 56.3% of the total respondents in this scoring group were accounted for by this error category while it accounted for only 41.1% of the total respondents. However, we must note that, even in this scoring range, error category A was not the most over-represented.

Note that error category E accounted for 25% of the respondents in this scoring group while accounting for only 14.5% of the total respondents. Comparing the ratios 25 to 14.5 and 56.3 to 41.1, it is clear that error category E was more over-represented than error category A in the modal scoring group (though both A and E were significantly over-represented there). In this achievement group, error categories B and D were under-represented while error category C was entirely unrepresented.

The median scoring group, the group scoring 45-41%, showed error category E as the most over-represented, followed by error categories C and B, in that order, as being over-represented. Error categories A and D were, however, under-represented in this median scoring group, with A more under-represented than D. We still have not observed a remarkable over-representation for error category A, and we are now down beyond the median on the achievement scale.

The group scoring in the range 40-36% contained 15 respondents, of which 11 were from error category A. In this group, error category A was the most over-represented having a slight edge over error category C in over-representation. Both categories B and D were under-represented in this scoring range, while category E was entirely unrepresented.

The group scoring in the range 35-31% consisted of eight respondents, of whom three were from category D, three from category A, one each from categories B and E, and none from category C. In this scoring group, error category D was over-represented, and error category B was only just over-represented. Categories A and E were under-represented, while category C was entirely unrepresented.

Nineteen respondents scored in the two scoring groups ranging from 30 to 21%. Of these respondents, 13 were from error category A, two from error category D, two from error category E, one from error category C, and one from

error category B. In this group, category A was clearly the most over-represented with 68.2% of the achievement group coming from that error category. All other error categories were under-represented in this achievement group. Thus, for the first time we have a scoring group in which error category A clearly dominates every other error category in over-representation. We must note, however, that this achievement group is almost entirely contained in the bottom quartile of achievement. This observation reinforces the working hypothesis formulated about errors due to semantics: Errors due to semantics affect achievement significantly more adversely than does any of the other four error categories proposed by Radatz.

Nineteen respondents scored in the three scoring groups ranging from 20 to 1%. Error categories A, B, C, D, and E were represented in this group by six, two, three, five, and three respondents, respectively. In this scoring group, error category C was grossly over-represented, accounting for 15.8% of the respondents in this group compared to the 7.5% it accounted for in the total sample membership. Error categories D and E were only slightly over-represented in this scoring group, while error categories A and B were under-represented. We must also note that this scoring group represented the bottom 13% on the achievement scale. We may formulate the following working hypothesis from this observation: Errors due to mastery are more likely to be found in very low achievers than elsewhere.

We are now in position to make an overarching observation with respect to answering research question three. To the extent that some error categories clearly show over-representation for each scoring range we have considered, it would seem reasonable to assert that students who scored comparably on the test also exhibited comparable error patterns. Our data in this study would seem to confirm our answer to research question three.

error category B. In this group, category A was clearly the most over-represented with 58.3% of the achievement group coming from that error category. All other error categories were under-represented in this achievement group. Thus, for the first time we have a [illegible] group in which error A clearly dominates every other [illegible] note, however, that [illegible] bottom qua[illegible] hypo[illegible]

Error Categories as Observed Across Gender,
Number of Years of High School Mathematics, and Age

Research question four asked, "Do error patterns vary significantly across (a) gender, (b) number of years of high school mathematics, and (c) age?" Our efforts to investigate this question consisted in running three chi-square tests as follows: (a) cross tabulation was called for error category by sex, (b) cross tabulation was called for error category by number of years of high school mathematics, and (c) cross tabulation was called for error category by age. As can be deduced from Tables 4.10, 4.11, and 4.12, our sample consisted of 70 males and 76 females; the errors were classified into five different categories; and the number of years of high school mathematics for our sample ranged from one to five, with 83 of 146 respondents having taken three years of high school mathematics, and 53 of 146 having taken four years. The remaining 10 respondents took one, two, or five years each. The age of respondents ranged from 15 to 23 years, and 75.9% were between 16 and 18 years inclusively.

Table 4.10
Error Category by Sex Test

	Category					
Sex	A	B	C	D	E	TOTAL
Female	34 57.6%	9 50.0%	6 50.0%	16 44.4%	11 52.4%	76 52.1%
Male	25 42.4%	9 50.0%	6 50.0%	20 55.6%	10 47.6%	70 47.9%
TOTALS:	59 40.4%	18 12.3%	12 8.2%	36 24.7%	21 14.4%	146 100.0%

Raw chi square = 1.621111 with four degrees of freedom.
Significance = 0.8050
Contingency coefficient = 0.10479

Table 4.11
Chi-Square Test for Error Category by Years of High School Mathematics

Math	Category A	B	C	D	E	TOTAL
2 years	4 6.8%	3 16.7%	1 8.3%	0 0.0%	1 4.8%	9 6.2%
3 years	34 57.6%	8 44.4%	4 33.3%	20 55.6%	17 81.0%	83 56.8%
4 years	22 35.6%	7 38.9%	7 58.3%	16 44.4%	3 14.3%	54 37.0%
TOTALS:	59 40.4%	18 12.3%	12 8.2%	36 24.7%	21 14.4%	146 100.0%

Raw chi square = 14.44646 with 8 degrees of freedom
Significance = 0.0708
Contingency coefficient = 0.30007

Table 4.12
Chi-Square Test for Error Category by Age

Age	Category A	B	C	D	E	TOTAL
16	17 28.8%	6 33.3%	3 25.0%	11 30.6%	9 42.9%	46 31.5%
17	18 30.5%	2 11.1%	2 16.7%	11 30.6%	4 19.0%	37 25.3%
18	11 18.6%	2 11.1%	4 33.3%	7 19.4%	6 28.6%	30 20.5%
19	3 5.1%	5 27.8%	1 8.3%	7 19.4%	2 9.5%	18 12.3%
20	10 16.9%	3 16.7%	2 16.7%	0 0.0%	0 0.0%	15 10.3%
TOTALS:	59 40.4%	18 12.3%	12 8.2%	36 24.7%	21 14.4%	146 100.0%

Raw chi square = 24.20594 with 16 degrees of freedom
Significance = 0.0851
Contingency coefficient = 0.37711

Error Category by Sex

The chi-square test for error category by sex showed no significant dependence of error category on sex at the 0.05 level of significance. Moreover, as Table 4.10 shows, an α level as low as 0.8050 would need to have been set by an investigator before the error category by sex chi-square test could be found significant. With such a high probability of type I error (80.5%), any claim that the error category a student belongs to depends on the sex of the student would be entertained only with a dearth of confidence.

Error Category by Number of Years of High School Mathematics

The chi-square test for error category by number of years of high school mathematics showed no significant dependence of error category on number of years of high school mathematics at the 0.05 level of significance. Table 4.11 shows, however, that an α level of 0.0708 would have needed to have been set by the investigator before the error category by number of years of high school mathematics chi-square test could have been found significant.

Error Category by Age

The chi-square test for error category by age was run to determine whether the error category to which a respondent belonged depended on the age of that respondent. The results of the test showed the dependence of error category on age not significant at $\alpha = 0.05$. Table 4.12 shows, however, that an α level of 0.0851 would have needed to have been set by the investigator before the error category by age chi-square test could have been found significant.

Based on the results of the three chi-square tests, the following can be said of research question four: At the 0.05 level of significance, there is no

Error Category by Sex

The chi-square value for error category by sex showed no significant dependence of error category on sex at the 0.05 level of significance. Moreover, as Table [illegible] shows, an α level as low as [illegible] would need to have been set by an investigator before the error category [illegible] significant. [illegible] the error [illegible] be enter[illegible] significant variation in error patterns across gender, number of years of high school mathematics, and age.

Individual Student Error Frequencies
as Observed Across Arithmetic, Algebra, and Geometry

Each respondent was assigned an error frequency vector as follows: a respondent who committed a errors in arithmetic, b errors in algebra, and c errors in geometry was associated with a vector (a,b,c). Then Pearson correlation tests were run to determine whether or not the number of errors students commit in one of the three areas of mathematics corrected with the number of errors students commit in each of the two other areas. Research question five sought to determine whether or not these correlations were significant. As can be seen from Table 4.13, the correlation between individual student error frequencies in arithmetic and in algebra was significant at α = 0.001 (r = .3902). However, the correlation between individual student error frequencies in arithmetic and in geometry was not significant at α = 0.01 (r = .1482). Nonetheless, the correlation between individual student error frequencies in algebra and in geometry were significant at α = 0.001 (r = 0.4094).

Table 4.13
Pearson Correlation Tests for Error Frequencies

	Arithmetic	Algebra	Geometry
Arithmetic	--		
Algebra	0.3902 (α = 0.001)	--	
Geometry	0.1482 (α = 0.037)	0.4094 (α = 0.001)	--

significant variation in error patterns across gender, school mathematics, and age.

Each

Thus at α = 0.01, individual error frequencies in algebra correlate significantly with those in arithmetic as well as those in geometry. However, at the same α - level, individual error frequencies in arithmetic fail to correlate significantly with individual error frequencies in geometry.

Comparison of Error Patterns Exhibited by the Fourth Quartile on the Achievement Scale with Those Exhibited by the First Quartile on the Achievement Scale

Research question six asked, "Is there a significant error pattern difference between the group that scores highest and the group that scores lowest?" Research efforts to answer this question were first focused on error behavior differences that could be readily discerned for the two achievement groups. For this purpose, we took the top 25% on the achievement scale as the highest scoring group and the bottom 25% on the achievement scale as the lowest scoring group. Following are the striking differences observed for the two groups (see Table 4.14 for these observations).

Table 4.14
Error Frequency for First and Fourth Quartiles

Error	First Quartile Frequency	Fourth Quartile Frequency
EE	11	16
SO	23	3
DR	20	8
NI	18	0
ST	18	3
CA	4	6
IN	11	4
SG	12	5
SS	14	4
AI	6	2
TP	3	9
TS	6	4
TOTALS:	146	61

1. Higher incidences of error type DR were observed for the lowest achievement group than were observed for the highest group.

2. A total of 18 NI errors were observed in the lowest achievement group while no NI errors were observed in the highest achievement group.

3. Errors EE, CA, and TP were more frequently observed in the highest achievement group than they were for the lowest achievement group.

4. All incidences of error FS were observed in the lowest scoring group so that no single incidence of FS was observed in the highest scoring group.

5. Error NR was conspicuously absent in the lowest scoring group but was observed four times in the highest scoring group.

6. it was often more difficult to make out what members in the lowest scoring group did on those items they missed than it was for the highest scoring group. Consequently, most errors the highest scoring group committed are likely to have been identified while a significant number of errors the lowest scoring group committed were most probably not identified.

After the above observations, error category membership patterns for the two achievement groups were critically examined in order to reveal any differences that might exist between the groups on this variable. Table 4.15 and 4.16 contain results of that analysis.

Table 4.15
Error Category Membership for Fourth Quartile

Error Category	A	B	C	D	E
# Respondents	8	6	3	13	6
% Membership in Quartile	22.22	16.66	8.33	36.11	16.66
% Membership Overall	41.1	12.3	7.5	24.6	14.5
Coeff. of Resp.*	0.54	1.35	1.11	1.46	1.14

$$*\text{Coefficient of representation} = \frac{\text{Percent Membership for Quartile}}{\text{Percent Membership for All Respondents}}$$

Table 4.16
Error Category Membership for First Quartile

Error Category	A	B	C	D	E
# Respondents	17	3	4	7	5
% Membership in Quartile	47.22	8.33	11.11	19.44	13.88
% Membership Overall	41.1	12.3	7.5	24.6	14.5
Coeff. of Resp.*	1.14	0.67	1.48	0.79	0.95

$$\text{*Coefficient of representation} = \frac{\text{Percent Membership for Quartile}}{\text{Percent Membership for All Respondents}}$$

One striking difference between the two groups on error category representation is the fact that error category A was grossly under-represented in the highest achieving group while the same error category was over-represented (though slightly) in the lowest achievement group. Thus we can state the following: The top 25% on the scoring scale was markedly less likely to commit errors due to semantics than were the bottom 25% on the scoring scale.

One can also observe that all error categories except A were over-represented in the highest achievement group, while only error categories A and C were over-represented in the lowest achievement group. Error category D was the most over-represented in the highest achievement group, while error category C was the most over-represented in the lowest achievement group. Thus we can assert that errors committed in the highest scoring group were most likely to be errors of association, while errors committed in the lowest scoring group were most likely to be due to mastery (to use Radatz's error categorization).

Summary

In this chapter, we have analysed data gathered for the study and stated some results accruing from those data. Under the nine sections of the chapter, the following have been done.

1. Respondent performance and achievement on the test was discussed. Under this section, we entertained the hunch that there may well be other important factors that contribute to the popularity of a test item besides the competence of the respondents to answer that test item correctly.

2. Errors committed by respondents were described and coded. A careful examination of responses from sample members enabled us to identify 52 different types of errors. Under this section, some very interesting aspects of a select group of error types were presented. The extent to which some error types were widespread across items as well as across respondents was examined, and from this exercise we were led to wonder whether an error that is potentially observable among most test items is necessarily an error most respondents commit.

3. An emerging picture of errors and achievements was described in broad terms. Under this section, four assertions warranted by evidence from the data were made and duly substantiated. These assertions were:

   a. high achievers appear to exhibit similar kinds of misconceptions, if any, on the same item;

   b. some errors seem to be inaccessible to those who score lowest on the achievement scale;

c. girls appear to suffer more from lack of partial credit than boys since the former tend to refrain from showing work whose accuracy they are not sure of more often than the latter; and

d. a large number of low achievers tend to perform simple arithmetic manipulations with numerical data in story problems even when these manipulations are totally unrelated to procedures that would lead to the correct answer.

4. Error types most frequently observed among the subjects of the study have been identified. A list of 12 such error types was made with frequency of occurrence ranging from 57 to 23. In this section, some light was shed on the relation between an error's being widespread across items and the same error's being widespread across respondents. We observed that "being widespread across items, for a given error, does not necessarily make that error nearly as widespread across respondents." Under this section, a working hypothesis was also formulated: Errors that are specific to individual items on the test are significantly more widespread across respondents than errors that can be committed on more than one test item. This section was particularly geared to answering research question one, and the listing of the 12 most frequently occurring error types was the research attempt to answer the question.

5. Efforts were directed toward determining whether a relationship exists between particular types of errors committed and mathematical achievement. Some analysis showed that certain error types are markedly distinguishing of achievement groups. For example, it was observed that error type NI remarkably distinguished achievement groups since, although it showed up 35

times over the whole membership of the sample, it did not show up even once in the top 25% scoring group. It was observed that some errors appear inaccessible to low achievers--a situation that might appear surprising to many. The emphasis of this section was to find an answer to research question two. In that regard, evidence was found in the data that indicated a relationship between particular types of errors a student commits and his/her mathematical achievement.

6. Some revealing information about error categories across achievement groups was gleaned from the data analysis pertinent to research question three. Each of the 52 observed error types, with an exception of error type BI, found an error category to belong to from one of the five error categories proposed by Radatz (1979). The criteria to assign particular types of errors to particular categories has been greatly influenced by Marshall (1981), since we leaned heavily on her interpretation of Radatz's error categories to decide which category a given error type belongs to. Each respondent was assigned an error category that contained the majority of his/her errors. That done, the construct of "comparability of error patterns" was delineated since two respondents who belonged to the same error category were said to exhibit comparable error patterns. Comparable achievement was delineated by achievement groups on the achievement scale. Motivated by the work relevant to this section, two working hypotheses were formulated: (a) errors due to semantics affect achievement significantly more adversely than does any of the other four error categories proposed by

times over the whole membership of the sample. It did not show up even once in the top 27% scoring group. It was observed that some errors appear inaccessible to low achievers--a situation that might appear surprising to many. The emphasis of this section was to find an [illegible]

Radatz, and (b) errors due to mastery are more likely to be found in very low achievers than elsewhere. Following from evidence in the data pertinent to research question three is this overarching observation: to the extent that some error category clearly shows over-representation for each scoring range we have considered, it would seem reasonable to assert that students who score comparably on the test also exhibit comparable error patterns. This observation answers in the affirmative research question three.

7. Error categories were then considered in an attempt to determine their relationship to gender, number of years of high school mathematics, and age of respondents. This is the substance of research question four. After running three chi-square tests--(a) error category by sex, (b) error category by number of years of high school mathematics, and (c) error category by age--it was found that the error category someone belongs to does not significantly depend on sex, number of years of high school mathematics, or age (none of the chi-square tests indicated significance at the 0.05 level).

8. Then individual student error frequencies as observed across arithmetic, algebra, and geometry were examined to determine if the frequencies were intercorrelated across these three content areas of mathematics. For this purpose, Pearson correlation tests were run and significance sought at = 0.01. Significantly correlated were (a) algebra error frequencies with arithmetic error frequencies and (b) algebra error frequencies with geometry error frequencies. However, arithmetic error

Raders, and (b) errors due to random [illegible]

in very low achievers

in the

frequencies were not significantly correlated with geometry error frequencies at $\alpha = 0.01$. Thus research question five was answered.

9. Finally, the top 25% and bottom 25% on the scoring scale were investigated for error pattern differences. This is the subject of research question six. Six striking differences in the error behavior of the two groups were stated in the section. These were followed by a critical examination of error category membership patterns in a bid to reveal any differences that might exist between the groups on this variable. It was found that error category A was grossly under-represented in the highest achieving group while the same error category was over-represented in the lowest achievement group. This led to the assertion that the top 25% on the scoring scale were markedly less likely to commit errors due to semantics than were the bottom 25% on the scoring scale. Other observations warranted this assertion: errors committed in the highest scoring group are most likely to be errors of association while errors committed in the lowest scoring group are most likely to be due to mastery (to use Radatz's error categorization).

CHAPTER V

SUMMARY, CONCLUSIONS, IMPLICATIONS, AND RECOMMENDATIONS

Introduction

The preceding chapters of this paper addressed the purpose and significance of the study, the current state of knowledge about student errors in mathematics as found in the literature, research questions and methodology, and data analysis and results, in that order. In this final chapter the following will be addressed: (a) a general summary of the study together with the findings, (b) some major conclusions together with discussion, (c) implications for mathematics education, and (d) recommendations for applications and subsequent studies.

Summary and Conclusions

Purpose and Significance of the Study

This study sought to identify, classify, and analyze students' errors in mathematics at the pre-college level. Bearing in mind that an effective mathematics teacher must get at the real source of a student's learning difficulty before the latter can learn from the former, research efforts that are most likely to assist mathematics teachers become more effective in class need to include aspects of error diagnosis. Thus the major significance of this study is the diagnosis of students' errors in order to inform appropriate remediation strategies. More specifically, this study set out to do the following:

1. contribute to knowledge about errors that cause difficulties in the learning of mathematics,

2. contribute to knowledge about errors that are useful in the development and teaching-learning process of mathematics, and

3. serve a unique role in mathematics education by dealing with students' errors in pre-college mathematics instead of just dealing with arithmetic errors at the elementary school level as the majority of previous studies in this area have done.

Literature Review

Results from many studies on students' errors in mathematics suggest a strong need to learn more about those errors in order to better understand their effect upon the acquisition of mathematical knowledge. Many scholars from countries like Germany, Russia, and the United States have carried out studies on students' errors in mathematics for more than 80 years. Although these scholars represent a diversity of research interests, their findings play a unified role of informing mathematical pedagogy.

Perhaps the one characteristic of errors that is most consistently reported across studies is error persistence. Through some important longitudinal studies that have been done to study persistence of errors, the types of errors that persist have been identified and documented.

Many teachers and students have come to view every error in mathematics as detrimental to mathematics learning. However, a substantial amount of literature exists which counters any stigma associated with errors that are, in fact, useful. Some errors can be used very effectively in class to lead to discoveries of interesting mathematical results. Literature abounds with reports of cases where what began as an error made by a student in class motivated a class discussion that eventually led to some interesting mathematical discovery. Reports also exist in the literature about some mathematically intriguing error techniques that led to correct results. These error techniques are described as "mathematical misteaks" by at least one scholar.

Some scholars have found that many errors students commit follow from strategies that make sense to the students at their cognitive levels. It is thus important that teachers understand their students' strategies before the former can hope to effectively help the latter learn. Individualization is one instructional technique that could be employed to help teachers understand their students' strategies.

Many studies in mathematics education have been done on sex-related differences in mathematics achievement. Findings from these studies conflict, but many of them report students' difficulties which seem to show sex-related differences. Moreover, some of these sex-related differences seem not to be country-specific since they are consistently reported in studies on subjects from different countries.

Research Questions and Methodology

The following research questions were posed.

1. What error types seem to be most frequent among students?
2. Is there a relationship between particular types of errors a student commits and his/her mathematical achievement as measured by test scores?
3. Do students who score comparably on the test also exhibit comparable error patterns in mathematics?
4. Do error patterns vary significantly across (a) gender? (b) number of years of high school mathematics? (c) age?
5. Do individual student error frequencies correlate across the three content areas of mathematics: arithmetic, algebra, and geometry?
6. Is there a significant error patterns difference between the group that scores highest and the group that scores lowest?

Research efforts into findings answers to these questions are bound to be rewarded by knowledge that will aid the development of remediation strategies,

on the one hand, and knowledge that is useful for the formulation of testable hypotheses on the other.

Ten main variables were involved in this study: (a) type of error, (b) mathematical achievement, (c) error pattern, (d) individual student error frequency, (e) gender, (f) number of years of high school mathematics, (g) age, (h) arithmetic, (i) algebra, and (j) geometry.

Methodology

This study was exploratory and largely descriptive although some quantitative analysis was applied to questions four and five. Radatz's error categories were used for questions three, four, and six to determine similarity of error patterns. One error type was identified which could not fit into any of Radatz's five categories, but it was not deemed necessary to create an additional category for this one error. The five categories were as follows:

1. errors in semantics (meaning, translation),
2. errors in spatial visualization (scale, shape confusion),
3. errors in mastery (operations),
4. errors in association (number of patterns, key words, formulas), and
5. errors characterized by use of irrelevant rules.

For question four $\alpha = 0.05$ was used as the level of significance, while for question five $\alpha = 0.01$ was used as the level of significance. A higher significance level was needed for question five in order to avoid statistical significance without substantive significance for the Pearson correlation tests.

The population for this study was eleventh grade Michigan students and Michigan State University students taking a pre-calculus or first calculus course in mathematics. Ninety-five eleventh grade students from two Michigan high schools and 51 Michigan State University students taking a pre-calculus or first

calculus course during the spring term of 1985 constituted the sample for this study. Each subject in the study responded to test items on an instrument designed to capture students' errors in mathematics. The instrument included 15 items to be answered in 50 minutes.

This study was delimited to those variables deemed appropriate for exploring the nature of errors committed and their possible effects on mathematics education. The study attempted virtually nothing at learning about error remediation since such a task needed more time than was available for this study. Several limitations of this study included difficulty in assigning errors to error categories validly, time constraints which limited the scope of the findings, and the complexity of the constructs being compared which tended to mask the clarity of analysis. However, some light was shed upon the nature and consequences of students' errors in mathematics as a result of this study in spite of the limitations.

Data Analysis and Results

A largely qualitative analysis of data was used; the only quantitative statistical tests done were for questions four and six, and these were chi-square tests and Pearson correlation tests, respectively. For the purposes of responding to questions three, four, and six, it was convenient to assign to each respondent an error category which contained the majority of errors that respondent had committed.

Results (Findings) and Conclusions

Research question one stated, "What error types seem to be most frequent among students? Research efforts to answer this question yielded 12 error types as those most frequently committed: EE, SO, DR, NI, ST, CA, IN, SG, SS, AI, TP, and TS, in that order (see Table 4.4).

Research question two stated, "Is there a relationship between particular types of errors a student commits and his/her mathematical achievement a measured by test scores?" Research efforts to answer this question yielded evidence in the data that indicated a relationship between particular types of errors a student commits and his/her mathematical achievement. For example, since error type NI did not show up at all in the top 25% scoring group, it can be said that a student who scores in the top 25% is unlikely to have committed error type NI.

Research question three stated, "Do students who score comparably on the test also exhibit comparable error patterns in mathematics?" Research efforts to answer this question yielded a defensible overarching statement to this effect: to the extent that some error category clearly shows over-representation for each scoring range considered, it would seem reasonable to assert that students who score comparably on the test also exhibit comparable error patterns.

Research question four stated, "Do error patterns vary significantly across (a) gender, (b) number of years of high school mathematics, and (c) age?" Research efforts to answer this question yielded no significance at the $\alpha = 0.05$ for all three chi-square tests. Thus, the finding is that the error category someone belongs to does not significantly depend on sex, number of years of high school mathematics, or age.

Research question four stated, "Do individual student error frequencies correlate across the three content areas of mathematics: arithmetic, algebra, and geometry?" Research efforts to answer this question yielded this information: algebra error frequencies significantly correlated with arithmetic error frequencies and with geometry error frequencies at $\alpha = 0.01$. However, arithmetic error frequencies did not significantly correlate with geometry error frequencies at $\alpha = 0.01$.

Research question six stated, "Is there a significant error pattern difference between the group that scores highest and the group that scores lowest?" Research efforts to answer this question yielded the following assertions:

1. the top 25% on the scoring scale were markedly less likely to commit errors due to semantics than were the bottom 25% on the scoring scale, and

2. errors committed in the highest scoring group are most likely to be errors of association while errors committed in the lowest scoring group are most likely to be due to mastery (to use Radatz's error categorization).

In light of these assertions, we are persuaded to answer research question six in the affirmative.

Discussion

Chapter IV of this paper contains several working hypotheses and assertions whose formulations were motivated by what appeared to be evidence from the data to warrant them. At this time, each will be visited and additional reflection will be done on the foundation of their worth. Each working hypothesis and assertion will be listed before discussion is done.

Working Hypotheses

1. Errors that are specific to individual items on the test are significantly more widespread across respondents than errors that can be committed on more than one test item.

2. High achievers are less likely to ignore details in data than are low achievers.

3. Very high achievers are less likely to exhibit errors in semantics than they are likely to exhibit errors in association.

4. Errors due to semantics affect achievement significantly more adversely than does any of the other four error categories proposed by Radatz.

5. Errors due to mastery are more likely to be found in the very low achievers' category than elsewhere.

Assertions

1. High achievers appear to exhibit similar kinds of misconceptions, if any, on the same item.

2. Some errors seem to be inaccessible to those who score lowest on the achievement scale.

3. Girls appear to suffer more from lack of partial credit than boys since the former more often tend to refrain from showing work whose accuracy they are not sure of.

4. A large number of low achievers tend to perform simple arithmetic manipulations with numerical data in story problems even when these manipulations are totally unrelated to procedures that would lead to the correct answer.

5. Errors committed in the highest scoring group are most likely to be errors of association while errors committed in the lowest scoring group are most likely to be due to mastery.

Discussion of Working Hypotheses and Assertions

We hasten to remark that we neither accord any of the above working hypotheses or assertions the status of a scientifically confirmed thesis nor deny any of them recognition as a potential subject of subsequent scientific inquiry. Our study was essentially an exploratory survey which set out to investigate possible answers to six research questions. The study had no hypotheses initially formulated for confirming or disconfirming and was, therefore, not designed for that task. It was, however, designed to generate some hypotheses and assertions which have the potential to stimulate subsequent research about them. Thus, the above five hypotheses and five assertions have been generated in that spirit. We now visit each hypothesis and assertion for some discussion.

Working hypothesis one intrigues the writer since, on the surface, it seems to run counter to ordinary sense. One would expect, it seems, that an error which can be "picked" from many items on a test has the potential of being "picked" by more respondents to the test than an error that can only be "picked" from one test item. What makes the working hypothesis interesting is that its

presence invites one to consider other factors that may influence the frequency of occurrence for a given error besides the fact that the error is widespread or not widespread across items. May it, for instance, not be the case that an error that is widespread across items has been observed more in the classrooms and thus remediated against more often than an error that is not so widespread across the items? Such a case makes for our hypothesis since an error that is limited to only one test item is likely to be committed by many respondents if only because not many people have been remediated against it.

Working hypothesis two was formulated after observing that the error types specific to items 8 (NI) and 13 (SS) were almost the monopoly of low achievers. We can only speculate, however, as to the possible thought processes a respondent who committed these errors had gone through since it was not possible for the writer to interview any respondent. However, it can be noted that both items 8 and 13 have some detail in their data set which, when glossed over by a respondent, will almost certainly lead to errors NI and SS. (We have already discussed what the crucial details are in the data for each item.) Habel (1958) described a similar kind of error behavior as the one we are speculating about as a possible antecedent to NI and SS.

Working hypothesis three was motivated by a clear under-representation of error category A and an over-representation of error category D in achievement groups from the very top down to an achievement group very close to the median achievement group. If evidence from an investigator's data is anything to go by in formulating hypotheses, then working hypothesis three was clearly inescapable! Hart (1978) made observations about error techniques that would lead to errors in category A and supplements this hypothesis with an elaborate discussion of possible causes for those kinds of error techniques.

Working hypothesis four is a more daring version of working hypothesis three as the former is couched in a language rather suggestive of causality. The worth of this hypothesis in this study is in the stimulus such a hypothesis must have for experimental studies.

Working hypothesis five was motivated by the same kind of observations in the low achieving group as those that motivated working hypothesis three in the high achieving group. Working hypotheses such as five are worth studying because our priority for remediation efforts in mathematics education would naturally go to the low achieving group. Should it be confirmed that working hypothesis five is, indeed, tenable, then remediation strategies geared to correcting errors due to mastery would be given primary consideration.

Assertion one was motivated by observing that the top eight students in achievement, for our sample, had the same kind of argument to justify their choice of "true" (which was the wrong choice) for fifteenth item 75% of the time. Six respondents of the eight argued to the effect that one could partition the equiangular pentagon into five isosceles congruent triangles, each with a vertex at the "center" of the pentagon and each with a base as a side of the pentagon. (These four respondents each committed at most one other error besides this one.) The worth of this assertion in this study is to draw the attention of mathematics educators to the fact that some very high achievers in high school mathematics can also entertain some serious misconceptions which should be looked for and remediated before they take their toll on subsequent mathematics achievement for these students.

Assertion two was motivated by realizing that errors such as TP appeared to be almost absent in the lowest scoring group and much over-represented in the highest achievement group. Ordinary sense would seem to suggest that an error that high achievers commit should be committed even more frequently by low

achievers. For some errors, this is true; but there are some errors also for which the reverse holds. One such error is TP. Another error that shows this behavior, but to a lesser extent, is EE. The worth of assertion two in this study consists in an invitation to always consider alternative explanations for our observations. the writer, for example, was able to find a convincing explanation for the behavior of TP (a disproportionately small number of low achievers attempted item six, the sole item where TP could be committed). However, this explanation cannot apply to EE since item 12 was attempted by all except two respondents. So we must continue looking for alternative explanations!

Assertion three was motivated by the writer's observation that a highly disproportionate number of female respondents had erased what could still be feintly deciphered as work that was worth substantial partial credit. Hardly any work worth partial credit was observed as having been erased by male respondents. Then this question arose: what if these female respondents had not erased their work? Would the earned partial credit not, in that case, have raised the mean achievement for the girls? Sex-related differences in mathematics achievement have been a popular subject of research in the recent past. Fenemma and Sherman (1977) and Benbow and Stanley (1980) are two pairs of scholars whose reports of sex-related differences in mathematics achievement conflict. The former report no significant differences, while the latter report a large difference in favor of boys. The present study has shown no sex-related dependencies for error categories. This assertion should help future researchers to control for such confounding variables as the "erasing partial credit factor" in order to obtain more meaningful results with respect to sex-related differences.

For assertion four enough discussion was given as an extension to arguments when the assertion was first made to warrant it. Its worth in this report consists of cause for needed remediation. The author holds the view achievers. For some errors, this is true; but there are some

the reverse holds. One such error

but to a lesser extent.

an invitation

the writer

behavior

that where errors such as those described in this assertion happen, very bad learning habits are in the making, if not already established. To have the attitude that it is worth one's while to simply carry out any conceivable manipulation with some data in the item, however irrelevant those manipulations are to a correct solution, is to be disposed to abusing the worth of time. Thus, whenever these kinds of error techniques are spotted in a student by a teacher, no effort on the part of the teacher should be spared to help the student out of this error behavior.

Assertion five really happens to be a combination of working hypotheses three and five. Thus, when we already have the two hypotheses, the only worth for assertion five in this study is that it is stated in a form more convenient for ethnographic studies than either of the two hypotheses.

Implications

This study has one central implication for mathematics education: in order to effectively remediate, teachers must reflect deeper upon the possible causes of students' learning difficulties--even a light hearted examination of the kinds of errors students commit in mathematics shows that there cannot be simple explanations with regard to the nature of these errors. Since remediation efforts without proper error diagnosis are, at best, misplaced, mathematics educators must give error analysis the place it deserves in mathematics education.

The five working hypotheses and five assumptions that have been formulated do certainly have non-trivial implications for mathematics education provided such hypotheses/assertions are confirmed by subsequent research. Say, for example, working hypothesis four is true. The implications of this truth for mathematics education would be to focus remediation efforts first toward correcting errors due to semantics. If, say, assertion three is true, then a

reflection is called for upon ways to help the girls it applies to out of the habit of "erasing partial credit."

Recommendations

Looking at all those untested hypotheses and assertions this study has generated, the first recommendation made is that research studies be designed to confirm or disconfirm them. The funding problem should not be allowed to kill studies in error analysis subsequent to this one considering the ultimate value of some of the findings these studies may produce to mathematics education.

Given that we have observed an interesting error type in NI, our immediate recommendation is that a carefully planned study, focused only on NI, should be pursued to learn more about it. Of course, there are many other interesting errors that also deserve whole studies to themselves!

DP was an error type that was both disappointing and interesting to the writer. It was a disappointing error type because as many as 16 of 146 respondents committed it--which was a high percentage for the age level of our respondents. (A person who commits DP indicates that s/he has probably not yet filled Van Heil level one--using the 0-4 nomenclature--in the learning of geometry). It was an interesting error to the writer because a serious reflection over when such a misconception of a pentagon first occurs to a learner was unavoidable. Our recommendations, therefore, are that errors similar to DP be subjected to appropriate research studies to get at their origins because these errors are formidable barriers to subsequent accurate conceptualizations.

Adequate funding or lack of it, adequate time or lack of it, the compelling need to gain more insights into students' errors should be answered positively--mathematics education is ultimately the better for this.

APPENDICES

APPENDIX A

M.S.U. MATHEMATICS DIAGNOSTIC EXAM

Please fill in your personal data as requested below:

(a) Name ______________________________

(b) Female ________ Male ________ (check one)

(c) Age ________

(d) The name of your school ______________________________

(e) Number of years of high school mathematics ________________ (including this year)

Now answer the following questions as completely as you can. For each question you answer, show all the work that leads to your answer.

1: Evaluate $\frac{x-y}{x+y} + 4$ given that y = -2X Answer [ ]

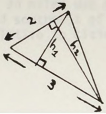
2: Solve the equation (x+3)(x+2) = (x-3)(x+2) Answer [ ]

3: Subtract x-3 from 3-x Answer [ ]

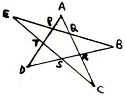
4: Simplify $\sqrt{(\frac{5}{6})^2 - (\frac{2}{3})^2}$ as much as you can. Answer [ ]

5: Evaluate $6-3[2(3-5)-3^2]$. Answer [ ]

6: In the figure, h_1 is the length of the perpendicular from one vertex of a triangle to the side 3 units long while h_2 is the length of the perpendicular from another vertex to another side 2 units long. Express h_1 in terms of h_2.



7: In the figure, PQRST is a pentagon. The star is made by extending each side of the pentagon in both directions. The acute angles at A, B, C, D, and E are measured in degrees. Find the sum of these five acute angles.



8: A pair of shoes, whose price was \$50 last month, now costs 10% more. The price of this pair of shoes next month will, however, be 10% less than the present price. How much will the pair of shoes cost next month?

9: Three ants P, Q, and R are crawling along a line segment AB of length 200 centimeters. Ant Q starts at A and moves toward B at 2 centimeters per second. At the same time, ant R starts at B and move towards A at 3 centimeters per second. 10 seconds later, ant P starts from A and moves toward B at X centimeters per second. Given that the three ants will meet at the same point between A and B, find the numerical value of X.

10: One pump empties a pool in 8 hours and a smaller pump empties the same pool in 10 hours. If the two pumps start working together at 7:00 AM, at what time will they empty the pool?

11: In making apple cider, one press of the piston squeezes out $\frac{2}{5}$ of the juice present in the apples prior to that press. Initially, there are 50 pints of juice in the apples. Compute the amount of juice that will be squeezed out <u>during the third press</u> of the piston.

12: City Q is 30 miles due north of City P while City R is 60 miles distant from City P and southeast of City Q. Draw a picture to illustrate the location of City R relative to cities P and Q.

13: Triangle ABC is isosceles with A = 90°. BCD is an equilateral triangle such that both triangles have the side BC in common. Draw a picture to show these triangles.

14: "If x, y and z are three integers, then $x^2 + y^2 + z^2 + x + y + z$ is always an even number." Say whether this statement is true or false supplying a reason/reasons for your answer.

15: "Any pentagon with all its interior angles equal has all its sides equal." Say whether this statement is <u>true</u> or <u>false</u> supplying a reason/reasons for your answer.

APPENDIX B

APPENDIX B

Table B-1
Female/Male/All Sample Achievement Group Frequencies

Achievement Group	Female Frequency	Male Frequency	All Frequency
100 - 91	0	0	0
90 - 81	1	5	6
80 - 71	3	3	6
70 - 61	9	9	18
60 - 51	12	12	24
50 - 41	23	9	32
40 - 31	11	12	23
30 - 21	9	10	19
20 - 11	7	8	15
10 - 01	1	2	3
TOTALS:	76	70	146

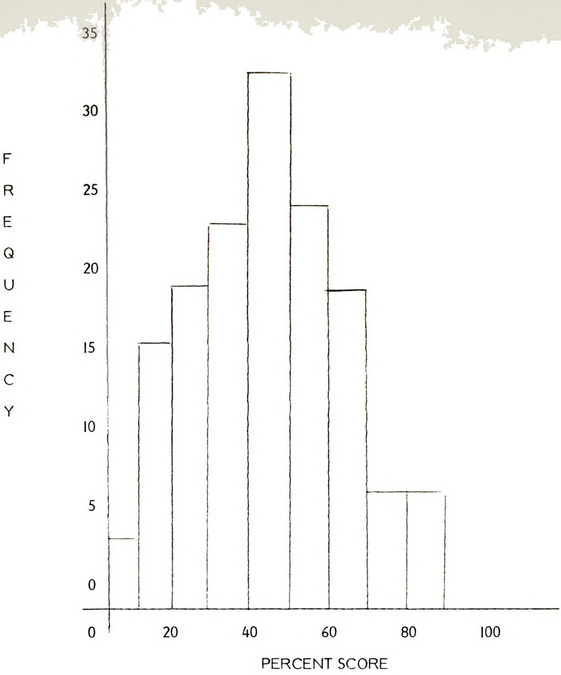


Figure B-1. Frequency histogram for sample achievement scores.

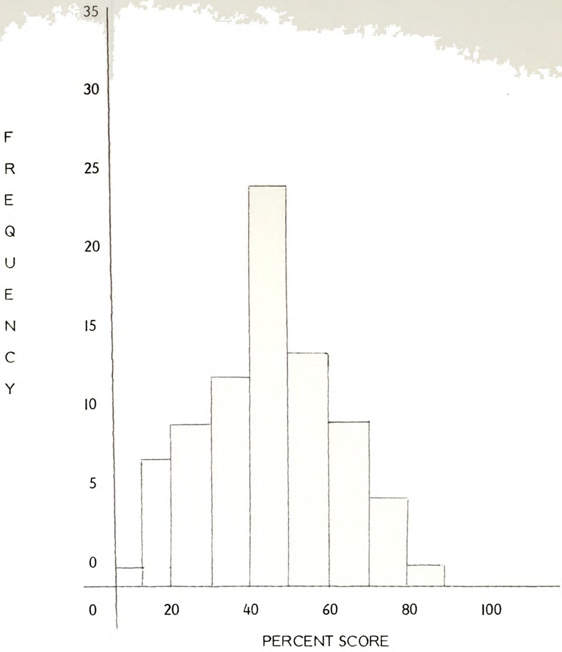


Figure B-2. Frequency histogram for sample female achievement scores.

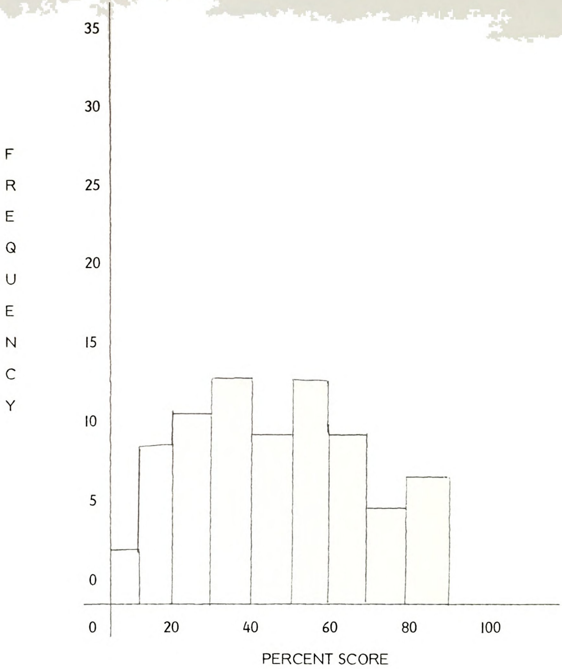


Figure B-3. Frequency histogram for sample male achievement scores.

APPENDIX C

Table C-1
Sex Distribution over Error Categories

Sex	A	B	C	D	E	TOTALS
Female	34	9	6	16	11	76
Male	25	9	6	20	10	70
TOTALS:	59	18	12	36	21	146

Table C-2
Age Distribution over Error Categories

Age	A	B	C	D	E	TOTALS
15	1	1	0	0	0	2
16	16	5	3	11	9	44
17	18	2	2	11	4	37
18	11	2	4	7	6	30
19	3	5	1	7	2	8
20	7	1	1	0	0	9
21	1	1	1	0	0	3
22	1	1	0	0	0	2
23	1	0	0	0	0	1
TOTALS:	59	18	12	36	21	146

Table C-3
Years of High School Math Distribution over Error Categories

Years	A	B	C	D	E	Totals
1	2	1	0	0	0	3
2	2	2	1	0	1	6
3	34	8	4	20	17	83
4	21	7	7	15	3	53
5	0	0	0	1	0	1
TOTALS:	59	18	12	36	21	146

APPENDIX D

Table D-I
Sex Differences in Achievement

	Achievement Aspects			
Sex	Mean	Median	Modal Class	Modal Frequency
Female	45.25	46.00	50-41	23
Male	44.26	42.00	60-51 and 40-31 (bimodal)	12 and 12

APPENDIX E

The Cash Discount Problem: Our Class Discussion

The Problem:	Alex purchases an item valued at $600.00 and makes a down payment of $300.00. Given that the cash discount rate for the item is 10%, how much does Alex still owe after making the down payment?
Method A of Solution (text method)	Since there is a 10% cash discount, $0.90 clears $1.00 of the bill during the discount period. Hence $300.00 cleared $\frac{\$300}{0.90}$; i.e. $333.33 of the bill. Thus Alex still owes $(600.00 - 333.33) or $266.67.
Method B of Solution (method most students used)	Alex should have been given a 10% discount on his $300.00 down payment. Thus, he should have been given $30.00 back. Since he paid all the $300.00, then he effectively cleared $330.00 of the bill. Thus, Alex still owes $270.00.
Method B' (the author's modification of student's Method B)	Note that the $30.00 that should have been paid back is now his cash which he can apply to remittance against his bill. Should he do this, then he is entitled to a 10% cash discount on this $30.00. This amounts to $3.00. Thus, he effectively cleared $330.00, but a balance of $3.00 is due him to his credit. He can apply this $3.00 to remittance against his bill, but he is still entitled to $0.30 as his 10% cash discount on the $3.00. Thus, he clears $333.00, but a balance of $0.30 is still due him to his credit. Continuing our argument in this fashion, we see that his $300.00 down payment must have effectively cleared 333.3\bar{3}$ of his bill (i.e., $333.33 rounded off two places after the decimal point). Thus, Alex still owes $266.67.

Remarks As stated earlier in the text of this paper, the author found it necessary to modify Method B of solution for the class for two main reasons. First, it was not enough for the author to simply suspend the students' method on the grounds that Method A, the text method, gave a correct answer which differed from the $270.00 obtained according to Method B. The students still found their method convincingly correct to them, and it is here that the duty of the instructor calls for convincing the students that there is a flaw in Method B. The flaw arose from ignoring the fact that the $30.00 should also be discounted in favor of Alex. Thus, Method B' was called for in order to help expose the flaw in Method B. Secondly, the recursive process of argument employed in Method B helped enhance the concept of cash discount within the students' stock of knowledge. Methods such as A, though elegantly short in comparison to B', have the disadvantage of being too elegant to sufficiently illustrate the cash discount concept. Thus, Method B' was necessary, at least to fill the pedagogical deficit methods such as A are bound to represent.

BIBLIOGRAPHY

BIBLIOGRAPHY

Badger. (1981, November). Why aren't girls better at math? A review of research. Educational Research, 24(1), 11-23.

Bainbridge, R. (1981, November). To err is human. Mathematics in School, 10(5). Mill, Great Britain: Longman.

Benbow, C., & Stanley, J. (1980). Sex differences in mathematical ability: Fact or artifact? Science, 210, 1262-1264.

Booth, L. R. (1981). Child methods in secondary mathematics. Educational Studies in Mathematics, 21, 29-41.

Brownwell, W. A., & Hendrickson, G. (1950). How children learn information, concepts, and generalizations. Learning and instruction, NSSE 49th yearbook, part I. Chicago: NSSE.

Burroughs, G. E. R. (1975). Design and analysis in educational research, 2nd ed.

Buswell, G. T., & Judd, C. H. (1925). Summary of educational investigations relating to arithmetic. Supplementary educational monograph, no. 27. Chicago: University of Chicago Press.

Carman, R. A. (1971). Mathematical misteaks. Mathematics Teacher, 64(2), 109-115.

Casey, D. P. (1979). An analysis of errors made by junior secondary pupils on written mathematical tasks. Unpublished master of education theses. Monash University.

Clarkson, P. C. (1980). The Newman error analysis--some extensions. In B. A. Foster (Ed.), Research in mathematics education in Australia, vol. I. Hobart: Mathematics Education Research Group of Australia.

Clarkson, P. C. (1983, November). Types of errors made by Papua New Guinean students. Educational Studies in Mathematics, 14(4), 355-367.

Clements, M. A. (1982). Analyzing children's errors on written mathematical tasks. Educational Studies in Mathematics, 11, 11-22.

Clements, M. A. (1982). Careless errors made by sixth grade children on written mathematical tasks. Journal for Research in Mathematics Education, 13, 136-144.

Fennema, E. A., & Sherman, J. (1977). Sex-related differences in mathematics achievement, spatial visualization, and affective factors. American Educational Research Journal, 14, 51-71.

Haddad, M. S. (1980). An error pattern leads to a discovery lesson. The Mathematics Teacher, 73(3).

Hall, V. M. (1980). Holiday cheer. The Mathematics Teacher, 73(9).

Hart, K. (1978). Mistakes in mathematics. Mathematics Teaching, 83, 38-40.

Hilton, T. L., & Berglund, G. W. (1974). Sex differences in mathematics achievement: A longitudinal study. Journal of Educational Research, 67, 231-237.

Kent. (1979). More about the processes through which mathematics is lost. Educational Research, 22(1).

Kent. (1978, November). Some processes through which mathematics is lost. Educational Research, 21(1), 27-35.

Laursen, K. W. (1978). Errors in first-year algebra. Mathematics Teacher, 71, 194-195.

Marshall, S. P. (1982, March). Why do boys and girls select different distracters on math problems? Paper presented at the meeting of the American Educational Research Association, New York.

Myeres, G. C. (1924, June). Persistence of errors in arithmetic. Journal of Educational Research, 19-28.

Newman, M. A. (1977). An analysis of sixth-grade pupils' errors on written mathematical tasks. In M. A. Clements and J. Foster (Eds.), Research in Mathematics Education in Australia, I, 239-258.

O'Donnel, W. J. (1980). Dividing fractions. The Mathematics Teacher, 73(9), 648.

Pallas, A. M. & Alexander, K. L. (1983, Summer). American Educational Research Journal, 20(2), 165-182.

Radatz, H. (1977). Schulerfehler und ihre moglichen ursachen in mathematikunterricht. Westermanns Padagogische Beitrage, 29, 365-369.

Rostad, A. (1971, January). Les-Tony theorem. Mathematics Teacher, 99.

Sadowski, B. R., & McIlveen, D. H. (1984, January). Diagnosis and remediation of sentence-solving error patterns. Arithmetic Teacher, 31(5), 42-45.

Schwarzenberger, R. L. E. (1984, October). The importance of mistakes. The Mathematical Gazette, 68(445).

Smith, D. E. (1895, June). Sex in mathematics. Educational Review, X, 84-88.

Teel, J. (1978), May. Response to "a new theorem." The Mathematics Teacher, 71(5).

Tubridy, A. F. (1978, January). A new theorem? Mathematics Teacher, 5.

West, T. (1971, November). Diagnostic pupil errors: Looking for patterns. Arithmetic Teacher, 467-469.

Wollman, W. (1983, May). Determining the sources of error in a translation from sentence to equation. Journal for Research in Mathematics Education, 14(3), 198-203.