A Meta-Analysis of the Effects of Modern Mathematics in Comparison with Traditional Mathematics in the American Educational System

Kuriakose K. Athappilly

Western Michigan University

Follow this and additional works at: https://scholarworks.wmich.edu/dissertations

Part of the Science and Mathematics Education Commons

Recommended Citation
https://scholarworks.wmich.edu/dissertations/2693

This Dissertation-Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Dissertations by an authorized administrator of ScholarWorks at WMU. For more information, please contact wmu-scholarworks@wmich.edu.
A META-ANALYSIS OF THE EFFECTS OF MODERN MATHEMATICS
IN COMPARISON WITH TRADITIONAL MATHEMATICS
IN THE AMERICAN EDUCATIONAL SYSTEM

by

Kuriakose K. Athappilly

A Dissertation
Submitted to the
Faculty of The Graduate College
in partial fulfillment
of the
Degree of Doctor of Education

Western Michigan University
Kalamazoo, Michigan
December 1978
ACKNOWLEDGMENTS

The present study as well as my entire doctoral program could not have been accomplished without the assistance and encouragement of many people. I wish to gratefully acknowledge my appreciation to the following people.

Dr. Uldis Smidchens, chair of my doctoral committee, Professor, Department of Educational Leadership, contributed most generously of his time, energy, and expertise as an educationist, researcher, and above all as a person of utmost understanding and singular dedication to the cause of the students. Dr. John Kofel, my former advisor and committee member, Director of International Education Center, had truly been to me "a friend, philosopher, and guide" from the beginning to the end of my doctoral program. Dr. James Riley, Professor, Department of Mathematics, provided his professional expertise for the research. To these members of my doctoral committee a debt of gratitude is most willingly, but inadequately, acknowledged.

I am extremely grateful to the authorities of Western Michigan University General Library. Although they must remain anonymous, it must be acknowledged that this research could not have materialized without their generous financial cooperation. I thank also my friends and acquaintances—students, faculty members, and administrators—who gave me their encouragement, concern, and understanding.

My mother, Mrs. Monica Athappilly, has immeasurably provided her moral and physical support to all my educational ventures. I am
grateful for her continued interest and prayers; for, in a sense, this work represents a fulfillment of one of her dreams.

My deepest appreciation goes to my wife, Sossa, and children, Geo, Geena, and Geetha. Sossa has for the past few years abandoned her profession to free me for this work. In addition to being wife and mother, she has most willingly assumed responsibilities that both parents should share. The cooperation, encouragement, and sacrifice of my family made this study a work of love.

This dissertation is dedicated to the memory of my father, Kunjuvarkey Athappilly, 1907-1967.

Kuriakose K. Athappilly
INFORMATION TO USERS

This was produced from a copy of a document sent to us for microfilming. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help you understand markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure you of complete continuity.

2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed the photographer has followed a definite method in "sectioning" the material. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.
ATHAPILLY, KURIAKOSE KUNJUVARKEY
A META-ANALYSIS OF THE EFFECTS OF MODERN MATHEMATICS IN COMPARISON WITH TRADITIONAL MATHEMATICS IN THE AMERICAN EDUCATIONAL SYSTEM.

WESTERN MICHIGAN UNIVERSITY, ED.D., 1978
# TABLE OF CONTENTS

LIST OF TABLES ........................................ vi
LIST OF FIGURES ......................................... viii

CHAPTER

I INTRODUCTION AND BACKGROUND INFORMATION .... 1
Purpose of the Study .................................. 1
The Evolution of Modern Mathematics .......... 2
The Role of Developmental Projects ............. 4
Toward Definitions .................................. 10
The Statement of the Problem .................... 11
Significance of the Study ......................... 12
Research Questions ............................... 13
Topics Excluded .................................. 14
Organization of Dissertation .................... 15

II RATIONALE FOR THE STUDY ....................... 17
Research Studies on Modern Mathematics ... 17
Techniques of Aggregation ....................... 28
Summary ........................................ 35

III DESIGN AND DATA ANALYSIS ..................... 37
Concept of Meta-analysis .......................... 37
Effect Size ..................................... 39
Techniques of Measuring Effect Sizes ............ 45
Multiple Effect Size ............................ 47
Sources of the Studies ............................. 48
Locating and Selecting Studies .................. 49

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Table of Contents—continued

CHAPTER

Identifying and Coding Variables ........ 53
Analysis of Data .......................... 55
Procedures for Judging the Magnitude
and Significance of Effect Size .......... 56
Summary ................................... 57

IV

RESULTS ....................................... 59
Description of the Studies Included ....... 59
Overall Effects of Modern Mathematics ... 62
Specific Effects by Subject
Characteristics .............................. 64
Specific Effects by Treatment
Factors ....................................... 69
Specific Effect Size by Study
Variables ..................................... 77
Specific Effects by Design Variables ..... 85
Summary ................................... 88

V

INTERPRETATIONS, IMPLICATIONS, GENERAL
RECOMMENDATIONS, AND SUMMARY ........ 92
Interpretations ............................ 92
Implications ............................... 100
General Recommendations .................. 104
Summary .................................. 107

BIBLIOGRAPHY ................................. 112
APPENDICES .................................... 132
Appendix A: The List of the Variables Identified .... 133
Appendix B: Code Sheet of the Variables Identified ... 137

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
# LIST OF TABLES

1. Overview of the Development of Mathematics Programs Since 1951 .......................... 5
2. Study I .......................................... 31
3. Study II .......................................... 31
4. Studies I and II Combined ....................... 32
5. Studies Classified by Source ..................... 60
6. Studies Classified by Year ....................... 61
7. Length of Treatment and Number of Subjects Per Study..................................... 62
8. Grand Mean Effect Sizes in Achievement and Attitude ................................. 63
9. Effect Size by Levels of Grade and Ability in Achievement ................................. 67
10. Effect Size by Socio-economic Status and Sex in Achievement ........................... 70
11. Effect Size by Content Area in Achievement ....................................................... 72
12. Effect Size by Type of Program in Achievement .................................................. 75
13. Effect Size by Length of Treatment in Achievement ............................................. 76
14. Effect Size by Number of Subjects Per Study in Achievement ................................... 78
15. Effect Size by Study Approach in Achievement and Attitude .................................... 80
16. Effect size by Source in Achievement ................................................................. 82
17. Effect Size by Subject Area of Investigation in Achievement .............................. 83
18. Effect Size by Affiliation of the Researcher in Achievement ................................... 84

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
List of Tables—continued

19. Effect Size by Design Variables in Achievement ................................ 87

20. Effect Sizes by All Categories of Variables in Achievement and/or Attitude . . . . . . . . . . 90

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Overview of the cluster approach</td>
</tr>
<tr>
<td>2.</td>
<td>Normal curves illustrating the effect of modern treatment in relation to traditional treatment—a hypothetical illustration</td>
</tr>
<tr>
<td>3.</td>
<td>Normal curves illustrating the effects of three different treatments of modern mathematics (modern algebra, modern geometry, and modern arithmetic) in relation to untreated traditional group—a hypothetical illustration</td>
</tr>
<tr>
<td>4.</td>
<td>Normal curves illustrating the aggregate effect of modern mathematics in achievement and attitude in relation to traditional mathematics group</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION AND BACKGROUND INFORMATION

Many innovations in the content, methodology, and materials related to mathematics education at the elementary, secondary, and post-secondary levels were introduced during the past three decades in an effort to make the American mathematics curriculum relevant to current needs. Studies on the effects of these innovative efforts were found primarily in dissertations, journals, and technical reports. Many curriculum changes in mathematics education were introduced under a common label called "new mathematics" or "modern mathematics." The terms "new mathematics" and "modern mathematics" were used interchangeably, since most of the literature related to the changes which occurred in mathematics curricula did not differentiate between the two terms.

Purpose of the Study

The studies under the titles "new mathematics" or "modern mathematics" were numerous and their findings were often dissimilar and conflicting (Dieudonné, 1973; Thom, 1971). Some attempts had been made to summarize and extract meaning from portions of the vast body of research data in the area of new mathematics. Recently, Hartley (1977) investigated the summative results of research pertaining to the effects of individually paced instruction in mathematics, which is a later offshoot of modern mathematics. No attempt had been made,
however, to summarize the body of data related to the various concepts encompassed within new mathematics. The purpose of this study, therefore, was (a) to summarize the effects which were unique to the new mathematics curricula and (b) to examine whether or not the new mathematics curricula led to improved academic achievement and attitude over "traditional mathematics."

The Evolution of Modern Mathematics

Toward the middle of the twentieth century, educational leaders in the colleges and universities, high school curriculum supervisors, mathematical organizations, and other interested individuals began challenging the traditional sequential pattern of mathematics teaching. In 1940, Betz cited several factors, such as general unawareness of the significance of mathematics, one-sided emphasis on the doctrine of "social utility," and mechanistic methods of teaching, as contributing to the poor status of mathematics in the schools.

In the same year, the Progressive Education Association (PEA) published a report, Mathematics in General Education, which aimed to develop a mathematics curriculum to meet the needs of the students (Bidwell & Clason, 1970, p. 531). Another report of the same year was the Place of Mathematics in Secondary Education, published by the joint commission of the Mathematical Association of America (MAA) and the National Council of Teachers of Mathematics (NCTM). The report emphasized "the goals of mathematics and the role of individual differences among students" (Bidwell & Clason, 1970, p. 532). Reports and publications of this nature asserted that the curriculum changes
in mathematics, which emphasized concepts, understanding, and insight without neglecting basic skills, were imperative to cope with the rapid advances in the field of science and technology.

World War II and Sputnik

Two critical events gave great stimuli for reform in mathematics education: World War II and the Russian launching of the first satellite (Sputnik) in 1957. World War II created a perceived need for trained manpower in scientific technology which, in turn, placed a new emphasis on improved mathematics education. The Education for All American Youth Report of 1945, the three reports of the Commission on Post-War Plans of the National Council of Teachers of Mathematics (NCTM) (1944, 1945, 1947), and The Steelman Report of 1947 were typical products of the times. All five reports reflected an increasing dissatisfaction with the content and approach of school mathematics, while calling for innovation in mathematics education.

The launching of Sputnik by the Russians and the consequent competition of the United States with Russia over space superiority created even greater pressure than World War II for the production of large numbers of qualified scientists, engineers, and mathematicians. As a result, the direction of mathematical research was toward the objective of developing mathematics curriculum adequate for the technological age rather than predictive studies toward success in the former sequential mathematics courses (Anglin, 1966).

National foundations, the federal government, and private organizations provided financial aid for curriculum improvement in
mathematics. As a result, mathematics scholars—in cooperation with educational specialists, supervisors, and teachers—began developing new mathematics programs for elementary and secondary schools. Sherman (1972) reviewed the purposes and accomplishments of all major mathematics programs established since 1951. An overview of those programs is shown in Table 1, including the year, range, and essential characteristics of 10 prominent programs.

The Role of Developmental Projects

Most of the developmental projects as seen in Table 1 were initiated primarily to correct perceived weaknesses in school programs. Consequently, the initial concerns were to produce adequate educational materials and to establish appropriate teacher training programs. The University of Illinois Committee on School Mathematics, "the progenitor of all current curriculum projects in mathematics" (NCTM, 1970, 32nd Yearbook, p. 257), was founded in 1951 to investigate problems concerning the content and teaching of high school mathematics and to correct weaknesses in that area. As a first step toward that goal, it produced materials which were pilot tested in classrooms, and it was ascertained that the success of the materials depended on the skills of the teacher.

In addition, the Madison Project placed a major emphasis on teacher training through demonstration centers and in-service programs. The three main objectives of the Greater Cleveland Mathematics Program were production of ample materials, teacher training, and evaluation. The Minnesota Mathematics and Science Teaching Project
Table 1
Overview of the Development of Mathematics Programs Since 1951

<table>
<thead>
<tr>
<th>Program</th>
<th>Year Started</th>
<th>Present Range</th>
<th>Support</th>
<th>Essentials</th>
</tr>
</thead>
<tbody>
<tr>
<td>University of Illinois Committee on School Mathematics (UICSM)</td>
<td>1951</td>
<td>Secondary</td>
<td>University of Illinois</td>
<td>Logical structure of mathematics; study of patterns</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Grades 7,8</td>
<td>USOE</td>
<td>Consistency; precise terminology</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NSF</td>
<td>Learning through discovery</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Carnegie</td>
<td>Early verbalization discouraged</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Nongraded units to be used in accordance with student background and experience</td>
</tr>
<tr>
<td>Boston College Mathematics Institute (BCMI)</td>
<td>(1953)</td>
<td>College</td>
<td>NSF</td>
<td>Structure of mathematics from the historical point of view</td>
</tr>
<tr>
<td></td>
<td>1957</td>
<td>Secondary</td>
<td></td>
<td>Cultural aspects of man's experience with numbers</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Elementary</td>
<td></td>
<td>Precise terminology</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Correspondence course with graduate credit for teachers</td>
</tr>
<tr>
<td>Ball State University (BSU)</td>
<td>1955</td>
<td>Secondary</td>
<td>Ball State University</td>
<td>Structure of mathematics</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Elementary</td>
<td></td>
<td>Interrelatedness of principles</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Logical development</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Deemphasis on social arithmetic</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Stress on mathematical ideas</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Intuitive and axiomatic approaches</td>
</tr>
<tr>
<td>Program</td>
<td>Year Started</td>
<td>Present Range</td>
<td>Support</td>
<td>Essentials</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>--------------</td>
<td>---------------</td>
<td>------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Madison Project</td>
<td>1957</td>
<td>Secondary</td>
<td>NSF, USOE</td>
<td>Discovery of structure through finding one's own solutions to problems</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Elementary</td>
<td></td>
<td>Discussion through careful questioning</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Unstructured tasks; social applications deliberately omitted</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ungraded material intended as a supplementary program</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Integration of arithmetic, algebra and geometry</td>
</tr>
<tr>
<td>University of Maryland Mathematics</td>
<td>1959</td>
<td>Grades 7,8</td>
<td>Carnegie, NSF, Funds from sale of texts</td>
<td>Mathematics as a language; precise terminology</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Elementary</td>
<td></td>
<td>Properties of a mathematical structure</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Unifying concepts; particular emphasis on number systems</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Learning through discovery</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Inductive and deductive reasoning</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Verbal and operational components simultaneously encouraged</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Fundamental learning processes a major concern of the project</td>
</tr>
<tr>
<td>Stanford University Sets and Numbers Project (SUSNP)</td>
<td>1959</td>
<td>Elementary</td>
<td>Carnegie, NSF, USOE</td>
<td>Concept of set and operations on set</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Precise language; mathematical laws</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Logic</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Relationship between set theory and foundations of arithmetic</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Algebraic and geometric principles</td>
</tr>
<tr>
<td>Program</td>
<td>Year Started</td>
<td>Present Range</td>
<td>Support</td>
<td>Essentials</td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>--------------</td>
<td>-----------------</td>
<td>------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>School Mathematics Study Group (SMSG)</td>
<td>1958</td>
<td>Secondary</td>
<td>NSF</td>
<td>Structure of mathematics</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Elementary</td>
<td></td>
<td>Concepts of mathematics as part of the whole of mathematics</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>not just to some subdivision</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Spiral approach</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>New content as well as conventional topics;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>opportunity for conventional practice and</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>review</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No strong position on verbal-operational question</td>
</tr>
<tr>
<td>Arithematic Project (UIAP)</td>
<td>1958</td>
<td>Elementary</td>
<td>University of Illinois Carnegie NSF</td>
<td>Mathematical exploration</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Topics with &quot;travel&quot;--an adventure</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Intuitive thinking; guessing; inventing; trying</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>things out</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Developing a feeling for mathematical ideas</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No stress on verbalization; not a complete</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>course of study; no grade levels</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Project seeks teachable alternatives for important mathematical ideas</td>
</tr>
<tr>
<td>Greater Cleveland Mathematics Program (GCMP)</td>
<td>1959</td>
<td>Secondary</td>
<td>The Council</td>
<td>Logical structure of mathematics</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Elementary</td>
<td></td>
<td>Mathematical laws</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Discovery approach</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Search for patterns and relationships</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Emphasis on a continuous and systematic flow of</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>mathematics concept formation K-12</td>
</tr>
<tr>
<td>Program</td>
<td>Year Started</td>
<td>Present Range</td>
<td>Support</td>
<td>Essentials</td>
</tr>
<tr>
<td>---------------------------------------------</td>
<td>--------------</td>
<td>---------------</td>
<td>---------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Minnesota Mathematics and Science Teaching Project (MINEMAST)</td>
<td>(1958)</td>
<td>Elementary</td>
<td>NSF</td>
<td>Three specific mathematics structures; the real number system, Euclidean space, a space with measure</td>
</tr>
<tr>
<td></td>
<td>1961</td>
<td></td>
<td></td>
<td>Concepts of an algebraic structure and a deductive science</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Connections between mathematics and science; cultural and historical aspects of mathematics</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Working toward a coordinated mathematics-science curriculum K-9 and undergraduate courses for pre-service education of teachers</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ungraded units</td>
</tr>
</tbody>
</table>

started with the goal of finding out what children could learn and then preparing teachers to teach it.

The Boston College Mathematics Institute was organized with the purpose of reeducating high school teachers in the elements of contemporary mathematics. The Ball State University Project was created to improve materials in secondary school geometry, and later in grades 7-12 and then in grades K-6. Summarily, improving materials and training teachers were the first steps toward introducing new mathematics in almost all the developmental projects.

Behavior modification in the minds of the pupils was effected by placing emphasis on logical thinking and structure instead of drill and rote. These concepts as described in Table 1 were introduced through different terminologies by different projects. The UICSM Project introduced the term "discovery approach." The University of Illinois Arithmetic Project stressed the ideas of thinking and concept development. The SMSG Project advocated a deeper understanding of basic concepts, structure of mathematics, and precise and sophisticated mathematical language.

The Madison Project recommended the unstructured task and the discovery approach. The purpose of the Greater Cleveland Mathematics Program was to develop a curriculum which could be presented in a logical, articulated, and sequential manner. It focused on patterns of relationships and the logical structure of mathematics. The BSU Project also gave prominence to the structure of mathematics. Creative thinking and structure of mathematics went hand in hand in almost all the developmental projects in the process of introducing new
mathematics.

There was a basic similarity in the outcomes of the various projects, although the different projects were not provided with any common set of objectives prior to the formulating of the modern mathematics programs.

**Toward Definitions**

Schaaf (1964) made a list of factors common to all the innovative projects and other activities: (a) increased emphasis on abstract ideas, (b) increased attention to logical region, (c) the use of contemporary terminology, (d) the insistence upon precise language, and (e) new "mathematical ideas" or "innovations of content."

The new ideas and content listed by Schaaf included the concepts and language of sets, Venn diagrams, systems of enumeration, base system, the real number system of algebra, modular arithmetic, inequalities, relations and properties, functions, axioms and postulates in geometry, the elements of logic, and the nature of measurement.

Ferguson (1964) pointed out that the modern mathematics programs had built-in motivation factors in that students experience the thrill of discovering mathematical principles, for example, in algebra seeing structure, order, and beauty.

According to a comprehensive analysis of six major programs for the elementary grade levels (Sherman, 1972), the common elements of new mathematics were a numeration system, measurement, geometric ideas, algebraic ideas, structure (laws and systems) sets, statistics, probability, number theory, and logic. Kline (1973) identified a set
of 12 major elements pertaining to the concepts of new mathematics. They were set theory, bases of number systems, the congruences, inequalities, matrices, symbolic logic, Boolean algebra, relations and functions, abstractions, structure, group and field, and statistics and probability.

It was, therefore, conceivable that another analysis of modern mathematics programs could produce another set of elements which were not identical with the previous sets. Hence, for the purposes of this study, the term "new mathematics" was used in referring to new developments in mathematics at the elementary, secondary, and post-secondary levels. The term applied to both new subject matter and new approaches to topics, such as those discussed above. "Traditional mathematics" referred to the content and teaching method prior to the introduction of the developmental programs.

The Statement of the Problem

"There can be little question that following a fifteen year period of unprecedented favor and prominence in the curricula of American schools, mathematics teaching today faces the future with less certainty of its goals and less optimism about its potential effectiveness," observed Hill (1969, p. 440), who chaired the National Advisory Committee on Mathematics Education.

The uncertainty about mathematics curricula stems from strong public criticisms and from the abundance of research findings of which some are dissimilar while others are conflicting in nature (Norland, 1971; Yasui, 1967). Because of the dissimilar and
conflicting nature of study findings, decision making at the levels of resource allocation, curriculum development, and classroom teaching becomes difficult and often misguided. Moreover, the money, time, and thought expanded on new mathematics have been considerable—perhaps even enormous (Kline, 1973).

These facts about the conflicting nature of research findings and public opinions about the outcomes of new mathematics, coupled with the scarcity of summative research resolving dissimilar and conflicting findings, formulate the problem for which this study was intended.

Significance of the Study

The conflicting nature of past findings made the study highly relevant and important. The study, which was a meta-analysis—an analysis of analyses—was precisely designed to resolve outcomes which were not in agreement with each other. Conclusive findings about the outcomes of modern mathematics were focused on three major areas. These areas were (a) the subject—the importance of an appropriate mathematics curriculum, (b) the methodology—the utilization of a meta-analysis technique to discover the message of the new mathematics in the midst of conflicting findings, and (c) the two levels of decision making—balanced resource allocation decisions and well-founded curriculum decisions yielding increased learning.
Research Questions

In order to summarize the effects of new mathematics and to examine whether or not the new mathematics curriculum led to improved academic achievement, the research sought answers to the following questions:

1. Did the aggregation of studies show any evidence of change in academic achievement attributable to new mathematics programs? This question was concerned with the overall impact of the modern mathematics as evidenced by the mean effect size of achievement when all achievement studies were combined.

2. Did the aggregation of studies show any evidence of attitudinal change attributable to new mathematics when compared to traditional mathematics programs? This was concerned with the grand mean effect size of attitude toward mathematics when all attitudinal studies were combined.

3. What was the differential effect of the new mathematics programs with respect to the following variables?

Subject characteristics:
- Grade
- Socio-economic status
- Ability
- Sex

Treatment factors:
- General content area
- Specific subject area
- Type of program
- Length of treatment
- Number of subjects per study
Study variables:

- Study type
- Study approach
- Author
- Source of study
- Affiliation of researcher
- Involvement of researcher
- General subject area

Design variables:

- Type of test
- Quality of design
- Novelty effect

Topics Excluded

In addition to the previously discussed projects, there were other developmental efforts that introduced innovative concepts in mathematics education. Some of these projects were excluded from this study. Primarily, they fell into one of the three following categories.

1. In the first category were projects which pertained to programmed instruction, individual learning packets, tutoring, and computer assisted instruction. The emphasis of these projects was purely on the methodology of teaching with little or no relationship to the new mathematics concepts as discussed in the definitions.

2. In the second category of excluded studies were those which posed a problem of definition. They pervaded the area of cognitive learning styles, pedagogical techniques, and teacher training. Existential approach, self supervision, hierarchical components, modeling theory, and resource personnel were some of them.

3. In the third category were the studies which were strictly
within the area of new mathematics, but were excluded for technical reasons, such as inadequate methodology, inappropriate design, and lack of pertinent quantitative data which were necessary for the purposes of comparison.

Organization of Dissertation

Chapter I discussed the purpose and significance of the study. It traced the evolution of modern mathematics and the role of the developmental projects. The chapter also addressed itself to the definition of terms, statement of the problem, research questions, and the topics excluded.

Chapter II reviews research studies on modern mathematics and techniques of aggregation. The section on research studies on modern mathematics has been subdivided into three categories: (a) appraisals of modern mathematics, (b) research studies on modern mathematics by primary analysis, and (c) research studies on modern mathematics by aggregate analysis.

Chapter III presents detailed description of the concept of meta-analysis, sources of the study, methods of locating and collecting studies, the process of identifying and coding the variables, and the techniques of measuring effect sizes. It also deals with the question of multiple effect sizes and the method of analysis of data.

Chapter IV reports the results of the data analysis. The results are in two major categories: (a) the overall effects and (b) the specific effects by different groups of variables, such as subject variables, treatment variables, study variables, and design variables.
Chapter V interprets the results of the study reported in Chapter IV, discusses the implications and the major recommendations based upon the research data, and gives a summary of the study.
CHAPTER II

RATIONALE FOR THE STUDY

The purpose of the chapter is to present the logical basis for the study through the review of the literature which focuses upon the area of the investigation. In this light, it may be expedient to devote a separate section of the chapter to the justification of the methodology, since the study utilized a new technique of aggregation—a meta-analysis procedure—as its research tool. Subsequently, the rationale for the study focuses upon two categories of topics: (a) research studies on modern mathematics and (b) the techniques of aggregation.

Research Studies on Modern Mathematics

To begin then, the effects of the new mathematics curriculum were reported in literally hundreds of research studies. A comprehensive treatment of the total studies conducted was not attempted in this chapter due to the number of studies being prohibitively large; however, a summary description of the studies on modern mathematics is given in the ensuing sections. The description consists of three parts: (a) appraisals of modern mathematics, (b) studies of modern mathematics by primary analysis, and (c) studies of modern mathematics by aggregate analysis.
Mathematics educators were divided in their opinions, points of view, and observations regarding the efficacy of the new mathematics programs. Supporters of modern mathematics advocated the relevance, appropriateness, and the distinct advantages of the new programs.

Fehr (1953) urged the adoption of the new mathematics because of its relevance to understanding the universe. In the same vein, MacLane (1954), Hartung (1955), Price (1957), Kennedy (1961), Mueller (1962), and Davis (1963) advocated the introduction of new mathematics programs at the elementary and secondary levels in order to meet the changing needs created by the advancement of science and technology.

Writers have also observed more specific advantages of the new programs. For example, Clark (1961) pointed out that according to the findings of Begle, Glennon, Gundlach, and others, the new mathematics programs permitted and even encouraged children to learn more mathematics at an earlier age. Similarly, Grossnickle (1964) suggested that the advantage of modern mathematics was in the introduction of new and more profound topics at all grade levels. Furthermore, some writers were of the opinion that slow learners benefitted more through increased learning when utilizing new mathematics programs rather than the traditional programs (Cunningham, 1965).

Conversely, the opponents of the new mathematics programs expressed different and opposing points of view. For example, Rosskopf (1954) objected to the new movement for a complete reorganization from traditional to contemporary mathematics instruction.
primarily because he believed very few teachers were competent to teach it. Similarly, Pingry (1956) and Kline (1959) criticized the outright revision of the mathematics content as the proper approach for improving student learning. Additionally, Kline (1958) echoed these sentiments when he said, "The modern mathematics stressed during the past few years is trivial in importance compared to topics currently taught and is hopelessly beyond young people, because of its abstractness and rigor" (p. 422).

About the same time other writers like Hannon (1959) believed that the traditional approach to teaching and the related materials had to form the basic core of the elementary curriculum. Several years later Rappaport (1962) expressed alarm over the haste with which untrained educators desired to teach modern mathematics. Also, Rappaport held a concern for the impracticality of the new math programs for the majority of the children.

It was apparent from the preceding sampling of viewpoints that there were conflicting appraisals regarding the worth of the modern mathematics programs. Again, the primary purpose for the study was to seek a resolution to the maze of conflicting and dissimilar findings about modern mathematics as illustrated by the examples which were cited.

Studies on Modern Mathematics by Primary Analysis

A review of the research on modern mathematics revealed two important aspects pertaining to the nature of the studies. One, in
attempting to study the effectiveness of new mathematics programs, investigators have adopted diverse methodologies for trying to gain a clearer perspective. Two, in spite of several attempts to gain a clear perspective, the study findings showed sharp contrasts and dissimilarities. The ensuing section seeks to point out the conflicts and disagreements among the study findings of the several different designs.

The most common methodologies utilized to study new mathematics ranged from simple descriptions to true experimental research designs. For example, Davis (1965) and Johntz (1967) used the descriptive process to investigate the effectiveness of the Madison Project. They concluded that the creative classroom experience of the new programs was superior to the traditional mathematics experience. Most of the studies compared the new programs with traditional ones by reviewing data from tests given to both experimental and control groups. Biddle (1967) used a course in programmed instruction and Berger and Howitz (1967), a discovery program for general mathematics students. After one year of study, both investigations found that there was "no significant difference" between the experimental and control groups.

Some researchers compared the effects of the experimental and the traditional programs over a longer period of time. Hungerman (1967) found that the School Mathematics Study Group (SMSG) students were superior to traditional students on a conventional test. Graft and Ruddell (1968) found a significant difference favoring the SMSG students in understanding, thought process, and transfer of learning.

Several studies compared the effectiveness of the treatment at
various levels in relation to factors such as ability, intelligence, and sex. Williams (1963) found no significant difference in achievement, in grade nine, between the traditional algebra text and the SMSG materials when used by average, high-average, and high ability students. Shuff (1962) noted that at the upper and middle ability levels, the students who received the traditional mathematics treatment in the seventh grade made better achievement gains than the students who received the SMSG instruction. Conversely, at the lower level, the SMSG students performed better than the traditional group.

Peck (1963) divided the sixth grade students studying modern mathematics and traditional mathematics into two groups. The high intelligent group consisted of students with IQ's of 115 or more and the low intelligent group consisted of students with IQ's less than 115. The findings indicated that there was no significant difference in achievement between the students who received modern mathematics and traditional mathematics at the high IQ levels or at the low IQ levels. About the same time, Harshmon (1962) made a comparison of programs, using modern manipulative materials and traditional materials. Groups receiving each program were subdivided into three intelligence levels: (a) those with IQ's of 99 and below, (b) those with IQ's between 100 and 125, and (c) those with IQ's of 125 and above. The results significantly favored the traditional program in the IQ subgroup of 100-125, but the opposite was true for the IQ ranges of 125 and above and 99 and below.

Wozencraft (1963) carried out a comprehensive study to determine the difference between boys and girls in arithmetic ability. She
found that there was a significant difference in arithmetic ability which favored the girls in total and average groups of children; however, the groups with high or low intelligence quotients did not show significant sex differences. Northcutt (1964) examined the comparative achievement of fifth grade boys and girls in two teaching approaches, a traditional approach and a programmed (modern) approach. He found no significant difference in gain between the sexes using either approach.

Longitudinal studies were yet another category of research in this area. Yasui (1967) reported that the mean achievement score of modern mathematics was significantly higher than the mean score of students of traditional mathematics. Conversely, the findings of Osborne (1965) revealed that the introduction of SMSG materials and their study for increasing period of time had "no significant effect" on arithmetic skills, algebra skills, and mathematics reasoning skills. The findings of Austin and Prevost (1972) confirmed Osborne's findings, but contradicted those of Yasui: After a four year experiment, the traditional mathematics group scored higher than the transitional and modern mathematics groups on the Metropolitan Achievement Test of Mathematics Computation.

In another longitudinal study, Norland (1971) observed that after four years of instruction in modern mathematics, achievement in arithmetic computation declined significantly in four of the five schools which participated in the study. Conversely, in 1973 the National Assessment of Educational Progress reported quite different findings.

One question about the new mathematics programs of the
1960's has been their impact on pupil's ability to compute . . . . But the evidence, limited though it may be, sug­gests that Johnny can add: computation with whole numbers is far from being a lost art. The performance of the thir­teen year olds on each of the four operations with the whole numbers was almost as good as the performance of the seventeen year olds and the adults . . . . The thirteen year old's responses indicate that computational algorithms for all four operations are well developed by the latter part of elementary school. (p. 457)

Thus, the findings of the research studies fell into three cate­gories: (a) findings showing no significant difference between the traditional and modern mathematics groups, (b) findings favoring the traditional mathematics, and (c) findings favoring the modern math­ematics. The findings, therefore, were so conflicting and dissimilar in nature that one could not make a sound generalization about the effects of new mathematics programs without a further investigation into the results of those studies conducted with a more appropriate technique of analysis.

Studies on Modern Mathematics by Aggregate Analysis

Studies on modern mathematics by aggregate analysis were exceed­ingly small in number. Some attempts were made in the generic area of mathematics to aggregate the studies and classify them under one or two major categories according to source, topic, or period of pub­lication; however, no analysis was made in any of the attempts to summarize the findings or draw any conclusions. Several of these attempts are summarized in the following paragraphs.

In the April 1957 issue of the Arithmetic Teacher, a bibliography appeared which included six years of research on arithmetic teaching

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
(1951-1956). The annual listing of research pertaining to elementary school mathematics continued in the *Arithmetic Teacher* for the next 14 years. In 1971 the *Journal of Research in Mathematics Education* began publishing a comprehensive listing of research which had been conducted in elementary and secondary mathematics education.


In 1972, Suydam published a *Review of Research on Secondary School Mathematics*. The review consisted of 780 research reports and 770 dissertations published between 1930 and 1970. The combined 1,550 documents were each categorized and annotated, but no effort was made in the reviews to summarize the findings nor draw any conclusions. The annual updates of the review's original bibliography appeared in the subsequent editions of the *Journal for Research in Mathematics Education* (Suydam & Weaver, 1973, 1974, 1975, 1976, 1977).

Collectively, the published annual bibliographies were— for the purposes of this study—the most valuable source for locating the majority of previous research studies. At the same time, it should be noted that no attempt was made in any of the publications to analyze the aggregate results of the research.

The first known attempt to analyze and synthesize the results of studies pertaining to selected areas in the teaching of arithmetic was made by Bartram in 1956. Some of the relevant areas of his investigation were problem-solving procedures, general teaching methods,
instructional materials, and provision for individual needs. Bartram selected 200 studies out of the 600 which were available. Bartram's criteria for selection of the 200 studies were (a) the restriction of experimental factors; (b) the control of variations in pupil factors; (c) the control of nonexperimental factors, such as instructional methods and materials; (d) the validity and reliability of the measuring instrument; (e) the adequacy of the sampling technique; and (f) the statistical significance of the findings. The selection criteria, employed by Bartram, especially statistical significance, resulted in the exclusion of several studies which contained valuable information.

Bartram followed three procedures in treating each area: (a) identification and analysis of points of view concerning the area, (b) detailed presentation of research, and (c) forming conclusions and suggesting teaching practices. His more significant conclusions in these areas are summarized as follows:

1. Arithmetic problem solving was learned better through the medium of an activity program than through one more rigidly organized; also, computational skills were learned equally well. In addition, activity programs contributed more to the child's social and emotional adjustment.

2. Inductive methods were superior to deductive for teaching process and generalizations.

3. Children profitted from meaningful practice. Such practice should be motivated by a social need and should follow the acquisition of understanding.

4. A combination of the activity method and grouping within the
classroom appeared to be the most satisfactory means of meeting the needs of the individual children.

Bartram's (1956) research focused on a critical examination of teaching practices, rather than on an evaluation of a new program. Consequently, Bartram's findings were not directly related to the present study.

A second effort was undertaken by Hartley in 1977, as mentioned in Chapter I. Hartley (1977) collected and synthesized 153 experimental studies which pertained to the efficacy of four techniques of mathematics instruction. Hartley utilized a meta-analysis methodology for conducting an aggregate analysis of the results.

The four techniques Hartley examined were computer-assisted instruction, cross-age and peer tutoring, individual learning packets, and programmed instruction. The effectiveness of each technique was determined by comparing the resulting mathematics achievement of students taught by the particular technique with the achievement of students taught by a traditional method of instruction. In addition to teaching methodology, other variables such as design characteristics, subject variables, and attributes of instructional technique were quantified. The resulting relationships to effectiveness of each technique were also examined.

Briefly, Hartley found, on the basis of the studies collected, that tutoring was the superior technique for increasing mathematics achievement. Peer tutoring, in which classmates tutor one another, was as effective as paid adult aides; however, cross-age tutoring, in which older students tutor younger students, was slightly better than
the peer and adult tutoring. Also, Hartley discovered that intensive supervision and instruction of the student tutors did not result in an increase in effectiveness, although some instruction was beneficial.

In addition, Hartley noted that computer-assisted instruction was less effective than tutoring in increasing the effectiveness of mathematics instruction, although it was considerably more effective than either individual learning packets or programmed instruction. Individual learning packets and programmed instruction were essentially comparable to traditional instruction; that is, these two techniques resulted in little or no achievement gain over traditional instruction. In many cases, traditional instruction was definitely superior to individual learning packets and programmed instruction.

The Hartley (1977) study was a later off-shoot of the innovative programs in mathematics and it did not strictly come under the definition of new mathematics given in Chapter I. Consequently, the results of the study were not directly applicable to the purposes of the present study.

It was clear, therefore, that no comprehensive attempt had been made to review the large volume of research studies on the effectiveness of modern mathematics by any aggregate technique of analysis. The present study attempted to utilize an appropriate technique of analysis to resolve the conflicting and dissimilar findings on modern mathematics programs.
Techniques of Aggregation

One of the early techniques of aggregation of studies was the narrative. The narrative was mainly a chronologically arranged verbal description; however, this technique failed to portray the accumulated knowledge of large numbers of studies due to the cumbersome length of the resulting narrative. As a consequence, the review of literature was created as an improved technique of extracting information from groups of studies.

The review of literature usually involved three processes. They were: (a) gathering the relevant empirical studies, (b) discarding studies with inadequate empirical designs, and (c) drawing conclusions from the remaining studies. In order to draw conclusions, the reviewers often made use of a listing of studies with a tally kept of the number of significant and nonsignificant results. Conclusions favored the results found in the larger number of studies in case of dissimilar or conflicting findings.

One of the drawbacks of the review of the literature was that it did not take into account the influence of the variables that could affect the results of a study in a systematic way. For example, some of these variables were the quality of design, ability level, age and sex of subjects, length of treatment, and date and source of study. A second drawback was that charts and tables were often used by reviewers to differentiate the results; however, when encountering large numbers of studies, the charts and tables proved to be inadequate means to fulfill the purpose.
A third drawback of the review of research was that it often eliminated the "poorly done" studies. Glass (1976) said that it was an empirical question whether the so-called "poorly done" studies gave results significantly at variance with those of the best designed studies. He believed that because the difference was so small, the integration of research results through eliminating the "poorly done" studies would result in discarding a vast amount of important data.

In addition to the narrative and the review of the literature, other techniques of aggregation have been utilized for summarizing accumulated data. Averaging, voting, and the cluster approach were the three major methods Light and Smith (1971) explained in their paper, "Accumulating Evidence." These three methods are briefly discussed in the ensuing paragraphs.

The Averaging Method

The averaging method was a more systematic approach than mere listing and simple narratives of the studies. The approach consisted of computing overall means for relevant statistics for a large number of studies. The measures of central tendency were used to compute the overall average which was used in the averaging method. The method could be considered to be the first attempt of quantification, but the weakness was that it threw away precisely the information needed most, namely, a true estimate of the treatment effect.

The Voting Method

A more systematic and widely used approach to aggregation was the
voting method. In the approach, all studies showing a relationship between a particular dependent variable and an independent variable were examined for their significance. The studies were classified under one of three categories of relationship—significantly positive, significantly negative, or neither significantly positive nor significantly negative. Tallies were made for each of the three categories. The category that had the maximum number of tallies was assumed to give the best estimate of the direction of the true relationship between the two variables under consideration.

The weakness in the method was threefold. First, valuable descriptive information was disregarded. Secondly, the method did not integrate measures to determine the strength of experimental effects or relationships among variables (according to whether the problem was basically experimental or correlational). Third, aggregation by voting did not take into account the possibility of Simpson's Paradox (Glass, 1977). Perhaps an example would illustrate the paradox.

Assume two studies on the effects of individualized instruction compared with the traditional method were made in regard to achievement. In Study I, 50 students received individualized instruction and 40 students received traditional instruction. In Study II 50 subjects were in the individualized instruction group and 300 were in the traditional instruction group. Tables 2 and 3 show the number of passes, failures, and the pass rates in each study. The pass rate for individualized instruction exceeded the traditional by 42% versus 40% in Study I. In Study II the pass rate for individualized instruction exceeded the traditional by 56% versus 55%. Utilizing the voting
method, the score would be 2-0 in favor of individualized instruction; however, an opposite result would be attained if one aggregated the raw data.

Table 2
Study I

<table>
<thead>
<tr>
<th></th>
<th>Individualized</th>
<th>Traditional</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pass</td>
<td>21</td>
<td>16</td>
<td>37</td>
</tr>
<tr>
<td>Failure</td>
<td>29</td>
<td>24</td>
<td>53</td>
</tr>
<tr>
<td>Total</td>
<td>50</td>
<td>40</td>
<td>90</td>
</tr>
<tr>
<td>Pass Rate</td>
<td>42%</td>
<td>40%</td>
<td></td>
</tr>
</tbody>
</table>

Table 3
Study II

<table>
<thead>
<tr>
<th></th>
<th>Individualized</th>
<th>Traditional</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pass</td>
<td>28</td>
<td>165</td>
<td>193</td>
</tr>
<tr>
<td>Failure</td>
<td>22</td>
<td>135</td>
<td>157</td>
</tr>
<tr>
<td>Total</td>
<td>50</td>
<td>300</td>
<td>350</td>
</tr>
<tr>
<td>Pass Rate</td>
<td>56%</td>
<td>55%</td>
<td></td>
</tr>
</tbody>
</table>

In Table 4 the pass rate was reversed. The traditional exceeded the individualized instruction by 53% versus 49%. The cause of this paradoxical situation was the problem of unbalanced experimental

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
design. Since the voting method did not pay attention to the sample size, there existed the possibility of misinterpretation of results.

Table 4

<table>
<thead>
<tr>
<th></th>
<th>Individualized</th>
<th>Traditional</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pass</td>
<td>49</td>
<td>181</td>
<td>230</td>
</tr>
<tr>
<td>Failure</td>
<td>51</td>
<td>159</td>
<td>210</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>340</td>
<td>440</td>
</tr>
<tr>
<td>Pass Rate</td>
<td>49%</td>
<td>53%</td>
<td></td>
</tr>
</tbody>
</table>

The Cluster Approach

As a more refined technique, the cluster approach had been elaborated upon by Light and Smith (1971) in an attempt to devise a technique for resolving contradictions among different studies. The approach stemmed from a pragmatic insight that many populations could be broken down into small, identifiable subpopulations which were called clusters. The clusters were not random samples from a population. Rather, they were natural aggregations within the population. The clusters usually differed in broad and systematic ways. Thus, a study could contain several clusters according to the various aggregations within the population.

The main difference between the cluster approach and the voting method was in the selection of the unit of analysis. In the voting
method the unit of analysis was a complete study. Conversely, in the cluster approach the unit of analysis was a cluster which could be a part of a study or a complete study.

The primary steps involved in the method of aggregation consisted of access to the original data, review of the quality of the studies, tests for differences among clusters, and classification and combination of the studies according to the availability of the explanation of differences.

Figure 1 gives an abbreviated overview of the basic logic of the cluster approach. The first block identified the five kinds of cluster differences for which one tested. If all the five tests failed to find differences, data from the several clusters could be combined. If few of the tests indicated differences, an explanation of these differences was sought. The data were combined after adjustments had been made to eliminate the differences. If no explanation was found for the differences, data could not be combined.

The cluster approach presented problems in its application. Since the cluster approach required access to the raw data, the study became overly restrictive. The approach also eliminated several studies because of the statistical and methodological restrictions. The subjects had to be selected from a known and precisely definable population. The dependent variables and the independent variables which were measured had to have a common measure or be subject to suitable conversion methods. The overall quality of the studies selected had to be comparable. Another serious problem was the inability to draw any conclusion in the absence of explanations for
Test for differences among clusters in:
(a) Means
(b) Variances
(c) Covariate relations
(d) Subject-treatment interactions
(e) Contextual effects

One or more differences found

Search for explanation:
(a) Selection effects
(b) Amplification effects
(c) Sensitization
(d) Different proportions of types of subjects in clusters
(e) Contextual differences
(f) Unmeasured variables
(g) Other

No explanation found

Adjust away explained cluster differences

Explanation found

Combine data from clusters

Cannot combine data from cluster

Figure 1. Overview of the cluster approach. (From "Accumulating Evidence: Procedures for Resolving Contradictions Among Different Research Studies" by R. J. Light and P. V. Smith, Harvard Educational Review, 1971, 41, 463.)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
differences. These restrictive criteria, therefore led to a loss of information.

Conversely, the present study attempted to include as many empirical findings as possible in order to analyze and summarize the voluminous body of research data pertaining to new mathematics curricula. The intended result was to determine the meaning of these findings without the loss of information. In this regard, a meta-analysis technique, as explained by Glass (1976), was adopted for the study. The primary reason for utilizing meta-analysis was that it diminished the weaknesses contained in the aforementioned techniques in many respects. The concept of meta-analysis is thoroughly discussed in Chapter III.

Summary

Chapter II dealt with rationale for the present study. The focus of the rationale was on research in modern mathematics and techniques of aggregation. The section pertaining to research on modern mathematics was subdivided into three areas: (a) appraisals of modern mathematics, (b) studies on modern mathematics by primary analysis, and (c) studies on modern mathematics by aggregate analysis. The discussion on appraisals of modern mathematics revealed that there existed conflicting viewpoints regarding the efficacy of modern mathematics. Since the studies on modern mathematics by primary analysis were numerous and their findings were diverse, an aggregate analysis was necessary to draw a conclusive meaning from the findings. The number of studies on modern mathematics by aggregate analysis was
small. Among the few existing ones, none focused upon the area of investigation of the present study.

In the discussion of the techniques of aggregation five methods were identified. They were narrative, reviews, the averaging method, the voting method, and the cluster approach. Strengths and weaknesses of each technique were discussed. The examination of these techniques revealed the necessity for a better technique of analysis for the purposes of this study. The present attempt, therefore, was the first aggregate analysis of the effects of modern mathematics by using a meta-analysis technique.
CHAPTER III

DESIGN AND DATA ANALYSIS

The purpose of Chapter III is to describe the theoretical and practical aspects of meta-analysis methodology. The theoretical aspect consists mainly of the explanation of the effect size and the techniques of calculating effect sizes. The practical aspect deals with identifying the sources, locating and selecting the studies, identifying and coding the variables, and the procedures for analyzing the data.

Concept of Meta-analysis

An aggregate analysis was deemed to be appropriate to resolve the dissimilar and conflicting nature of the study findings about the new mathematics. Many of the techniques of aggregation used in the past were inadequate for the purposes of the present study. A meta-analysis procedure, introduced by Glass (1976), was found to be an improvement on the techniques discussed in Chapter II.

In an attempt to explain the concept of meta-analysis, Glass (1976) made a distinction among primary, secondary, and meta-analysis. Primary analysis was defined as the original analysis of data in a research study. Secondary analysis was thought of as the reanalysis of data for the purposes of answering the original research question with better statistical techniques. Also, secondary analysis could be utilized for answering new questions with old data. Conversely,
meta-analysis was the analysis of analyses. In the present study, meta-analysis, as defined by Glass, was adopted as an alternative technique to analyze the aggregate findings on the effects of modern mathematics.

The literature on meta-analysis consisted mainly of the writings of Glass. Two of his papers which specifically focused on meta-analysis were "Primary, Secondary, and Meta-analysis of Research" (Glass, 1976) and "Integrating Findings: The Meta-analysis of Research" (Glass, 1978). The former provided an overview of the meta-analysis methodology and the latter pertained to the different techniques involved in the meta-analysis procedure.

Glass explained meta-analysis as a statistical technique which analyzed a large collection of analysis results from individual studies for the purpose of integrating the findings. The method connoted a rigorous alternative to the casual narrative discussions of research studies which typified attempts to make sense out of the rapidly expanding body of research literature. A special feature of Glass' meta-analysis technique was that it statistically aggregated, rather than merely listed, the results of a large number of studies. Also, it used summary statistics instead of raw data for aggregation.

One of the advantages of meta-analysis was that conflicting or dissimilar studies were not eliminated on the basis of a lack of explanations for differences. Large numbers of conflicting studies were combined to yield an overall conclusion by calculating what Glass called the "effect size" (Glass, 1976). Theoretically, effect size as used in this study was a standardized measure of the
effectiveness of the treatment when compared to an untreated control group.

Effect Size

To elaborate, the key concept in Glass' meta-analysis methodology was the effect size. The effect sizes were calculated for each study in a measure that was common for all studies. Then the measures were combined to yield more information than just a significant-not-significant count. The result indicated the difference between the effect of a treatment and its comparison treatment. Thus, the effect size yielded information pertaining to the specific treatments in the study. Specifically, each effect size measured how much more or less effective the modern mathematics was than the traditional mathematics counterpart.

Effect size was defined as the difference between the means of the experimental and control groups, divided by the standard deviation of the control group (Glass, 1978). The formula was as follows:

\[
\text{Effect size} = \frac{\text{mean of treatment} - \text{mean of control}}{\text{standard deviation of the control group}} \quad (1)
\]

Symbolically, the formula can be written as:

\[
ES = \frac{\bar{X}_E - \bar{X}_C}{S_X} \quad (2)
\]

where \(\bar{X}_E\) = the mean of the experimental group

\(\bar{X}_C\) = the mean of the control group
\[
\frac{S}{X} \quad \text{the standard deviation of the control group}
\]

and \( ES \) = the effect size

A positive measure of effect size indicated that the findings favored the treatment group while a negative measure indicated that results supported the control group. For example, if the effect size was \( k \) standard deviations, then an average person in the control group was \( k \) standard deviations below the average person in the treatment group.

By dividing the difference of means by the standard deviation, the effect sizes were standardized and were readily converted into z-score units. Consequently, the effect sizes were appropriate for comparison even though they were derived from different studies. Also, corresponding to each effect size, a percentile score could be obtained from the table which indicated the areas and ordinates of the unit normal distribution.

For example, let the effect size comparing the effectiveness of two treatments—a traditional and a modern—be 0.49. Then, the value 0.49 indicated that an average person receiving the modern treatment had been raised from the 50th percentile to the 69th percentile of the traditional group, as depicted in Figure 2.

In Figure 2, the normal curve of the dotted lines represented the control group or the traditional group. An average student in this group occupied the midpoint of the baseline. This point coincided with the zero point of the z-score. The average person, therefore, was said to be at the 50th percentile of this normal curve,
CONTROL GROUP (TRADITIONAL)  

TREATMENT GROUP (MODERN) 

Average Effect Size is 0.49 

Figure 2. Normal curves illustrating the effect of modern treatment in relation to traditional treatment—a hypothetical illustration.
which represented the traditional mathematics population. The normal curve of the solid line represented the treatment group.

The assumed value of the effect size in Figure 2, namely 0.49, being a measure of the difference between the means, was represented by the distance between the midpoints of the two normal curves. The midpoint of the curve which represented the modern group coincided with the 69th percentile of the traditional group. (The percentile value corresponding to the z-score of 0.49 was 69.)

Thus, given an effect size of 0.49, a person who was at the 50th percentile of the modern group was, at the same time, at the 69th percentile of the traditional group. In other words, the treatment of the modern program under the average conditions could be expected to move an average student from the 50th to 69th percentile of the traditional mathematics population, if the mean effect size was 0.49. The same explanation was applicable to all specific mean effect sizes, except for the fact that the effect sizes were different in each particular case.

The mean effect size was a powerful tool in summarizing the results. The findings could be summarized at the general as well as specific levels into which the independent variables of this study were categorized, by using various measures of mean effect size. At the most general level, average of all the effect sizes were found. This was called the "grand mean effect size" or the "overall mean effect size." The terms "mean" and "average" were used interchangeably.

In the present study, modern mathematics population represented
the treatment group and the traditional mathematics population, the control group. The outcome variable was either achievement or attitude measured in terms of effect size.

The hypothetical effect sizes of three different subject areas are compared in Figure 3. In all three cases, a modern mathematics group was compared with the traditional mathematics group. Since the traditional mathematics group was not given any treatment in modern mathematics, it was considered to be an untreated control group whose average effect size was always zero. The four normal curves, therefore, represented the typical three treated populations in relation to the untreated control subjects in the subject areas of arithmetic, algebra, and geometry. For example, 102 effect size measures from modern algebra studies averaged about 1.18 standard deviation.

The major impression derived from Figure 3 was that modern algebra and geometry were not substantially different from each other in their average impact on learning. Also, one could deduce that both modern algebra and geometry groups performed better than the traditional mathematics group.

In addition, data from Figure 3 illustrated that the children who received the modern arithmetic treatment performed worse than the traditional mathematics group. Thus, the various measures of effect size illustrated the comparative effectiveness of different modern mathematics treatments among themselves as well as their comparative effectiveness with the traditional mathematics treatment.
Figure 3. Normal curves illustrating the effects of three different treatments of modern mathematics (modern algebra, modern geometry, and modern arithmetic) in relation to untreated traditional group—a hypothetical illustration.
Techniques of Measuring Effect Sizes

Different study designs required different techniques to measure the numerator and the denominator in the equation of the effect size. Glass (1978) discussed many techniques in his paper, "Integrating Findings: The Meta-analysis of Research." In the ensuing sections the techniques which were used for the present study are enumerated.

Techniques for determining the numerator were as follows.

1. The numerator of a typical experiment (in which the subjects were randomly assigned to treatment and control groups, or the pretest scores indicated that the groups were equivalent) was the difference between the means of the posttests. The pretest scores did not enter into the computation.

2. If the design made use of analysis of covariance, and if the adjusted means were reported, then the numerator was the difference between the adjusted means.

3. If a design reported the gain scores—pretest scores subtracted from the posttest scores—then the numerator of the effect size was the difference between the averages of the two gain scores.

4. If a design reported only grade equivalent scores from a standardized test, then the numerator was calculated by the length of the treatment. This procedure assumed that theoretically an untreated average group should gain one point for each month in the school. Thus, if a study indicated a pre-post difference of 2.0 grade equivalents after the treatment of eight months, the numerator of the effect size was 2.0 - (1 x 0.8) = 1.2.
Correspondingly, the methods for determining the denominator of the effect size were as follows.

1. Whenever it was possible to obtain the standard deviation of the control group, then that was used as the denominator of the effect size.

2. If only the means based on scores from a standardized test were given, the standard deviation of the norm group was taken as the denominator of the effect size.

3. If only the t-statistic was available, then effect size was derived from the formula:

\[
ES = t \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}
\]  \hspace{1cm} (3)

where

- \( n_1 \) = the number of subjects in the control group
- \( n_2 \) = the number of subjects in the treatment group

The equation was derived from the formula (2) given earlier and from the equation for the t-statistic, which was given by the formula:

\[
t = \frac{\bar{X}_E - \bar{X}_C}{\sqrt{S_X^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}}
\]  \hspace{1cm} (4)

Formula (2) was:

\[
ES = \frac{\bar{X}_E - \bar{X}_C}{S_X}
\]

By combining formulas (2) and (4), formula (3) was derived.
4. If a report contained only the information that the mean of the experimental subjects exceeded the mean of the control subjects at certain \( \alpha \)-level significance, then an approximate value of the effect size was calculated, assuming the t-value for the particular \( \alpha \)-level. For example, if \( \alpha = .05 \), then the most conservative value of \( t \) was equal to 1.96. The value 1.96 was obtained from the table showing the percentile points of t-distributions in standard textbooks dealing with statistics. By substituting this value for \( t \) in formula (3) the effect size was obtained from the following formula:

\[
ES = 1.96 \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}
\]  

(5)

5. If the report indicated only an F-value and there were only two groups, the positive square root of F had to be substituted for \( t \).

Glass (1978) discussed other techniques, such as probit analysis and logit transformation, in his paper, "Integrating Findings: The Meta-analysis of Research." However, these techniques were not required in the present study since all of the data reported were readily convertible to effect size by the previously discussed techniques.

Multiple Effect Size

In the process of calculating effect sizes, care was taken to prevent the problem of multiple effect size. The problem of multiple effect size arose when a study had more than one effect size. When a study had more than one effect size, certain effect sizes could be
repeated either in part or whole. For instance, if a study dealt with two grade levels, then effect sizes could be found for each grade level separately and for the entire group combining the two grades. Thus, three effect sizes were possible from a study consisting of two grade levels. If all three effect sizes were considered for the grand mean effect size, the study would be contaminated by the multiple effect size, because the effect sizes of each grade level and the effect size of the combined group represented the same value. The problem could be avoided by choosing either the two effect sizes obtained from each grade levels—one from each grade—or the effect size of the combined group, but not both sets.

The problem of multiple effect sizes became more acute when the design utilized was a factorial study involving sex or ability. For example, if a study included four grades and sex as a factor, then it would have four effect sizes on grade, one for each grade, and two on sex. In the overall analysis of studies, either the first four effect sizes or the last two would be included, and not both sets.

Sources of the Studies

The success of meta-analysis depended on collecting large numbers of research and evaluation studies from as many sources as available. The first task, therefore, was to identify the sources where studies had been done in the area of interest. Accordingly, a wide variety of sources were examined in the search for appropriate studies to be utilized in the meta-analysis. Information pertaining to potential studies to be included was obtained primarily from the
reference entitled *Dissertation Abstracts*. *Index to Education, Education Resources Information Center (ERIC) documents, mathematics education and psychological journals, research reviews, and government publications were also helpful in providing information about research studies which pertained to the area under investigation.

**Locating and Selecting Studies**

The sources helped advance progress toward the general goal of conducting a meta-analysis—a quantitative aggregation of a large number of research studies. After the primary sources of studies were identified, the next task in the process was to locate specific studies which pertained to new mathematics. This was accomplished through a systematic and thorough search of the key sources—*Index to Education, Dissertation Abstracts Index,* and *ERIC* for the years from 1951 to 1977. To systematize the process of locating and collecting studies, the researcher adopted certain criteria for study selection. The studies had to satisfy all criteria.

Broadly, the criteria for study selection were divided into primary and secondary. The primary set of criteria were the following: (a) the studies had to be experimental as opposed to descriptive or theoretical, (b) the dependent variables were to be limited to achievement or attitude, (c) tests of achievement or attitude were either standardized or experimenter developed or teacher developed, and (d) the area of experimentation had to be within the limits of the range of elements discussed in the definition of new mathematics in Chapter I.
The primary set of criteria was intended to select studies from abstracts and summaries of research studies in the areas of investigation; however, the criteria were not adequate to detect other characteristics which were necessary for the final selection of studies. Hence, a set of secondary criteria were established for the final selection process. The secondary criteria were the following: (a) studies comparing two or more groups had a control group which received traditional mathematics treatment, (b) studies adopted adequate method of analysis, and (c) studies related to the concepts introduced into mathematics instruction during the 1950's and 1960's.

Once the criteria for selection were established, the search for locating the studies started. The search was conducted in two stages. Initially, a variety of approaches was utilized to identify studies from different sources. For example, separate computer searches were made of the ERIC documents and Dissertation Abstracts by using appropriate descriptors. The searches yielded approximately 300 studies which were potentially usable.

In addition to dissertations and ERIC, a genuine attempt was made to locate and collect studies or results of studies from textbook publishers. Parenthetically, textbook publishers who played an active role in the production and promotion of modern mathematics materials frequently conducted research on the methods and techniques incorporated within their textbooks. The publishers were identified from the list of modern mathematics in EL-HI Textbooks in Print 1978. Unfortunately, the search elicited responses which were predominantly unusable.
The search for studies on modern mathematics programs from developmental projects was revealing on several dimensions. One, many of the developmental projects were no longer in existence. Among those which are in operation, ironically, no project indicated that it possessed any studies which measured the effectiveness of that particular program. A few of the research centers, such as the Ohio State University ERIC Center for Science-Mathematics and Environmental Education, proved to be useful in citing some of the relevant studies and locating their sources. Approximately 50 studies were located through this source.

At the end of the search through research centers, the investigator made a thorough search of the Dissertation Abstracts. In the Dissertation Abstracts more than 300 studies were identified which pertained to the area of investigation. Ninety-two dissertations were selected from the 300 studies for meta-analysis in accordance with the primary criteria. Copies of the 92 dissertations were obtained from the University Microfilms, Ann Arbor, Michigan.

As mentioned, a second search was completed for the additional studies. The search was based on the information gathered from the bibliographies and references of the previous set of studies. In the second search, 68 additional dissertation studies were identified and the copies of these dissertations were obtained from the University Microfilms, Ann Arbor, Michigan.

Many additional studies were located through a variety of sources. Some of them were identified in more than one source. Eliminating the repetitious studies, more than 200 studies were
originally identified through the brief descriptions given in the form of abstracts or summaries as potentially providing relevant data for the present analysis. The identification of these studies was accomplished by the primary set of criteria mentioned earlier.

While the investigator reviewed the original studies, it was necessary to eliminate those which met all the primary criteria, but did not satisfy some of the secondary criteria. The studies which were excluded fell into one of the following categories.

1. Studies comparing two or more groups which received new mathematics treatments, but did not have a control group which received traditional treatment. For example, a study which compared a guided discovery group with a hinted discovery group belonged to this category.

2. Studies which adopted an inadequate method of analysis. A study which showed the comparative effect only by the difference in percentage of pass or failure of the two groups typified this category.

3. Studies which belonged to the later developments in mathematics instruction were also excluded. This category included studies pertaining to the existential approach, self-supervision, individualized instruction, television-aided instruction, computer-assisted instruction, etc.

In the final screening, 134 studies were selected for meta-analysis from three different sources. The sources were the ERIC documents, journals, and doctoral dissertations. Of the three sources, doctoral dissertations contributed more than three-fourths
Identifying and Coding Variables

After the studies were selected through the processes of locating, collecting, and screening, information about the individual studies was extracted. Compilation of information about the studies was carried out through an identification and coding process for the relevant variables. Since meta-analysis research is concerned with the description of the relationships among findings and the characteristics of the studies, care was taken to identify as many characteristics as possible and classify them as specific variables.

Identification of the variables was a continuous process. Prior to reviewing each study in depth, the descriptive variables, such as number of students, grade level, type of treatment, length of treatment, year of study, rating, and the quality of design were identified. Later these variables were classified under four categories: the subject characteristics, treatment factors, study variables, and design variables. Identification of a variable consisted of three steps: (a) naming the variable, (b) defining the purpose the variable serves, and (c) assigning values to the variable.

Originally, 31 variables were identified for coding. While reviewing the studies, it was necessary to make modifications in the original list of variables. Major modifications were introduced in the range of values for certain variables. Two new variables were added to the original list: the generic content area and the sub-achievement area. Variables such as retention, time, transfer time,
the dependent variable, the length of treatment per day, the length of the treatment per minute, subachievement area, and new versus field-tested, were not considered for analysis, since the number of effect sizes in those cells was too small. The list of variables and the coding sheet are presented in Appendices A and B, respectively. The list of variables provided the name, purpose, and range of the values, while the code sheet consisted of the code, the name of the variable, the columns, and the values.

Coding the necessary information from the selected studies involved a great deal of concentration. For convenience, the collected studies were first of all arranged in a chronological order. Then, each study was reviewed in its entirety to identify and assign the proper values for the quantitative and descriptive characteristics which were listed in the code sheet. The descriptive data were spread throughout the study, while the quantitative data were mostly located in chapters dealing with methodology and results. Occasionally, tables of basic summary statistics, such as mean, standard deviation, and sample size, were provided in the appendix.

Each code sheet contained one effect size and all other variables related to the effect size. Consequently, the data in one code sheet formed the unit of analysis. From most of the studies more than one effect size was calculated. When there were many effect sizes in one study, the photostat copies of the pages in the original studies, which contained the necessary data for the calculation of effect sizes, were obtained. The copies were attached with the respective sets of code sheets which comprised the study. The procedure
facilitated the calculation of effect sizes and further reference.

Analysis of Data

After the information about the studies was coded in a systematic manner, the findings were analyzed. The analysis of the results of the present study was expected to fall under four major categories. The first category was the overall effect of new mathematics programs. This was divided into two subcategories: (a) the effect in terms of achievement and (b) the effect in terms of attitude. The overall effect size was measured by calculating the average effect size of all the studies taken together.

The second category of results consisted of effect sizes which were computed for each relevant independent variable. These effect sizes were calculated to show the relative significance of the different characteristics under consideration.

The calculations involved in the computation of effect sizes were completed with the use of a hand calculator. The analysis of data, such as the mean effect size, correlations, and regressions were done by computer. The computer analysis cards were punched from the code sheets. The data were entered into the computer from the cards. The computer analysis was performed by using Statistical Package and Bank Data Management Package, prepared and programmed by Houchard (1974) for the Western Michigan University Computer Center.
As previously discussed, the main tool for the analysis of the results was the mean effect size. Consequently, judgments on the results of the findings were heavily dependent upon the mean effect sizes. In order to make sound judgments about the results, it was necessary to have a clear perception of the magnitude and significance of the mean effect sizes.

Procedures to judge the magnitude and significance of effect size of a group of studies were not detailed in literature on meta-analysis. The most obvious and direct meaning of the effect size was the impact of a new treatment in terms of a shift from the mean point toward the positive or negative direction of the unit normal curve. The shift was measured in standard deviations which were readily convertible to percentiles. Any effect size which corresponded to a percentile greater than the 50th suggested a positive impact or an improvement by the new treatment over the untreated group; conversely, any effect size which corresponded to a percentile less than the 50th indicated a negative impact or a deterioration. The greater the average effect size was, the greater the impact was.

To determine whether the effect size of a group of studies was a chance factor or not, other procedures were necessary. In the present study, t-values and the exact probabilities were calculated corresponding to each average effect size of a group of studies. This was done in order to determine the significance of the effect size. The t-values were calculated under the assumptions that studies
on modern mathematics formed a sampling distribution and the hypo-
thetical mean of the traditional mathematics population was at the
zero point. Thus, the t-values were obtained from the formula:

\[ t = \frac{(\text{Effect size}) - (\text{Zero})}{(\text{Standard error of mean of the effect sizes})} \]

The degrees of freedom needed in addition to the t-value for the
calculation of the exact probability were obtained by subtracting one
from the number of effect sizes. Since the studies on modern mathe-
matics indicated positive as well as negative impacts, the probabil-
ities corresponding to the t-values were calculated for the two-
tailed test. The effect sizes were then considered to be significant
or not significant at a certain level of significance. The probabil-
ity of committing type I error for the present study was taken to be
0.05.

To help visualize more of the nature of the central tendency
and of the dispersion of the distribution of the effect sizes, the
values of the median and of the standard deviation were provided
along with the average effect size.

Summary

The discussion of the methodology for the present study focused
mainly in two areas: the theoretical aspects of meta-analysis and
the practical steps involved in its implementation. The theoretical
section consisted mainly of an explanation of the concept of meta-
analysis, the effect size, and the description of different techniques.
of measuring effect size. The practical steps involved in the implementa-
tion of meta-analysis consisted of identifying the source of the study, locating and collecting studies, identifying and coding the variables, and the analysis of the data.
CHAPTER IV

RESULTS

The purpose of Chapter IV is to present the results of the analysis of the data. The presentation of the results consists of a brief description of the studies included and a detailed discussion of the effects of the variables for which the effect sizes were calculated. The effects were classified into five groups: (a) the overall effects of modern mathematics in achievement and attitude, (b) the specific effects by subject characteristics, (c) the specific effects by treatment factors, (d) the specific effects by study variables, and (e) the specific effects by design variables.

Description of the Studies Included

A description of the studies was intended to give a picture of the general nature of the studies included in the present analysis. The factors that describe the general nature are the sources from which the studies were collected, the year in which the studies were published, the length of treatment of the studies, and the number of subjects per study.

Source of Study

One hundred and thirty-four studies were identified from three different sources. From the 134 studies, 810 effect sizes were calculated. The exact number of studies and effect sizes corresponding
to each source is shown in Table 5.

### Table 5

Studies Classified by Source

<table>
<thead>
<tr>
<th>Source</th>
<th>Number of Studies</th>
<th>Number of Effect Sizes</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERIC Documents</td>
<td>4</td>
<td>39</td>
</tr>
<tr>
<td>Journal Articles</td>
<td>16</td>
<td>127</td>
</tr>
<tr>
<td>Doctoral Dissertations</td>
<td>114</td>
<td>644</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>134</strong></td>
<td><strong>810</strong></td>
</tr>
</tbody>
</table>

### Year of Study

The studies utilized came from the period between 1951 and 1977. The majority of the studies were conducted during the second half of the 1960's. The median year for the studies and the effect sizes was 1968. In Table 6, the number of studies and the number of effect sizes corresponding to each year are listed.

### Length of Treatment and Number of Subjects Per Study

The length of treatment of studies ranged from 1 week to 144 weeks. The average length of treatment was approximately 33 weeks. The number of subjects per study ranged from 4 to 2,862. The average number of subjects per study was 108. In order to obtain a clearer perspective of the central tendencies along with the mean of the two
Table 6
Studies Classified by Year

<table>
<thead>
<tr>
<th>Year of Publications</th>
<th>Number of Studies</th>
<th>Number of Effect Sizes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1955</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>1956</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>1958</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1959</td>
<td>6</td>
<td>20</td>
</tr>
<tr>
<td>1960</td>
<td>2</td>
<td>19</td>
</tr>
<tr>
<td>1961</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>1962</td>
<td>4</td>
<td>20</td>
</tr>
<tr>
<td>1963</td>
<td>5</td>
<td>59</td>
</tr>
<tr>
<td>1964</td>
<td>5</td>
<td>19</td>
</tr>
<tr>
<td>1965</td>
<td>10</td>
<td>47</td>
</tr>
<tr>
<td>1966</td>
<td>9</td>
<td>96</td>
</tr>
<tr>
<td>1967</td>
<td>13</td>
<td>73</td>
</tr>
<tr>
<td>1968</td>
<td>10</td>
<td>83</td>
</tr>
<tr>
<td>1969</td>
<td>13</td>
<td>70</td>
</tr>
<tr>
<td>1970</td>
<td>14</td>
<td>107</td>
</tr>
<tr>
<td>1971</td>
<td>8</td>
<td>47</td>
</tr>
<tr>
<td>1972</td>
<td>5</td>
<td>30</td>
</tr>
<tr>
<td>1973</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>1974</td>
<td>3</td>
<td>25</td>
</tr>
<tr>
<td>1975</td>
<td>6</td>
<td>26</td>
</tr>
<tr>
<td>1976</td>
<td>7</td>
<td>20</td>
</tr>
<tr>
<td>1977</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>Median Year:</td>
<td>1968</td>
<td>1968</td>
</tr>
</tbody>
</table>
variables, the measures of the median and mode are provided in Table 7. The maximum and minimum values of the variables are also provided in the same table. This was done in order to indicate the direction of the range of observations under consideration.

Table 7
Length of Treatment and Number of Subjects Per Study

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Median</th>
<th>Mode</th>
<th>Maximum</th>
<th>Minimum</th>
<th>Number of Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length of treatment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(in weeks)</td>
<td>33.84</td>
<td>24</td>
<td>36</td>
<td>144</td>
<td>1</td>
<td>810</td>
</tr>
<tr>
<td>Number of subjects per study</td>
<td>206.90</td>
<td>108</td>
<td>24</td>
<td>2,862</td>
<td>4</td>
<td>810</td>
</tr>
</tbody>
</table>

Overall Effects of Modern Mathematics

Description of the studies provided a picture of the general nature of the studies included. The ensuing sections will deal with the quantified measures of the effects of modern mathematics in terms of achievement and/or attitude.

Chapter III discussed the concept that the "effect size" was a common measure of treatment effectiveness. This measure was defined as the difference between the means of the modern mathematics group and the traditional mathematics group divided by the standard deviation of the traditional mathematics group. Based on the preceding definition, 660 effect sizes in achievement and 150 effect sizes in
attitude were calculated from 134 controlled outcome studies. It must be remembered that in some studies more than one effect size was calculated.

The grand mean effect size in achievement was 0.24 and the grand mean effect size in attitude was 0.12. Table 8 shows the average effect sizes in achievement and attitude.

Table 8
Grand Mean Effect Sizes in Achievement and Attitude

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Achievement</th>
<th>Attitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.24*</td>
<td>0.12*</td>
</tr>
<tr>
<td>Median</td>
<td>0.09</td>
<td>0.04</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.77</td>
<td>0.56</td>
</tr>
<tr>
<td>Standard error of mean</td>
<td>0.03</td>
<td>0.05</td>
</tr>
<tr>
<td>Number of effect sizes</td>
<td>660</td>
<td>150</td>
</tr>
<tr>
<td>t-value</td>
<td>8.00</td>
<td>2.40</td>
</tr>
<tr>
<td>Probability</td>
<td>0.00</td>
<td>0.018</td>
</tr>
</tbody>
</table>

Note. Degrees of freedom = Number of effect sizes - 1.

*The null hypothesis, stating the mean effect size is zero, is rejected with the probability of committing type I error being less than or equal to 0.05.

This meant that the average person receiving some form of modern mathematics treatment was 0.24 standard deviation more improved on the achievement measure and 0.12 standard deviation more improved on the attitude measure than the average traditional mathematics group.
member (see Figure 4).

The interpretation of Figure 4 followed the reasonings presented while discussing Figure 2 in Chapter III. The dotted lines represented the normal distribution of the traditional mathematics population and the solid line represented the normal distribution of the modern mathematics population. Since the effect size in achievement was 0.24, a student at the 50th percentile of the modern mathematics group was equivalent to a student at the 59th percentile of the traditional mathematics group. Similarly, a student at the 50th percentile of the modern mathematics group was equivalent to a student at the 55th percentile of the traditional mathematics group in attitude. The effect size of attitude was 0.12. In other words, the modern mathematics program raised an average student from the 50th to the 59th percentile in achievement and from the 50th to the 55th percentile in attitude over the traditional mathematics population. The basic assumption was that an average student of the traditional mathematics population occupied the 50th percentile in a normal curve.

It is interesting to note that the improvements, reported in terms of effect sizes which resulted from the treatment of modern mathematics, indicated a great difference in scores on the achievement and attitude dimensions. In the following sections, treatment of the specific effects are provided.

Specific Effects by Subject Characteristics

The subject variables described the important characteristics of the students who were involved in the studies that were utilized.
Figure 4. Normal curves illustrating the aggregate effect of modern mathematics in achievement and attitude in relation to traditional mathematics group.
Four main subject characteristics were analyzed. They were the grades which were chosen for the experiments, ability levels, the levels according to socio-economic status, and sex. Classification of the subjects on the basis of the variables discussed was intended to enable the reader to form an idea of the effectiveness of modern mathematics under varying conditions and circumstances. Since the classification of levels within ability and socio-economic status was arbitrary in nature, the effect sizes in those subcells were less specific and reliable than the effect sizes of grade levels and sex.

**Effect Size by Grade**

Grade levels considered for the present study ranged from kindergarten to post-secondary. The post-secondary group consisted mostly of students of the community colleges and their freshmen and sophomore counterparts at four year colleges and universities. Among the five grade levels, the elementary had the largest average effect size in achievement. It was 0.32. An average effect size which was close to the elementary grades was from the senior high grades. Conversely, the average effect size at the junior high grades was remarkably low in comparison to the elementary and senior high grade levels. Table 9 indicates that the average effect size of the kindergarten and the post-secondary grade levels in achievement was insignificant. This was possibly due to the limited number of effect sizes in each category.

An attempt was made, in observing the average effect sizes, to find the relationship between the significant effect sizes and grade
Table 9

Effect Size by Levels of Grade and Ability in Achievement

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
<th>Standard Deviation</th>
<th>Standard Error of Mean</th>
<th>Number of Effect Sizes</th>
<th>t-value</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Levels of Grade</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kindergarten</td>
<td>0.06</td>
<td>0.00</td>
<td>0.47</td>
<td>0.14</td>
<td>11</td>
<td>0.43</td>
<td>0.676</td>
</tr>
<tr>
<td>Elementary</td>
<td>0.32</td>
<td>0.14</td>
<td>0.96</td>
<td>0.07</td>
<td>216</td>
<td>4.57</td>
<td>0.000*</td>
</tr>
<tr>
<td>Junior high</td>
<td>0.17</td>
<td>0.02</td>
<td>0.58</td>
<td>0.04</td>
<td>249</td>
<td>4.25</td>
<td>0.000*</td>
</tr>
<tr>
<td>Senior high</td>
<td>0.28</td>
<td>0.18</td>
<td>0.77</td>
<td>0.06</td>
<td>140</td>
<td>4.67</td>
<td>0.000*</td>
</tr>
<tr>
<td>Post-secondary</td>
<td>0.12</td>
<td>0.09</td>
<td>0.47</td>
<td>0.07</td>
<td>44</td>
<td>1.71</td>
<td>0.095</td>
</tr>
<tr>
<td><strong>Levels of Ability</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>0.35</td>
<td>0.00</td>
<td>1.25</td>
<td>0.13</td>
<td>89</td>
<td>2.69</td>
<td>0.009*</td>
</tr>
<tr>
<td>Middle</td>
<td>0.25</td>
<td>0.07</td>
<td>0.64</td>
<td>0.07</td>
<td>79</td>
<td>3.57</td>
<td>0.001*</td>
</tr>
<tr>
<td>High</td>
<td>0.21</td>
<td>0.00</td>
<td>0.79</td>
<td>0.08</td>
<td>88</td>
<td>2.63</td>
<td>0.010*</td>
</tr>
<tr>
<td>Unspecified</td>
<td>0.22</td>
<td>0.14</td>
<td>0.63</td>
<td>0.03</td>
<td>404</td>
<td>7.33</td>
<td>0.000*</td>
</tr>
</tbody>
</table>

Note. Degrees of freedom = Number of effect sizes - 1.

*The null hypothesis, stating the mean effect size is zero, is rejected with the probability of committing type I error being less than or equal to 0.05.
levels. This was done by fitting a parabola with a downward bend. The multiple correlation coefficient of the curvilinear regression was only 0.09. The value of the multiple correlation coefficient indicates that there is no significant relationship between the levels of grades and the effect sizes in achievement.

Effect Size by Ability Level

Most of the studies which compared the effectiveness of modern mathematics based on ability divided the subjects into three groups, such as low, middle, and high. In a few instances, the comparison was made between the high and low ability levels only. In this study, all three levels of ability were considered. Effect sizes were classified into low, middle, or high levels of ability according to the designation in the original studies. A fourth group, designated as "unspecified" ability level, consisted of studies which did not make a comparative analysis by ability levels.

The average effect sizes were considered to be statistically significant at $\alpha = 0.05$. Table 9 indicates that the average effect size in achievement was the largest for the low ability group, which was 0.35. This value was substantially higher than that of the grand mean effect size in achievement which was 0.24 (see Table 8). The average effect sizes of the middle, high, and the unspecified ability levels did not deviate from the grand mean effect size in any significant manner. In general, Table 9 shows a decrease in mean effect size as one goes from the low to the high level of ability.
Effect Size by Socio-economic Status and by Sex

Socio-economic status and sex were the other two variables which belonged to the category of subject characteristics. There were only 72 effect sizes out of 660 in achievement where the socio-economic status of the students was identified (see Table 10). The 72 effect sizes were not representative of the total population of studies insofar as the mean of the 72 effect sizes was not equal to the grand mean effect size. The only effect size which was significant at the $\alpha = 0.05$ level was that of the students belonging to middle class parents. The effect size of the middle group was 1.05 which was conspicuously higher than the grand mean effect size of the present study.

As in the case of the analysis of socio-economic status, the cell sizes in achievement were quite small when the data were analyzed by sex. As Table 10 indicated, the effect sizes of each category in this particular subset did not represent the total population, since the average mean for either sex was substantially lower than those studies not grouped for sex.

Because of the small number of studies in the socio-economic status and sex categories, caution must be exercised in the interpretation of the results.

Specific Effects by Treatment Factors

The previous section dealt with the characteristics of the students who were involved in the studies that were utilized; whereas,
Table 10
Effect Size by Socio-economic Status and Sex in Achievement

<table>
<thead>
<tr>
<th>Levels of Socio-economic Status</th>
<th>Mean</th>
<th>Median</th>
<th>Standard Deviation</th>
<th>Standard Error of Mean</th>
<th>Number of Effect Sizes</th>
<th>t-value</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>0.15</td>
<td>0.01</td>
<td>0.47</td>
<td>0.09</td>
<td>25</td>
<td>1.67</td>
<td>0.108</td>
</tr>
<tr>
<td>Middle</td>
<td>1.05</td>
<td>0.42</td>
<td>1.77</td>
<td>0.31</td>
<td>33</td>
<td>3.39</td>
<td>0.002*</td>
</tr>
<tr>
<td>High</td>
<td>0.22</td>
<td>0.35</td>
<td>0.50</td>
<td>0.13</td>
<td>14</td>
<td>1.69</td>
<td>0.115</td>
</tr>
<tr>
<td>Unspecified</td>
<td>0.21</td>
<td>0.78</td>
<td>0.66</td>
<td>0.03</td>
<td>588</td>
<td>7.00</td>
<td>0.000*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Difference in Sex</th>
<th>Mean</th>
<th>Median</th>
<th>Standard Deviation</th>
<th>Standard Error of Mean</th>
<th>Number of Effect Sizes</th>
<th>t-value</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>0.11</td>
<td>-0.05</td>
<td>0.59</td>
<td>0.12</td>
<td>26</td>
<td>0.92</td>
<td>0.366</td>
</tr>
<tr>
<td>Female</td>
<td>0.03</td>
<td>-0.05</td>
<td>0.66</td>
<td>0.12</td>
<td>28</td>
<td>0.25</td>
<td>0.805</td>
</tr>
</tbody>
</table>

Note. Degrees of freedom = Number of effect sizes - 1.

*The null hypothesis, stating the mean effect size is zero, is rejected with the probability of committing type I error being less than or equal to 0.05.
the treatment factors described the type and content of application of modern mathematics in the various experiments selected for the study. Two factors which related to the content of application were the general content area and the specific subject area. There were three variables which described the type of treatment. They were the type of program that had been selected for the treatment of modern mathematics, the length of treatment, and the number of subjects per study.

**Effect Size by General Treatment Area**

Table 11 reveals that 304 effect sizes in achievement were identified as belonging to one of the three selected categories. The categories were in accordance with the emphasis given in the treatment of modern mathematics in the generic area. The three areas of emphasis were concepts, computation, and application. Studies which could not be identified as belonging to one of these three categories were treated in the fourth category, designated as the "combination."

Among the four categories, the highest average effect size was that of the concepts. Almost similar achievement gain was evidenced in computation. While the mean effect sizes of the concepts and of computation were 0.36 and 0.31, respectively, the mean effect size of the application was only 0.06. In other words, the improvement attributable to modern mathematics in the area of application of the mathematical concepts was next to zero, whereas the improvements in the areas of concepts and of content was quite substantial.
### Table 11

Effect Size by Content Area in Achievement

<table>
<thead>
<tr>
<th>General Content Area</th>
<th>Mean</th>
<th>Median</th>
<th>Standard Deviation</th>
<th>Standard Error of Mean</th>
<th>Number of Effect Sizes</th>
<th>t-value</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concepts</td>
<td>0.36</td>
<td>0.27</td>
<td>0.78</td>
<td>0.07</td>
<td>134</td>
<td>5.14</td>
<td>0.000*</td>
</tr>
<tr>
<td>Computation</td>
<td>0.31</td>
<td>0.08</td>
<td>1.05</td>
<td>0.11</td>
<td>95</td>
<td>2.82</td>
<td>0.006*</td>
</tr>
<tr>
<td>Application</td>
<td>0.06</td>
<td>0.00</td>
<td>0.45</td>
<td>0.05</td>
<td>75</td>
<td>1.20</td>
<td>0.234</td>
</tr>
<tr>
<td>Combination</td>
<td>0.22</td>
<td>0.09</td>
<td>0.71</td>
<td>0.04</td>
<td>356</td>
<td>5.50</td>
<td>0.000*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Specific Subject Area</th>
<th>Mean</th>
<th>Median</th>
<th>Standard Deviation</th>
<th>Standard Error of Mean</th>
<th>Number of Effect Sizes</th>
<th>t-value</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arithmetic</td>
<td>0.21</td>
<td>0.05</td>
<td>0.88</td>
<td>0.05</td>
<td>305</td>
<td>4.20</td>
<td>0.000*</td>
</tr>
<tr>
<td>Algebra</td>
<td>0.43</td>
<td>0.31</td>
<td>0.68</td>
<td>0.08</td>
<td>74</td>
<td>5.38</td>
<td>0.000*</td>
</tr>
<tr>
<td>Geometry</td>
<td>0.14</td>
<td>-0.01</td>
<td>0.70</td>
<td>0.08</td>
<td>85</td>
<td>1.75</td>
<td>0.084</td>
</tr>
<tr>
<td>Other</td>
<td>0.25</td>
<td>0.11</td>
<td>0.62</td>
<td>0.04</td>
<td>196</td>
<td>6.25</td>
<td>0.000*</td>
</tr>
</tbody>
</table>

Note. Degrees of freedom = Number of effect sizes - 1.

*The null hypothesis, stating the mean effect size is zero, is rejected with the probability of committing type I error being less than or equal to 0.05.
Effect Size by Specific Subject Area

An attempt was made to classify effect sizes into specific subject areas, such as arithmetic, algebra, and geometry. The fourth section consisted of treatments pertaining to sets, groups, fields, logic, and some specific areas of mathematics, such as calculus, statistics, and probability, which were introduced in the language of modern mathematics.

Among the specific subject areas, algebra had the highest average effect size with 0.43 (see Table 11) in achievement. The average effect size in arithmetic was almost close to the grand mean effect size of the total number of studies in achievement (0.21 versus 0.24). Table 11 also points out the average effect size of achievement in geometry (0.14) was lower than that of arithmetic (0.21) and of algebra (0.43). The mean effect size in achievement of the combination group was typical of the grand mean effect size in achievement as reported in Table 8. Generally, the mean effect sizes of arithmetic and of the combination group represented the grand mean effect size. The larger mean effect size of algebra was offset by the smaller mean effect size of geometry.

Effect Size by Type of Program

Initially, an attempt was made to identify studies according to the specific development programs discussed in Chapter I; however, the number of studies belonging to each specific development program other than the School Mathematics Study Group (SMSG) was either very small.
or nil. The type of program therefore, was broadly divided into SMSG programs and non-SMSG programs. The average effect size in achievement of studies belonging to SMSG was considerably lower than the average effect size in achievement of all other studies put together (see Table 12). Other studies denoted as non-SMSG, included a few studies on modern mathematics as introduced by other developmental programs such as the University of Illinois Committee on School Mathematics Program, the Ball State University Program, the Greater Cleveland Mathematics Program, and a number of studies on modern mathematics as introduced by innovators other than the reformers in the developmental programs.

**Effect Size by Length of Treatment**

Table 13 reports the average effect size in achievement according to the length of treatment. Studies with a length of treatment which was less than 9 weeks and studies with a length of treatment which was greater than or equal to 36 weeks had the same average effect size (0.25). The average effect size reached the maximum point, which was 0.48 in this context, when the length of treatment was between 9 and 18 weeks. Studies with the length of treatment ranging from 18 to 36 weeks had average effect sizes as low as 0.09. Thus, the effect sizes indicated a gradual increase until the length of treatment became 18 weeks or more. After 18 weeks, the effect sizes gradually decreased until the length of treatment was 36 weeks or greater. Thereafter, the effect sizes in achievement rose and remained steady at 0.25, which was close to the grand mean effect.
Table 12

Effect Size by Type of Program in Achievement

<table>
<thead>
<tr>
<th>Program</th>
<th>Mean</th>
<th>Median</th>
<th>Standard Deviation</th>
<th>Standard Error of Mean</th>
<th>Number of Effect Sizes</th>
<th>t-value</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMSG</td>
<td>0.14</td>
<td>0.00</td>
<td>0.52</td>
<td>0.05</td>
<td>127</td>
<td>2.80</td>
<td>0.006*</td>
</tr>
<tr>
<td>Other</td>
<td>0.27</td>
<td>0.10</td>
<td>0.81</td>
<td>0.04</td>
<td>533</td>
<td>6.75</td>
<td>0.000*</td>
</tr>
</tbody>
</table>

Note. Degrees of freedom = Number of effect sizes – 1.

*The null hypothesis, stating the mean effect size is zero, is rejected with the probability of committing type I error being less than or equal to 0.05.
### Table 13

Effect Size by Length of Treatment in Achievement

<table>
<thead>
<tr>
<th>Length of Treatment (weeks)</th>
<th>Mean</th>
<th>Median</th>
<th>Standard Deviation</th>
<th>Standard Error of Mean</th>
<th>Number of Effect Sizes</th>
<th>t-value</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than 9</td>
<td>0.25</td>
<td>0.17</td>
<td>0.57</td>
<td>0.05</td>
<td>133</td>
<td>5.00</td>
<td>0.000*</td>
</tr>
<tr>
<td>Between 9 and 18</td>
<td>0.48</td>
<td>0.00</td>
<td>1.43</td>
<td>0.17</td>
<td>69</td>
<td>2.82</td>
<td>0.006*</td>
</tr>
<tr>
<td>Between 18 and 36</td>
<td>0.09</td>
<td>0.04</td>
<td>0.64</td>
<td>0.06</td>
<td>124</td>
<td>1.50</td>
<td>0.136</td>
</tr>
<tr>
<td>36 or more</td>
<td>0.25</td>
<td>0.09</td>
<td>0.66</td>
<td>0.04</td>
<td>334</td>
<td>6.25</td>
<td>0.000*</td>
</tr>
</tbody>
</table>

Note. Degrees of freedom = Number of effect sizes - 1.

*The null hypothesis, stating the mean effect size is zero, is rejected with the probability of committing type I error being less than or equal to 0.05.*
size in achievement.

**Effect Size by Number of Subjects Per Study**

As indicated in Table 14, the number of subjects per study was categorized into five groups. The upper limits of each class interval of the distribution of number of subjects per study were in geometric progression with common ratio equal to 2 and the base equal to 50. With this arrangement, the effect sizes were relatively evenly distributed in each subgroup. The average effect size in achievement was fairly high for the studies in which the number of subjects was less than 50. Studies in which the number of subjects was greater than 200 and less than 400 had a typical average effect size (0.26). It was only 0.02 standard deviation greater than the grand mean effect size. The average effect sizes in the other cells of this subgroup ranged from 0.13 to 0.16.

**Specific Effect Size by Study Variables**

The study variables outlined characteristics such as the emphasis in the approach of the study, the source of the studies, the affiliation of the researcher, and the author who introduced an innovation. The variables in this category were not manipulative in nature, but they were related to the outcomes. Collectively, they described those characteristics of the study which were not treated under the subject characteristics, treatment factors, or design variables.
Table 14

Effect Size by Number of Subjects Per Study in Achievement

<table>
<thead>
<tr>
<th>Subjects Per Study</th>
<th>Mean</th>
<th>Median</th>
<th>Standard Deviation</th>
<th>Standard Error of Mean</th>
<th>Number of Effect Sizes</th>
<th>t-value</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 or less</td>
<td>0.49</td>
<td>0.25</td>
<td>1.10</td>
<td>0.09</td>
<td>143</td>
<td>5.44</td>
<td>0.000*</td>
</tr>
<tr>
<td>Between 50 and 100</td>
<td>0.16</td>
<td>0.05</td>
<td>0.78</td>
<td>0.06</td>
<td>153</td>
<td>2.67</td>
<td>0.008*</td>
</tr>
<tr>
<td>Between 100 and 200</td>
<td>0.15</td>
<td>0.07</td>
<td>0.53</td>
<td>0.04</td>
<td>178</td>
<td>3.75</td>
<td>0.000*</td>
</tr>
<tr>
<td>Between 200 and 400</td>
<td>0.26</td>
<td>0.10</td>
<td>0.67</td>
<td>0.06</td>
<td>118</td>
<td>4.33</td>
<td>0.000*</td>
</tr>
<tr>
<td>More than 400</td>
<td>0.13</td>
<td>0.08</td>
<td>0.34</td>
<td>0.04</td>
<td>68</td>
<td>3.25</td>
<td>0.002*</td>
</tr>
</tbody>
</table>

Note. Degrees of freedom = Number of effect sizes - 1.

*The null hypothesis, stating the mean effect size is zero, is rejected with the probability of committing type I error being less than or equal to 0.05.
Effect Size by Study Approach

The variable of study approach allowed the studies to be distinguished according to one of four areas of emphasis: methodology approach, content approach, materials approach, and combination approach. The combination approach consisted of studies with no special emphasis on any of the three areas previously discussed. The majority of the studies which emphasized the methodology belonged to the discovery approach, heuristic method, and deductive reasoning. Studies which emphasized the content dealt mostly with topics such as sets, logic, structure, and nondecimal bases. Most of the studies which emphasized the materials belonged to experiments with manipulative materials, the Cuisenaire approach, and the mathematics laboratory.

Table 15 shows the average effect sizes of each of the study approaches in achievement and attitude. Obviously, there were large achievement gains in all the approaches except that of methodology. Conversely, there were no substantial attitudinal gains in any of the four approaches except that of methodology. In the methodology approach the achievement gain was only 0.04, which was the least of all the mean effect sizes. Conversely, the attitudinal gain in the methodology approach was 0.17, which was the largest mean effect size. The best result in achievement was that of the materials approach with a mean effect size of 0.51. The content approach was second with an average effect size of 0.35.
Table 15

Effect Size by Study Approach in Achievement and Attitude

<table>
<thead>
<tr>
<th>Approach</th>
<th>Mean</th>
<th>Median</th>
<th>Standard Deviation</th>
<th>Standard Error of Mean</th>
<th>Number of Effect Sizes</th>
<th>t-value</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Achievement</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methodology</td>
<td>0.04</td>
<td>0.08</td>
<td>0.59</td>
<td>0.04</td>
<td>181</td>
<td>1.00</td>
<td>0.319</td>
</tr>
<tr>
<td>Content</td>
<td>0.35</td>
<td>0.10</td>
<td>0.94</td>
<td>0.06</td>
<td>229</td>
<td>5.83</td>
<td>0.000*</td>
</tr>
<tr>
<td>Material</td>
<td>0.51</td>
<td>0.21</td>
<td>0.86</td>
<td>0.14</td>
<td>39</td>
<td>3.64</td>
<td>0.001*</td>
</tr>
<tr>
<td>Combination</td>
<td>0.25</td>
<td>0.05</td>
<td>0.61</td>
<td>0.04</td>
<td>211</td>
<td>6.25</td>
<td>0.000*</td>
</tr>
<tr>
<td><strong>Attitude</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methodology</td>
<td>0.17</td>
<td>0.21</td>
<td>0.70</td>
<td>0.07</td>
<td>84</td>
<td>2.43</td>
<td>0.017*</td>
</tr>
<tr>
<td>Content</td>
<td>-0.07</td>
<td>0.00</td>
<td>0.21</td>
<td>0.05</td>
<td>19</td>
<td>-1.04</td>
<td>0.312</td>
</tr>
<tr>
<td>Material</td>
<td>0.12</td>
<td>0.16</td>
<td>0.36</td>
<td>0.09</td>
<td>17</td>
<td>1.33</td>
<td>0.202</td>
</tr>
<tr>
<td>Combination</td>
<td>0.08</td>
<td>0.00</td>
<td>0.26</td>
<td>0.05</td>
<td>30</td>
<td>1.60</td>
<td>0.120</td>
</tr>
</tbody>
</table>

**Note.** Degrees of freedom = Number of effect sizes - 1.

*The null hypothesis, stating the mean effect size is zero, is rejected with the probability of committing type I error being less than or equal to 0.05.
Effect Size by Source of the Study

Three different sources from which the studies were gathered were the ERIC documents, journal articles, and doctoral dissertations. Table 16 lists the effect sizes by each source. The ERIC documents consisted of school district reports, final reports from federal grants, and other miscellany. The studies from the ERIC documents had the largest average effect size in achievement—0.77—an effect size which was three times bigger than that of the grand mean effect size. The journal articles had an effect size (0.42) which was about twice the grand mean effect size in achievement. The mean effect size of the doctoral dissertations was the least in the category of source (0.16).

Effect Size by Other Descriptive Variables

Other descriptive variables analyzed for the present study were the subject area of investigation, the affiliation of the researcher, and the author. The subject area variable had two categories: (a) mathematics alone and (b) mathematics and other subjects combined. In Table 17, the average effect sizes by subject area are reported. The average effect sizes in both categories were almost identical.

"Affiliation of researcher" identified the organization the researcher had an association with while conducting the study. The variable was broadly divided into affiliation with a university and affiliation with any organization other than a university. Table 18 indicates that the average effect sizes in achievement were almost
Table 16

Effect Size by Source in Achievement

<table>
<thead>
<tr>
<th>Source</th>
<th>Mean</th>
<th>Median</th>
<th>Standard Deviation</th>
<th>Standard Error of Mean</th>
<th>Number of Effect Sizes</th>
<th>t-value</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERIC documents</td>
<td>0.77</td>
<td>0.02</td>
<td>1.70</td>
<td>0.27</td>
<td>39</td>
<td>2.85</td>
<td>0.000*</td>
</tr>
<tr>
<td>Journal articles</td>
<td>0.42</td>
<td>0.34</td>
<td>0.52</td>
<td>0.05</td>
<td>121</td>
<td>8.40</td>
<td>0.000*</td>
</tr>
<tr>
<td>Doctoral dissertations</td>
<td>0.16</td>
<td>0.04</td>
<td>0.67</td>
<td>0.03</td>
<td>500</td>
<td>5.33</td>
<td>0.000*</td>
</tr>
</tbody>
</table>

Note. Degrees of freedom = Number of effect sizes - 1.

*The null hypothesis, stating the mean effect size is zero, is rejected with the probability of committing type I error being less than or equal to 0.05.
<table>
<thead>
<tr>
<th>Area of Investigation</th>
<th>Mean</th>
<th>Median</th>
<th>Standard Deviation</th>
<th>Standard Error of Mean</th>
<th>Number of Effect Sizes</th>
<th>t-value</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mathematics</td>
<td>0.24</td>
<td>0.09</td>
<td>0.78</td>
<td>0.03</td>
<td>609</td>
<td>8.00</td>
<td>0.000*</td>
</tr>
<tr>
<td>Other</td>
<td>0.23</td>
<td>0.09</td>
<td>0.61</td>
<td>0.08</td>
<td>51</td>
<td>2.88</td>
<td>0.006*</td>
</tr>
</tbody>
</table>

*The null hypothesis, stating the mean effect size is zero, is rejected with the probability of committing type I error being less than or equal to 0.05.*
Table 18

Effect Size by Affiliation of the Researcher in Achievement

<table>
<thead>
<tr>
<th>Affiliation</th>
<th>Mean</th>
<th>Median</th>
<th>Standard Deviation</th>
<th>Standard Error of Mean</th>
<th>Number of Effect Sizes</th>
<th>t-value</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>University</td>
<td>0.24</td>
<td>0.09</td>
<td>0.79</td>
<td>0.03</td>
<td>583</td>
<td>8.00</td>
<td>0.000*</td>
</tr>
<tr>
<td>Other</td>
<td>0.25</td>
<td>0.09</td>
<td>0.49</td>
<td>0.06</td>
<td>77</td>
<td>4.17</td>
<td></td>
</tr>
</tbody>
</table>

Note. Degrees of freedom = Number of effect sizes - 1.

*The null hypothesis, stating the mean effect size is zero, is rejected with the probability of committing type I error being less than or equal to 0.05.
identical for both the studies affiliated with a university and those affiliated with other organizations. The similarity was also true for the subject area of investigation.

An attempt was made to identify the author who introduced either a new content, a new methodology, a new material, or any combination of the three. The purpose of this variable was to identify the direction of bias in the results due to the influence of personal interest in the topic of experimentation. For this purpose, the variable (author) was subdivided into four categories (commercial, school districts, experimenter, and university); however, it was impossible in most of the cases to identify the real author of the innovations.

An attempt was made to determine the extent of each researcher's personal involvement in the experiment, but it was uncertain in most of the cases exactly how involved the researcher was in the results of the study. In some cases the researcher was either the director of a project or a member of the staff of a project; however, information of this nature was available in only a few cases. Consequently, a determination could not be made about the extent of the researcher's actual involvement in the study.

Specific Effects by Design Variables

The fourth category of the specific variables dealt with the design variables. The function of the variables in this category was to measure the different characteristics of the research design used in each study. Three variables were selected for this purpose. They were the type of test used to measure the achievement and/or attitude,
the quality of the design used for each experiment, and the novelty effect which was assumed to be a factor that depends on the length of treatment.

**Effect Size by Type of Test**

Basically, there were only two categories of tests that were used in the studies. One category consisted of a variety of standardized norm-referenced mathematics tests. The other category consisted of teacher or experimenter developed tests. Table 19 shows that about two-thirds of the effect sizes in achievement were measured by commercial tests. The average effect size in achievement of the commercial test was only 0.15 which was lower than the grand mean effect size in achievement; however, the average effect size of the experimenter developed test in achievement was about three times greater than that of the commercial tests.

**Effect Size by Quality of Design**

The quality of design variable had two categories: studies where subjects were randomly assigned to treatment and studies where subjects were not randomly assigned to treatment. As indicated in Table 19, 552 effect sizes in achievement were obtained from studies with random assignments. The average effect size in achievement of the studies with random assignment was 0.26, whereas the average effect size of studies without random assignment was only slightly more than half (0.14) of the former.
Table 19
Effect Size by Design Variables in Achievement

<table>
<thead>
<tr>
<th>Types of Test</th>
<th>Mean</th>
<th>Median</th>
<th>Standard Deviation</th>
<th>Standard Error of Mean</th>
<th>Number of Effect Sizes</th>
<th>t-value</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commercial</td>
<td>0.15</td>
<td>0.05</td>
<td>0.62</td>
<td>0.03</td>
<td>438</td>
<td>5.00</td>
<td>0.000*</td>
</tr>
<tr>
<td>Experimenter</td>
<td>0.42</td>
<td>0.25</td>
<td>0.97</td>
<td>0.07</td>
<td>222</td>
<td>6.00</td>
<td>0.000*</td>
</tr>
<tr>
<td>developed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality of Design</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random assignment</td>
<td>0.26</td>
<td>0.09</td>
<td>0.80</td>
<td>0.03</td>
<td>552</td>
<td>8.67</td>
<td>0.000*</td>
</tr>
<tr>
<td>No random assignment</td>
<td>0.14</td>
<td>0.05</td>
<td>0.53</td>
<td>0.05</td>
<td>108</td>
<td>2.80</td>
<td>0.006*</td>
</tr>
<tr>
<td>Novelty Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Present</td>
<td>0.33</td>
<td>0.13</td>
<td>0.96</td>
<td>0.07</td>
<td>202</td>
<td>4.71</td>
<td>0.000*</td>
</tr>
<tr>
<td>Absent</td>
<td>0.20</td>
<td>0.08</td>
<td>0.66</td>
<td>0.03</td>
<td>458</td>
<td>6.67</td>
<td>0.000*</td>
</tr>
</tbody>
</table>

Note. Degrees of freedom = Number of effect sizes - 1.

*The null hypothesis, stating the mean effect size is zero, is rejected with the probability of committing type I error being less than or equal to 0.05.
Effect Size by Novelty Effect

Since the novelty effect variable depended on the length of treatment, a critical measure was arbitrarily established to measure the novelty effect. For the purposes of the present study, the novelty effect was assumed to be present whenever the length of treatment was less than 18 weeks. For those studies in which novelty effect was judged to be present, the average was significantly higher than that of the studies not having novelty effect. Table 19 lists the effect sizes in achievement by novelty effect. About one-third of the number of effect sizes had a novelty effect. The average effect size in achievement of the studies without novelty effect was less that two-thirds of the effect size in achievement of the studies with novelty effect.

Summary

Table 20 gives an overview of the mean effect sizes in achievement and/or attitude of all the variables analyzed. In this section only the major results are highlighted.

When the subjects were divided into five grade levels, the elementary and the senior high had the larger effect sizes in achievement, which were respectively 0.32 and 0.28. Among the low, middle, and high ability groups, the low ability had the largest effect size (0.35) in achievement. In a similar grouping based on the socio-economic status, the effect size of the middle group in achievement was 1.05 which was conspicuously higher than that of the low and high
Table 20

Effect Sizes by All Categories of Variables in Achievement and/or Attitude

<table>
<thead>
<tr>
<th>Variable</th>
<th>Effect Size</th>
<th>Variable</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Achievement by grade</td>
<td></td>
<td>Achievement by number of subjects</td>
<td></td>
</tr>
<tr>
<td>Kindergarten</td>
<td>0.06</td>
<td>50 or less</td>
<td>0.49*</td>
</tr>
<tr>
<td>Elementary</td>
<td>0.32*</td>
<td>Between 50 and 100</td>
<td>0.16*</td>
</tr>
<tr>
<td>Junior high</td>
<td>0.17*</td>
<td>Between 100 and 200</td>
<td>0.15*</td>
</tr>
<tr>
<td>Senior high</td>
<td>0.28*</td>
<td>Between 200 and 400</td>
<td>0.26*</td>
</tr>
<tr>
<td>Post-secondary</td>
<td>0.12</td>
<td>More than 400</td>
<td>0.13*</td>
</tr>
<tr>
<td>Achievement by ability</td>
<td></td>
<td>Achievement by study approach</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>0.35*</td>
<td>Methodology</td>
<td>0.04</td>
</tr>
<tr>
<td>Middle</td>
<td>0.25*</td>
<td>Content</td>
<td>0.35*</td>
</tr>
<tr>
<td>High</td>
<td>0.21*</td>
<td>Material</td>
<td>0.51*</td>
</tr>
<tr>
<td>Achievement by SES</td>
<td></td>
<td>Attitude by study approach</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>0.15</td>
<td>Methodology</td>
<td>0.17*</td>
</tr>
<tr>
<td>Middle</td>
<td>1.05*</td>
<td>Content</td>
<td>-0.07</td>
</tr>
<tr>
<td>High</td>
<td>0.22</td>
<td>Material</td>
<td>0.12</td>
</tr>
<tr>
<td>Achievement by sex</td>
<td></td>
<td>Achievement by source</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>0.11</td>
<td>ERIC document</td>
<td>0.77*</td>
</tr>
<tr>
<td>Female</td>
<td>0.03</td>
<td>Journal article</td>
<td>0.42*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dissertation</td>
<td>0.16*</td>
</tr>
<tr>
<td>Achievement by general content area</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Concepts</td>
<td>0.36*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Computation</td>
<td>0.31*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Application</td>
<td>0.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Achievement in specific subject area</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arithmetic</td>
<td>0.21*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Algebra</td>
<td>0.43*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Geometry</td>
<td>0.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Achievement by program</td>
<td></td>
<td>Achievement by area of</td>
<td></td>
</tr>
<tr>
<td>MSG</td>
<td>0.14*</td>
<td>investigation</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>0.27*</td>
<td>Mathematics</td>
<td>0.24*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Other</td>
<td>0.23*</td>
</tr>
<tr>
<td>Achievement by length of</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>treatment in weeks</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than 9</td>
<td>0.25*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between 9 and 18</td>
<td>0.48*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between 18 and 36</td>
<td>0.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>36 and more</td>
<td>0.25*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grand mean achievement</td>
<td>0.24*</td>
<td>Grand mean attitude</td>
<td>0.12*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*The null hypothesis, stating the mean effect size is zero, is rejected with the probability of committing type I error being less than or equal to 0.05.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
groups.

In the general content area, the effect sizes of concepts (0.36) and of computation (0.31) were very large in achievement, while that of the application was as low as 0.06. Among the specific subject areas, algebra had a high effect size of 0.43 in achievement; however, geometry had a poor effect size of 0.14.

Of the three approaches such as methodology, content, and materials the effect sizes in the content was 0.35 and in the materials 0.51. The methodology had the lowest effect size (0.04) in achievement. Conversely, methodology had the highest effect size in attitude.

The effect sizes when averaged according to the different sources, studies from dissertations showed the least effect size (0.16). The journal articles and the ERIC documents indicated superior effect sizes (0.42 and 0.77, respectively).

Caution must be exercised in the interpretation of the nonsignificant results reported in the tables presented in this chapter. Two or more effect sizes having the same magnitude may differ in their significance. For example, in Table 20, the grand mean effect size in attitude (0.12) is significant, whereas the mean effect size of the post-secondary grades in achievement (0.12) is insignificant. It must be noted that in many instances, it is very likely that the results are insignificant due to small number of effect sizes. Achievement by sex and by socio-economic status are other striking instances of the previously discussed phenomenon.

In total, the grand mean of 660 effect sizes in achievement was
0.24 and the grand mean of 150 effect sizes in attitude was 0.12. In other words, an average student who received some form of modern mathematics instruction was raised from the 50th to the 59th percentile in achievement and from the 50th to the 55th percentile in attitude, over the traditional mathematics population.
CHAPTER V

INTERPRETATIONS, IMPLICATIONS, GENERAL
RECOMMENDATIONS, AND SUMMARY

Interpretations

The interpretations generally followed the sequence of the dis­
cussion of the results in Chapter IV. Some of the mean effect sizes
reported in Chapter IV were not significant. As a rule, interpreta­
tions were focused primarily upon results which were statistically
significant.

The Grand Mean Effect Sizes

The obtained results revealed that there were improvements in
achievement and attitude which were attributable to modern mathemat­
ics. Improvement in achievement of an average student was almost
double the improvement in attitude. The discrepancy between improved
rates of achievement and attitude is typical of innovations where the
appeal is intellectual in nature and demanding in its application.
New mathematics fits this mode in that it is mostly abstract. Con­
sequently, modern mathematics requires a strict discipline of the
mind in order to comprehend its operations. It is consoling to note
that the effects of modern mathematics are worth its demands. Also,
attitude toward an innovative program is often a result of the per­
ceived success or failure of the innovation. Dissimilar findings in
the absence of summarized results helped maintain a low morale toward
modern mathematics for the past 30 years. People were uncertain about whether or not "new math worked." Hopefully, this study has revealed those areas where modern mathematics is effective. Consequently, a more positive attitude toward new mathematics should occur, an attitude which will be comparable to the achievement level.

Subject Variables

Of the four variables identified as subject characteristics, three lend themselves to interpretations. The three variables are grade level, ability level, and socio-economic status. In general, the findings in this group were not in agreement with the expectations and viewpoints of the critics of new mathematics. Consequently, interpretations of these findings must be interesting to the opponents as well as proponents.

Grade level. Among the different grade levels, the mean effect size in achievement of the junior high pupils was the least significant. One explanation of the low mean effect size of the junior high grades is obvious. Confusion may have entered the minds of pupils due to the sudden transition in their thinking patterns as they switched from traditional mathematics in the elementary grades to modern mathematics in the junior high.

The greater mean effect size of the elementary grades where the students were introduced to only one kind of treatment (new mathematics) explains the clear perception of new mathematics concepts in the absence of treatments antithetical in nature. In the case of the senior high students, the large effect size could be assumed to be

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
the result of a development of minds into a stage in which they could
discriminate one treatment from another without confusion.

Ability level. Summarily, the greater the effect, the lower the
ability level was. This finding explains the comparative effectiveness
of the two treatments of mathematics. High ability students,
because of their innate intellectual acumen, do not depend upon the
kind of treatment as much as the low ability students do in the
understanding of mathematics concepts. Consequently, the introduc-
tion of new mathematics effected a comparatively smaller gain in
achievement among the high ability groups while the low ability
groups experienced a large growth in achievement.

Knowing this trend, one would not be tempted to conclude that
modern mathematics was an easy approach. More likely, it was a mean-
ingful approach which produced better results than the traditional
mathematics among all ability levels, especially at the lower levels
of ability.

Socio-economic status. The striking result in the application
of modern mathematics among different levels of socio-economic status
was the outstanding effect size of the pupils coming from middle
class parents. One of the main reasons for the high achievement of
pupils from middle socio-economic levels could be parental help. It
must be pointed out that among the parents who are most attuned to
the trends in mathematics, many are also teachers who generally
belong to the middle socio-economic class. These parents are often
more aware of the need for providing help to their children in under-
standing modern mathematics.
Treatment Variables

The interpretations in the ensuing sections might help to find some answers to the heated criticisms levelled against certain aspects of modern mathematics. These interpretations pertain to the general content area, specific subject area, type of program, and the length of treatment.

**General content area.** In the general content area, which consisted of concepts, computation, and application, the highest mean effect size was that of the concepts. Obviously this is a likely outcome of the treatment of modern mathematics because one of the unique characteristics of modern mathematics is the emphasis on the understanding of concepts. The weakest score in the general content area was in application. The poor result in this area perhaps explains a lack of adequate time for the pupils to exercise their minds. Possibly, a longer treatment of modern mathematics with sufficient emphasis on the concrete presentation of abstract concepts would strengthen the students in the area of application of mathematical concepts.

**Specific subject area.** The specific content area was divided into arithmetic, algebra, and geometry. A fourth category was included to accommodate subject areas other than those previously specified. Among the four subject areas, algebra had the highest mean effect size in achievement. The explanation for this result is subject to speculation. It could be that among all the specific subject areas investigated, algebra was the most expressive medium to
bring out the message of modern mathematics. Similar results can be expected in other subject areas, especially in geometry, if the mode of presentation of those subject areas are more attuned to the axiomatic approach of modern mathematics.

Type of program. The types of programs were grouped in School Mathematics Study Group (SMSG) and non-SMSG. This was due to the small number of studies in all developmental programs except SMSG. The average effect size in achievement of the non-SMSG group was almost twice that of the SMSG group. A number of studies in the non-SMSG group were conducted by individual experimenters. In those experiments, factors such as enthusiasm of the experimenter, attraction of a new theme, and pride of the students in belonging to a selected group, might have been favorable for the great success in modern mathematics treatment. Conversely, in the treatment of SMSG programs, most of the factors mentioned did not exist because they were generally adopted on a massive scale by school districts; therefore, the effect size of non-SMSG programs explains more of an initial enthusiasm. Conversely, the results of the SMSG program indicate a conservative, but steady and longstanding, effect of modern mathematics treatment.

Length of treatment. The effect sizes, when averaged according to the length of treatment, demonstrated an interesting pattern. The average effect in achievement was maximum when the length of treatment was between 9 and 18 weeks and minimum for a period between 18 and 36 weeks. The maximum and minimum values were 0.48 and 0.09, respectively. In the studies with the length of treatment greater
than 36 weeks, the effect sizes were averaged to be 0.25, the mid-value between the two extremes. Therefore, the pattern was a rise, fall, and a gradual steadiness toward the mean of the two values. The rise possibly demonstrates the initial enthusiasm and fascination for the new mathematics programs. The fall manifests the setback the pupils experienced when they were confronted with the rigor and abstractness of the modern math treatment. Once the pupils got over the period of a confrontation, they began making steady improvements which were reflected in the effect sizes of treatments over 36 weeks.

Most of the criticisms that were levelled against modern mathematics during the past decades could have been based on the results obtained during the time of the initial setback. The ebb and tide is over. Now, the "new math" flows steady. This is the time for a mature judgment on the effect of modern mathematics.

**Study Variables**

The interpretations of the study variables highlight some possible explanations of the differences in effectiveness in three major areas. The areas are the study approach, author, and source of the study.

**Study approach.** A polarity existed between the scores in achievement and attitude in the four different types of study approach. Among the four types, namely, methodology, content, materials, and combination, all types except methodology had significantly larger effect sizes in achievement. On the contrary, effect sizes of all types of approaches, except that of methodology, were relatively
small in attitude.

The polarity reveals a striking point. What worked better with modern mathematics was due to an appeal to reason, rather than an appeal to the senses. The methodology approach, which mostly consisted of discovery approaches and its subcategories, might have appealed to the susceptibilities of the young minds; however, little was accomplished in terms of achievement through the new methodologies. What made modern mathematics effective in terms of achievement was its content area characterized by the structure and abstractness.

The mean effect size on the materials is even more revealing. The high achievement effect size of the materials approach makes it explicit that the abstract concepts of modern mathematics, when presented to the students in concrete forms through adequate materials, produced the best results in achievement.

**Author.** The effect sizes in this group were associated with school districts, universities, and private experimenters according to the affiliation of the researcher. A comparatively large mean effect size in achievement for the experimenter group indicates, perhaps, the direction of bias caused by self-interest.

The experimenter group consisted mainly of textbook publishers and other commercial agencies. This group possibly tried to present the most positive results for monetary purposes. Consequently, they would be the most vulnerable in the direction of bias. To a lesser degree, the school districts might have been affected by their preoccupation to build up an exemplar. The results of the studies whose authors were affiliated with universities indicated the least effect
size in achievement. On the whole, effect sizes in that group represented the most conservative and most dependable figures.

**Source of the study.** Sources of the study revealed a similar situation. ERIC documents which consisted of school district reports, final reports from federal funds, etc., had the highest average effect size in achievement. The reason for the high effect size could be that studies from the above sources are often attempts to convince the funding agencies to continue funding of a specific program by presenting the results as attractively as possible. The journal articles also had a large effect size in achievement. Perhaps the journals reveal the increased tendency to publish studies favoring the new approach. Conversely, dissertations with no such self-motives, contained more balanced and neutral findings.

Since the present analysis utilized 85 percent of the total number of studies from complete dissertations, the findings presented here would be considered to be highly conservative and most dependable in nature.

**Design Variables**

Interpretations of the design variables deal with the type of test and the quality of design. The elucidations will help one evaluate the comparative effectiveness of different types in tests and designs.

**Type of test.** The types of tests utilized were commercial and experimenter developed. Effect sizes in achievement of studies measured by commercial tests were very small in comparison with effect
sizes in achievement measured by experimenter developed tests. One of the weaknesses of the standardized tests was that they were not constructed specifically for a study to test the objectives of a particular treatment. As a rule, most of the commercial tests were more suited to measure the objectives of the traditional treatments than those of the modern treatment of mathematics. Consequently, the effect size of the commercial test group did not adequately represent the achievement of students in modern mathematics. It is obvious, therefore, that the overall mean effect size would have been significantly higher than the present one, if all the studies utilized tests which adequately measured the performance according to the objectives specified for the particular treatment.

Quality of design. In the present study, the experiments in which the subjects were randomly assigned to treatments had a higher effect size in achievement than the experiments without random assignment. According to the methodologists, experiments with random assignment should yield results with less bias than the experiments without random assignment. On that basis, the greater effect size in achievement obtained from the random assignment group should be considered to be more objective than the smaller effect size obtained from the studies without random assignment.

Implications

The data presented in Chapter IV and their interpretations in the early sections of Chapter V have several important implications for sound decision making and needed improvements in mathematics.
programs.

The aggregate findings of the experiments with modern mathematics for the last 30 years show the critics and opponents wrong with regard to the detrimental effects of the "new math." On the contrary, the evidence shows the beneficial effects in terms of the improvement in overall achievement and attitude. In the light of the findings of the present study, it would be unwise to advocate the superiority of the traditional mathematics over the new mathematics. Subsequently, the current retrenchment in mathematics under the banner of the back-to-basics philosophy can be regarded only as an aberration brought about by the short sightedness of the doomsayers and the misjudgments of the decision makers in the absence of reliable research findings about the effectiveness of new mathematics.

The beneficial effects of new mathematics call for the promotion of the new programs throughout the nation. The findings should encourage the higher authorities throughout the country, on the federal, state, and local levels of mathematics education, to re-introduce new mathematics curriculum in order to meet the changing needs of the society.

The overall findings also indicated that the improvement of attitude did not compare with that of achievement. The discrepancy between the gains in achievement and attitude has serious implications for the implementation of the new mathematics programs. In the first place, the decision makers should be aware of the existence of the discrepancy. Secondly, in assessing the new mathematics programs, the decision makers should take the achievement score as the index of
effectiveness. Thirdly, in the decisions regarding the implementation of the new mathematics program, they should not be guided by the score in attitude which partly echoes the deteriorating influence of the back-to-basics movement.

The implications based on the results which pertained to the specific variables considered in the study are numerous. An attempt is made in the following sections to discuss the major considerations of the important findings. Among the grade levels, the explicitly superior gains in achievement in the elementary and senior high, in contrast with the junior high, have a serious implication. The superior scores suggest that in the process of introducing new mathematics, the most appropriate grade levels are either the elementary or the senior high. Unfortunately, in the past the majority of the experiments conducted with modern mathematics were at the junior high level. Much of the "anti-feelings" about new mathematics might not have come to the surface, if the new programs were introduced either at the elementary or senior high levels.

One of the allegations against modern mathematics has been that it is good only for the highly motivated students who will go on to major in mathematics in college. The findings of this study proved the allegation wrong. The average mean scores, when categorized according to the ability levels, show that although new mathematics was beneficial for bright students, it was more beneficial for the slow learners. It will be a serious mistake, therefore, to teach the new mathematics only to a chosen ability group—a practice that has been prevalent in the past decades. The consequences would be to
deprive the slow learners of the benefits of the new mathematics. Moreover, when a choice is available regarding the levels of ability and introducing new mathematics, the decision should be to support the low ability groups utilizing new mathematics.

The findings in the general content area are capable of solving much of the heated controversy over the concept approach and the computational skills. The results indicated remarkably superior achievement attributable to new mathematics in concepts and computation. The criticism that modern mathematics is hopelessly beyond young people because of its abstractness and rigor carries no weight in the light of these findings. Also, computation with whole numbers is far from being a lost art. More surprisingly, now Johnny adds better. Knowing that Johnny adds better makes one skeptical of the back-to-basics movement.

As the back-to-basics movement gains momentum, one should be deeply concerned about the dangers it can hold. The movement, insofar as it emphasizes the computational skills in isolation, might eliminate teaching for mathematical understanding. What good can computational skills in isolation do, if a student does not know what to compute when? The modern math has been addressing this problem by placing the computational skill within a total mathematics program. The findings of this study supported the assertion that this approach has been successful, and that it is far superior to the traditional approach. In reality, going back to basics demarcates a deterioration in the learning of mathematics.

The results pertaining to application revealed that the
improvement attributable to new mathematics in the area of application was remarkably low in comparison with the scores in concepts and computation. Unless and until mathematics can operate to clarify and solve human problems, it has indeed a narrow value. Hence, major emphasis must be placed to strengthen the area of application in the new mathematics curriculum. As a result, the students should be able to apply the concepts and skills they learned in the practical field.

The relatively low geometry score in the specific subjects revealed another area which requires serious consideration. Since geometry occupies a central position in mathematics, according to many mathematicians, presenting geometry as meaningful and understandable as the rest of the subject areas is a proven necessity to fulfill the goal of total mathematics program.

On the whole, the new mathematics has been proven to be superior in all areas investigated in this paper. In a few instances, such as geometry, application in the generic content area, and methodology in the general approaches, the improvement attributable to new mathematics is negligible; however, "new math" is still on its way. Further, it is hoped that the leaders in the new mathematics curriculum will recognize these areas as their unfinished task and propose alternatives for improvement.

General Recommendations

The discussion of recommendations pertains to three points. They are recommendations for utilization of the study findings, recommendations for further meta-analyses in the field of mathematics, and
recommendations for improved techniques for meta-analysis.

Utilization of the Study Findings

The present study attempted to resolve several problematic issues pertaining to modern mathematics education. In general, the study tried to find an answer to the central issue: whether the treatment of new mathematics resulted in an improvement or decline of mathematical achievement. The overall analysis of 810 conflicting and dissimilar findings showed that the "new math" raised a typical student from the 50th to the 59th percentile of the traditional mathematics group in achievement. Though to a lesser degree, the introduction of new mathematics also improved the attitude toward mathematics. The attitude of a typical student rose from the 50th to the 55th percentile.

The study also showed the relative strengths and weaknesses of new mathematics in several specific areas. These were categorized under four sets of variables: subject characteristics, treatment factors, study attributes, and design variables. The overall findings and the findings in those specific areas will enable mathematics educators, educational specialists, curriculum developers, classroom teachers, commercial agencies, and other interested persons or agencies to make sound decisions pertaining to such things as grade, ability level, socio-economic status, sex, general content area, specific content area, different subject areas, and type of program in the process of implementing modern mathematics programs.

In summary, the authorities in mathematics education should
realize that a great deal of improvement has been made in the learning of mathematics since World War II. Sputnik is due credit for having the greatest influence on the new mathematics. It would be unwise to ignore these benefits derived from new mathematics and rush "back-to-basics" which holds a promise for only a retarded growth.

Further Meta-analyses on Mathematics

There are many areas in mathematics in which there exists an abundance of studies with dissimilar findings. Meta-analysis would prove to be extremely useful for extracting the message of those diverse findings. A meta-analysis would be indicated in such questions as the relative effectiveness of the application of Piaget's theories and the traditional treatment. Likewise, the literature on instructional television, use of calculators, and such other innovations hold promise for statistical meta-analysis.

Improvements in the Techniques of Meta-analysis

Improvement in the techniques of meta-analysis is another area of exploration. For example, the mean effect size as defined and applied in the literature does not enable one to evaluate its significance. A smaller mean effect size could be of a greater significance than a larger mean effect size if the number of effect sizes is significantly small or the standard deviation is significantly high. An attempt is made in the present study to evaluate the relative merits of the mean effect sizes. The t-values and the corresponding probabilities were calculated to find the significance of each mean effect
Since the concept of meta-analysis is relatively recent, improvements in the techniques such as those discussed above are quite possible and certainly rewarding.

Summary

During the last three decades many innovations in mathematics education were introduced under a common label called "new mathematics." The studies on new mathematics were numerous and their findings were conflicting and dissimilar in nature. No attempt had been made to summarize those studies in order to extract the message of the diverse findings about "new mathematics." The present study attempted to summarize 134 studies on new mathematics, which were located mainly from dissertations, abstracts, journal articles, and ERIC documents.

In order to explain the term "new mathematics," the author traced the origin and evolution of the movement of modern mathematics. There were two critical events, World War II and the Russian Sputnik, which gave great stimuli for reform in mathematics education. The reform movement was initiated mostly by developmental projects which emerged during the 1950's and early part of the 1960's. The role of the developmental programs was to correct perceived weakness in school programs, such as to improve materials, to train teachers, and most importantly to introduce new concepts in the area of content, methodology, and materials in the treatment of mathematics. The new concepts placed great emphasis on logical thinking, structure, and abstractedness, instead of drill and rote of the traditional
mathematics. Innovations of these developmental programs were known under a common label, "new math."

The facts about the conflicting nature of research findings and public opinions about the outcomes of new mathematics, coupled with the scarcity of summative research resolving the dissimilar and conflicting findings, presented the problem for which the study was intended. Research questions were formulated to answer the problematic issues in the new mathematics curriculum. The issues pertained to the examination of overall outcome—decline or growth in the learning of mathematics—attributable to modern mathematics.

Since the study adopted a new technique for summarizing the aggregate findings about new mathematics, the rationale for the study focused on two categories of topics: the justification of the research topic and the justification of the research tool. To justify the topic, three major areas in the literature of modern mathematics were examined. They were: (a) appraisals of modern mathematics, (b) studies on modern mathematics by primary analysis, and (c) studies on modern mathematics by aggregate analysis. Treatment on appraisals of modern mathematics disclosed conflicting viewpoints regarding the efficacy of modern mathematics. To make a sound judgment on the efficacy of modern mathematics, research studies were utilized. An overview of the research studies on modern mathematics by primary analysis revealed that there were numerous studies in that area; however, the findings were diverse. Hence, an aggregate analysis was deemed to be necessary in order to resolve the conflicting findings. The existing aggregate analysis did not strictly pertain to the area of investigation...
of the present study. Therefore, the researcher attempted to utilize an aggregate analysis methodology specifically to resolve the conflicting findings about new mathematics.

The literature search for techniques of aggregation exposed five different methodologies used in the past. They were narratives, reviews, the averaging method, the voting method, and the cluster approach. The drawbacks of the review methodology were: (a) it did not evaluate the influence of the independent variables, (b) it adopted inadequate techniques as means of aggregation, and (c) it eliminated the so called "poorly done" studies. The weakness of the averaging method was that it threw away precisely the information needed most, namely, a true estimate of the treatment effect. The defects of the voting method were: (a) it disregarded valuable descriptive information, (b) it did not integrate the measures of the strength of experimental effects, and (c) it did not solve the problem of unbalanced experimental design. The major drawbacks to the cluster approach were that the study became overly restrictive because: (a) it required access to raw data, (b) it eliminated studies which were not measured in a common scale, and (c) it eliminated studies which lacked explanation for differences.

Because of the drawbacks mentioned previously, none of the techniques was found to be adequate for the present analysis. Glass (1976) introduced another technique of aggregation called meta-analysis. Meta-analysis, in many respects, was an improvement on the previously discussed techniques.

Meta-analysis methodology stood for a statistical aggregation of
a large body of similar, dissimilar, and conflicting findings without loss of information. The key concept in the technique of meta-analysis was the effect size, which was defined as the difference between the means of the experimental and control groups measured in standard deviations of the control group. Since the effect sizes were standardized measures, comparison of different effect sizes were possible. Practical steps to conduct the present study consisted of the identification of the sources of the study, locating and collecting the studies, identifying and coding variables, and methods of analyzing the data.

The overall mean effect sizes of 810 substudies revealed that modern mathematics showed evidence of change attributable to new mathematics programs. The change was reflected in the improvement of academic achievement and in the improvement of attitude toward mathematics. The results also revealed differential effects of specific characteristics of subject variables, treatment variables, study variables, and design variables. Spectacular improvements attributable to modern mathematics have been made in the understanding of the concepts and computational skills. Among the specific subject areas, algebra had an outstandingly superior achievement score. Of the different approaches of modern mathematics, the one that emphasized the concrete presentation of abstract ideas was indicated to be the most rewarding. The major areas which revealed a need for substantial improvements were geometry among the specific subjects and application in general content area.

The findings of the study have several important implications for
sound decision making and needed improvements in many specific areas. On the whole, the findings of the study strongly encourage the promotion of the new mathematics throughout the nation. At the same time, the back-to-basics movement should certainly be regarded as an aberration in the advancement of the learning of mathematics.

In summary, the present study did show evidence that the experiments with modern mathematics of the past three decades in its content, methodology, materials, or in any combination, have, in general, led to improved learning of mathematics. This is revealed by the progress in achievement and attitude toward mathematical concepts, computation, and application on all grade and ability levels. Therefore, findings of the present study should motivate the decision makers to dispel their unfounded uncertainties and misguided apprehensions about the effectiveness of new mathematics and to bring forth the new mathematics to a respectable place in the curriculum for the benefit of the students and for the betterment of the society.
BIBLIOGRAPHY


*Entries that are preceded by an asterisk were used as sources of data for the analyses of the present study.

112

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


Bartram, C. E. An analysis and synthesis of research relating to selected areas in the teaching of arithmetic (Volumes I and II) (Doctoral dissertation, Ohio State University, 1956). Dissertation Abstracts, 1956, 16, 2101A. (University Microfilms No. 56-3,334)


Betz, W. The present situation in secondary mathematics, with particular reference to the new national reports on the place of mathematics in education. The Mathematics Teacher, 1940, 33 (8), 339-360.


*Bush, W. J. An analysis of the outcomes of functional mathematics and formal mathematics as measured by certain objective tests after completion by the students of two years of study (Doctoral dissertation, University of Arkansas, 1959). Dissertation Abstracts, 1959, 29, 920A. (University Microfilms No. 59-3036)


Clark, J. R. Looking ahead at instruction in arithmetic. The Arithmetic Teacher, 1961, 8, 388.

*Cronin, R. E. The effects of varying amounts of traditional and modern mathematics instruction relative to sex and intellectual ability on both the traditional and modern mathematics achievement of eighth grade pupils (Doctoral dissertation, The Catholic University of America, 1967). Dissertation Abstracts, 1968, 28, 2551A. (University Microfilms No. 67-17,139)


*Cummins, K. B. A student experience-discovery approach to the teaching of calculus (Doctoral dissertation, Ohio State University, 1958). Dissertation Abstracts, 1959, 19, 2292A. (University Microfilms No. 59-367)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


DeVault, M. V., & Kriewall, T. E. Perspectives in elementary mathematics. Columbus, Ohio: Charles E. Merrill, 1969.


Fehr, H. F. Reorientation in mathematics education. Teachers College Record, 1953, 54, 430-439.


*Greathouse, J. J. An experimental investigation of relative effectiveness among three different arithmetic teaching methods (Doctoral dissertation, The University of New Mexico, 1965).* Dissertation Abstracts, 1966, 26, 5913A. (University Microfilms No. 66-4465)

*Greitzer, S. L. A comparison of the postulational approach and the traditional approach in teaching selected topics in algebra to above average students (Doctoral dissertation, Yeshiva University, 1959).* Dissertation Abstracts, 1960, 21, 1137A. (University Microfilms No. 60-02108)


Hill, S. Issues from the National Advisory Committee on Mathematics Education (NACOME) report. The Mathematics Teacher, 1976, 69, 440-446.

*Hollis, L. Y. A study to compare the effects of teaching first grade mathematics by the Cuisenaire-Gattegno method with the traditional method (Doctoral dissertation, Texas Technological College, 1964). Dissertation Abstracts, 1965, 26, 904A. (University Microfilms No. 65-1458)


*Klingbeil, D. H.* An examination of the effects of group testing in mathematics courses taught by the small group discovery method (Doctoral dissertation, University of Maryland, 1974). *Dissertation Abstracts International*, 1975, 36, 849A. (University Microfilms No. 75-17,886)


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


Muller, D. J. Logic and the ability to prove theorems in geometry (Doctoral dissertation, The Florida State University, 1975). Dissertation Abstracts International, 1975, 36, 51A. (University Microfilms No. 75-17,944)


Nichols, D. E. *Comparison of two approaches to the teaching of selected topics in plane geometry* (Doctoral dissertation, University of Illinois, 1956). *Dissertation Abstracts,* 1956, 16, 21906A. (University Microfilms No. 18,178)


*Phelps, J. A study comparing attitude towards mathematics of SMSG and traditional elementary school students (Doctoral dissertation Oklahoma State University, 1963). Dissertation Abstracts, 1964, 25, 1052A. (University Microfilms No. 64-8942)


Pingry, R. E. For a better mathematics program in the junior high school. The Mathematics Teacher, 1956, 49, 120.


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


Rosskopf, M. F. Trends in content of high school mathematics in the United States. Teachers College Record, December 1954, 56, 137.

*Rowlett, J. D. An experimental comparison of direct detailed and direct discovery methods of teaching orthographic projection principles and skills (Doctoral dissertation, University of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


SchAAF, W. L. How modern is modern mathematics? The Mathematics Teacher, February 1964, 57, 90.


*Steadman, D. G. An evaluative study comparing achievement and attitude between the SSMCS unified mathematics programs and traditional accelerated mathematics programs in the Utah county public


*Tucker, J. A junior high school level experiment on developing algebra readiness by the use of finite non-numerical algebraic systems (Doctoral dissertation, Auburn University, 1969)


*UrWiller, S. L.* A comparative study of achievement, retention, and attitude towards mathematics between students using spiral homework assignments and students using traditional homework assignments in second year algebra (Doctoral dissertation, University of Nebraska, 1971). *Dissertation Abstracts International,* 1971, 32, 845A. (University Microfilms No. 71-19,521)


Wozencraft, M. Are boys better than girls in arithmetic? The Arithmetic Teacher, December 1963, 10, 486-490.


### Appendix A

**The List of the Variables Identified**

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Purpose</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Study number</td>
<td>To group the effect sizes by category</td>
<td>One assigned to each study</td>
</tr>
<tr>
<td>2. Study approach</td>
<td>To specify the modern mathematics treatment in accordance with the emphasis placed on the approach</td>
<td>1 = methodology, 2 = content, 3 = materials, 4 = combination</td>
</tr>
<tr>
<td>3. Study type</td>
<td>To differentiate between achievement and attitude studies</td>
<td>1 = achievement, 2 = attitude</td>
</tr>
<tr>
<td>4. Author</td>
<td>To identify the innovator</td>
<td>1 = commercial, 2 = school district, 3 = experimenter, 4 = university, 5 = other</td>
</tr>
<tr>
<td>5. Year of study</td>
<td>To determine the year of publication</td>
<td></td>
</tr>
<tr>
<td>6. Source of study</td>
<td>To determine where study was located</td>
<td>1 = public school report, 2 = ERIC, 3 = journal articles, 4 = dissertations</td>
</tr>
<tr>
<td>7. Affiliation of researcher</td>
<td>To specify the organization the researcher worked for</td>
<td>1 = staff of a project, 2 = school district evaluation staff, 3 = university, 4 = other</td>
</tr>
<tr>
<td>8. Involvement of researcher</td>
<td>To determine researcher's personal dependency on the outcome of the study</td>
<td>1 = evaluating on material or project, 2 = none</td>
</tr>
<tr>
<td>9. Area of study</td>
<td>To define the area of investigation</td>
<td>1 = mathematics, 2 = other subject areas included</td>
</tr>
</tbody>
</table>
### Variable Name  
**Purpose**  
**Values**

10. Specific subject area  
To describe what specific subject was taught  
1 = arithmetic  
2 = algebra  
3 = geometry  
4 = sets  
5 = logic  
6 = statistics  
7 = calculus  
8 = trigonometry  
9 = other

11. Type of program  
To identify the specific innovative program  
1 = UICSM  
2 = SMSG  
3 = GCMP  
4 = MADISON  
5 = BCMI  
6 = BSU  
7 = UMMaP  
8 = MINNEMAST  
9 = OTHER

12. Length of treatment  
To determine the duration of the treatment  
Number of weeks until posttest

13. Length of treatment  
To determine the duration of the treatment  
Days per week

14. Length of treatment  
To determine the duration of the treatment  
Minutes per day

15. Time  
To find the total number of instruction hours  
Weeks x Days x Minutes / 60

16. Number  
To find the sample size  
The number of subjects that means are based on

17. New versus field-tested  
To determine if the treatment was field-tested or new  
1 = new—never used before  
2 = field tested

18. Grade  
To determine the grade or subjects  
1 = KG  
2 = Elementary  
3 = JH  
4 = SH  
5 = Post-secondary
<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Purpose</th>
<th>Values</th>
</tr>
</thead>
</table>
| 19. SES       | To determine the socio-economic status of subjects | 1 = unknown  
2 = low  
3 = middle  
4 = high |
| 20. Ability   | To determine the ability levels of subjects | 1 = unknown  
2 = low  
3 = middle  
4 = high |
| 21. Sex       | To determine the sex of subjects | 1 = unspecified  
2 = female  
3 = male |
| 22. Test      | To determine the type of test | 1 = commercial  
2 = experimenter developed  
3 = teacher developed  
4 = other |
| 23. Retention | To identify the length of retention | Weeks between end of treatment and testing |
| 24. Effect size | To calculate measures effectiveness | |
| 25. Method for effect sizes | To determine the technique to be used to measure effect size | 1 = control group  
2 = ANOVA  
3 = gain scores  
4 = grade equivalent scores  
5 = standardized tests  
6 = t-test  
7 = F-test  
8 = ANCOVA |
| 26. Random assignment | To determine the quality of design | 1 = random assignment  
2 = no random assignment |
| 27. Novelty effect | To determine the novelty effect | 1 = present  
2 = not present |
<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Purpose</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>28. Dependent variable</td>
<td>To identify overlapping studies</td>
<td>1 = unique data</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 = subtests when total included</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 = subtests when total not included</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4 = total when subtests included</td>
</tr>
<tr>
<td>29. General content area</td>
<td>To determine the general content area emphasized</td>
<td>1 = concepts</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 = computation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 = application</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4 = combination</td>
</tr>
<tr>
<td>30. Nature of test</td>
<td>To determine the category of test</td>
<td>1 = retention</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 = recall</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 = transfer</td>
</tr>
<tr>
<td>31. Pool</td>
<td>To determine the number of effect sizes in each study</td>
<td></td>
</tr>
</tbody>
</table>
## Appendix B

Code Sheet of the Variables Identified

| Reference |

<table>
<thead>
<tr>
<th>CODE</th>
<th>COLUMN</th>
<th>VARIABLE</th>
<th>VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-4</td>
<td>STUDY NO</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>STUDY TYPE</td>
<td>1 Met 2 Cont 3 Mat 4 Comb</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>SUBSTUDY NO</td>
<td>1 Ach 2 Att</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>AUTHOR</td>
<td>1 Com 2 Sc D 3 Exp 4 Uni 5 Ot</td>
<td></td>
</tr>
<tr>
<td>8-9</td>
<td>YEAR OF STUDY</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>SOURCE</td>
<td>1 P Sc 2 ERIC 3 JA 4 Diss 5 Ot</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>STUDY RATING</td>
<td>1 Rand 2 ECG 3 NECG 4 NC</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>AFFI OF RES</td>
<td>1 SP 2 Sc ES 3 Uni 4 Ot</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>INV OF RES</td>
<td>1 0 wn 2 None</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>SUB AREA</td>
<td>1 Math 2 Sc ES 3 SS 4 Lang 5 Comb</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>METHOD</td>
<td>1 Reas 2 Act 3 Gru 4 Ot</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>CONTENT</td>
<td>1 Arit 2 Alg 3 Geo 4 Set 5 Log</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>TYPE OF PGM</td>
<td>1 UICSM 2 SMSG 3 GCMP 4 MAD 5 BCMI</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>6 BSU 7 UMMaP 8 MIN 9 OT</td>
<td></td>
</tr>
<tr>
<td>18-20</td>
<td>LENGTH OF TREATMENT</td>
<td>WEEKS</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>LENGTH OF TREATMENT</td>
<td>DAYS</td>
<td></td>
</tr>
<tr>
<td>22-24</td>
<td>LENGTH OF TREATMENT</td>
<td>MINUTES</td>
<td></td>
</tr>
<tr>
<td>25-27</td>
<td>TIME</td>
<td>WEEKS X DAYS X MINUTES</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>28-31</td>
<td>NUMBER AVERAGE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>FIELD-TESTED</td>
<td>1 New 2 Field T 3 Unk</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>GRADE</td>
<td>1 KG 2 Ele 3 JH 4 SH 5 Post Sec 6 Ot</td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>SES</td>
<td>1 Unk 2 L 3 M 4 H</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>ABILITY</td>
<td>1 Unk 2 L 3 M 4 H</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>SEX</td>
<td>1 Both 2 Male 3 Female</td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>TEST</td>
<td>1 Comm 2 Ex D 3 Tea D 4 Ot</td>
<td></td>
</tr>
<tr>
<td>38-40</td>
<td>RETENTION TIME</td>
<td>WEEKS</td>
<td></td>
</tr>
<tr>
<td>41-48</td>
<td>EFFECT SIZE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>METOD ES</td>
<td>1 CT 2 ANOVA 3 GAIN SC 4 GES 5 ST</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>6 t-T 7 F-T 8 ANCOVA 9 Ot</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>ASSIGNMENT</td>
<td>1 Rand 2 No Rand</td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>NATURE OF TEST</td>
<td>1 Ret 2 Rec 3 Tran</td>
<td></td>
</tr>
<tr>
<td>52</td>
<td>NOVELTY EFFECT</td>
<td>1 Yes 2 No</td>
<td></td>
</tr>
<tr>
<td>53</td>
<td>DEP VAR</td>
<td>1 Uni 2 STWTI 3 STWTNI 4 TWSI</td>
<td></td>
</tr>
<tr>
<td>54</td>
<td>CONT GEN</td>
<td>1 Con 2 Comp 3 Appl 4 Gen</td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>TRANSFER TIME</td>
<td></td>
<td></td>
</tr>
<tr>
<td>56-60</td>
<td>POOL</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.