

INFORMATION TO USERS

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

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

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

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

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

U·M·I

University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700 800/521-0600

Order Number 9803278

**A CBA model's effect on middle school students in math
achievement**

Bornfield, Alva Jo Anne Gail, Ph.D.

The University of Arizona, 1992

Copyright ©1992 by Bornfield, Alva Jo Anne Gail. All rights reserved.

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

A CBA MODEL'S EFFECT ON MIDDLE SCHOOL STUDENTS
IN MATH ACHIEVEMENT

by

Alva Jo Anne Gail Bornfield

Copyright © Alva Jo Anne Gail Bornfield 1992

A Dissertation Submitted to the Faculty of the
DEPARTMENT OF SPECIAL EDUCATION AND REHABILITATION

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY
WITH A MAJOR IN SPECIAL EDUCATION

In the Graduate College
THE UNIVERSITY OF ARIZONA

1 9 9 2

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Final Examination Committee, we certify that we have read the dissertation prepared by A. J. Gail Bornfield entitled EFFECTS OF A CBA MODEL ON MIDDLE SCHOOL STUDENTS IN MATH ACHIEVEMENT

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy

William C. Healey 6/25/92
Date
William C. Healey

James C. Chalfant 6/25/92
Date
James C. Chalfant

Joel E. Macht 6/25/92
Date
Joel E. Macht

M. Angele Thomas 6/25/92
Date
M. Angele Thomas

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copy of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

William C. Healey 6/25/92
Date
Dissertation Director William C. Healey

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under the rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgement of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the copyright holder.

SIGNED: 

ACKNOWLEDGEMENTS

I would like to express my appreciation to Dr. William Healey, my major advisor, mentor, and director of this work, for his expertise and guidance on this project and through my doctoral studies. I would also like to express my gratitude to Dr. James Chalfant and Dr. Joel Macht for their assistance with this project. Appreciation also goes to Dr. Angele Thomas and Dorothy Friedman for their support and assistance.

A very special thank you goes to my children, Suzy, Cole, Jeremiah, and Joshua, for the patience, support, encouragement, and love they have provided me. They never doubted that I would complete this course of study and supported me on every adventure my studies led me through.

TABLE OF CONTENTS

	Page
LIST OF ILLUSTRATIONS.....	7
LIST OF TABLES.....	9
ABSTRACT.....	10
1. INTRODUCTION.....	11
Statement of Need	16
Purpose of Study	19
Research Questions.....	19
Significance of the Study.....	19
Limitations & Definitions.....	20
Limitations.....	20
Definitions.....	20
2. REVIEW OF LITERATURE.....	22
Mathematics Performance of Middle School Students.....	22
Types of Mathematical Problems.....	23
Theories of Development.....	24
Piaget.....	24
Bruner.....	29
Dienes.....	30
Problems Intrinsic to the Student.....	31
Problems Extrinsic to the Student.....	34
Approaches Related to the Assessment of Mathematical Performance.....	37
Standardized Achievement Tests.....	40
Definition & Purpose.....	40
Uses & Outcomes of Standardized Tests.....	43
Authentic Assessment.....	46
Definition & Purpose.....	46
Uses & Outcomes of Authentic Assessment.....	48
Curriculum-Based Assessment.....	49
Uses & Outcomes of Curriculum-Based Assessment.....	51
3. METHODS.....	55
Research Design.....	55
Variables.....	56
Subjects.....	57
Instrumentation.....	59
Treatment Groups.....	60
Experimental Treatment.....	61
Group A.....	61
Baseline (Phase A).....	61
Treatment Phase (Phase B).....	64
Maintenance Phase (Phase C).....	65
Group B.....	65

TABLE OF CONTENTS - Continued

	Page
4. RESULTS.....	67
Subject 1.....	69
Subject 2.....	72
Subject 3.....	78
Subject 4.....	80
Subject 5.....	86
Subject 6.....	90
Subject 7.....	96
Subject 8.....	102
Subject 9.....	108
Summary.....	112
5. CONCLUSIONS & FUTURE IMPLICATIONS.....	113
Research Question 1: Does the use of CBA methods increase mathematical skills of middle school students who are at-risk of failing mathematics?.....	113
Research Question 2: Are differential effects in mathematical achievement observed between two heterogenous groups of middle school students when the Experimental Group instruction uses the CBA and traditional methods, while the Control Group has traditional instruction only?.....	115
Research Question 3: Are gains achieved through the CBA Model maintained after the assessment and instruction are terminated for a period of 14 to 18 days?.....	117
Problems Encountered in this Study.....	118
Implications.....	119
Recommendations for Future Research.....	120
APPENDIX A - HUMAN SUBJECTS APPROVAL.....	123
REFERENCES.....	126

LIST OF ILLUSTRATIONS

Figure	Page
1. Subject 1 - The Effects of Treatment on Addition of Unlike Fractions.....	71
2. Subject 2 - The Effects of Treatment on Subtraction with Regrouping.....	73
3. Subject 2 - The Effects of Treatment on Addition of Like Fractions.....	75
4. Subject 2 - The Effects of Treatment on Reducing Fractions.....	77
5. Subject 3 - The Effects of Treatment on Simple Division.....	79
6. Subject 4 - The Effects of Treatment on Unlike Fractions.....	81
7. Subject 4 - The Effects of Treatment on Changing Fractions to Percents.....	83
8. Subject 4 - The Effects of Treatment on Multiplication of Fractions.....	85
9. Subject 5 - The Effects of Treatment on Division with Zeros.....	87
10. Subject 5 - The Effects of Treatment on Addition of Unlike Fractions.....	89
11. Subject 6 - The Effects of Treatment on Addition of Unlike Fractions.....	91
12. Subject 6 - The Effects of Treatment on Changing Fractions to Percents.....	93
13. Subject 6 - The Effects of Treatment on Multiplication of Fractions.....	95
14. Subject 7 - The Effects of Treatment on the Addition of Like Fractions.....	97
15. Subject 7 - The Effects of Treatment on Addition of Unlike Fractions.....	99
16. Subject 7 - The Effects of Treatment on Reducing Fractions.....	101

LIST OF ILLUSTRATIONS - Continued

Figure		Page
17.	Subject 8 - the Effects of Treatment on Changing Fractions to Percents.....	103
18.	Subject 8 - The Effects of Treatment on Multiplication of Fractions.....	105
19.	Subject 8 - The Effects of Treatment on Division of Fractions.....	107
20.	Subject 9 - The Effects of Treatment on Changing Fractions to Percents.....	109
21.	Subject 9 - The Effects of Treatment on the Multiplication of Mixed Numbers.....	111

LIST OF TABLES

Table	Page
1. Percentage Correct for Each Skill by Subject	
Under Each Condition.....	112
Group A.....	112
Group B.....	112

ABSTRACT

The present study was an investigation of the effects of a CBA model on mathematical achievement of middle school students. Nine subjects in the seventh and eighth grades were selected to participate in the study. A multiple baseline single subject design was used.

Results indicated that a CBA model designed in the form of a pull-out program can be very effective in identifying and remediating problems in mathematics for middle school students who are at-risk for failing mathematics. Implications of the current findings for the use of a CBA model were discussed. The use of a CBA model for students identified as at-risk for failing mathematics in middle school was supported.

CHAPTER I

INTRODUCTION

... there are very few persons, either male or female, ... who have not occasions daily to make use of arithmetic ... And the person most ready in calculation is much the most likely to succeed in business of any kind. As our country becomes more thickly peopled, and further progress is made in the arts, and new arts are discovered, knowledge of all kinds is brought into acquisition; and none more so than that of arithmetic, and the higher branches of mathematics, of which arithmetic is the foundation.

Warren Colburn, 1830

During 1990, the Mathematical Sciences Education Board and the National Research Council issued a joint report stating, " ... As the economy adapts to information-age needs, workers in every sector - from hotel clerks to secretaries, from automobile mechanics to travel agents - must learn to interpret intelligently computer-controlled processes ... students need more mathematical power in school as preparation for routine jobs ... The changing demands of the workplace exert extraordinary burdens on mathematics education, burdens that we have not successfully borne."

During the early years of the twenty-first century, when today's students will enter the workforce, most jobs will require greater mathematical skills (Johnson & Packer, 1987). Simultaneously, a global economy is emerging as a direct force in American society. However, many reports indicate that students of the United States are not competitive in their mathematical accomplishments with students of other countries (Stevenson, Lee, & Stigler, 1986; McKnight, Crosswhite, Dossey, Kifer, Swafford, Travers, & Cooney, 1987; Stigler & Perry, 1988; Lapointe et al.,

1989). The mathematical skills of our nation's children are generally insufficient to cope with either on-the-job demands for problem-solving or college expectations for mathematical literacy (Steen, 1989).

The National Assessment of Educational Progress (NAEP) conducted during 1990 is a nationally representative and continuing assessment of what students in the United States can do in the mathematics area. The sample selected for the NAEP was comprised of 26,000 students in 1300 schools across 40 states. Three grade levels (fourth, eighth, and twelfth) were sampled in the content areas of numbers and operations; measurement; geometry; data analysis, statistics, and probability; and algebra and functions. The results of the NAEP Mathematics Assessment are as follows:

Fourth Grade. Approximately 72 percent of the fourth graders demonstrated the ability to consistently solve simple addition and subtraction problems with whole numbers. This material is typically taught in the third grade. Eleven percent of these students demonstrated a grasp of multiplication and two-step problems - material usually covered in the fifth grade. No fourth graders demonstrated a consistent grasp of fractions, decimals, percents, or simple algebra. These data reflect that fourth graders have learned skills appropriate to their grade level.

Eighth Grade. Ninety-eight percent of all eighth graders demonstrated a grasp of adding and subtracting with whole numbers (third grade material). Approximately 66 percent demonstrated consistent success with multiplication and division of whole numbers which includes problems involving more than one-step (fifth grade content). Only 14 percent consistently demonstrated successful performance with problems involving fractions, decimals, percents, and simple algebra (seventh grade material). No eighth graders demonstrated success with

reasoning and problem-solving involving geometry, algebra, and beginning statistics and probability.

Twelfth Grade. All high school seniors demonstrated success with third grade material, while 91 percent demonstrated mastery of fifth grade content. This would indicate that nine percent of all students graduating from high school do not know how to apply the four basic arithmetic operations to solve simple problems with whole numbers. Forty-six percent of the seniors demonstrated a consistent grasp of decimals, percents, fractions, and simple algebra. Only five percent demonstrated an understanding of geometry and algebra at a level that suggested preparedness for the study of advanced mathematics. To assess practical application of mathematics to daily living, students were asked to determine the cost of a meal from a menu. Thirty-seven percent of the fourth graders, 66 percent of the eighth graders, and 77 percent of the twelfth graders accurately calculated the cost of the meal. Given the task involved which is only addition of menu items, it would be expected that twelfth graders would perform the task correctly as they demonstrated mastery of third grade material.

In summary, the mathematical skills of our nation's children are insufficient to cope with either on-the-job demands for problem-solving or college expectations for mathematics literacy (Steen, 1989). Due to the importance of mathematics in education, citizenship, and careers, business and industry are spending billions of dollars to train personnel, while colleges and universities are donating large amounts of resources to remediation (Willoughby, 1990).

The National Research Council (1989) reports that the growth of technology, increased applications, impact of computers, and the expansion of mathematics have combined in the past 25 years to extend both the scope and application of the

mathematical sciences. The mathematics of today involves more than calculation. It involves clarification of the problems, deduction of consequences, facilitation of alternatives, and the development of tools.

The learning of mathematics is a vital component of the education of every student (Mathematical Sciences Education Board, 1990). The study of mathematics is important for several reasons: to learn practical skills for daily living, to understand quantitative aspects of public policy, to develop problem-solving skills, and to prepare for careers.

Over the past 30 years, the traditional mathematics curriculum has undergone a series of changes (Capps & Cox, 1991). The movement has encompassed new math, back-to-basics, problem-solving, and currently change emphasized in the *Curriculum and Evaluation Standards for School Mathematics* (National Council for Teachers of Mathematics, 1989). The current changes emphasize that mathematics include problem-solving, communication, reasoning, and relatedness of ideas internal and external to the discipline.

Instructional practices affect mathematics achievement. Currently, teachers and students report that textbooks and worksheets are the predominant materials used in teaching mathematics (Mullis et al, 1991). Over one-half of the fourth graders sampled in the NAEP worked problems from their textbooks on a daily basis and completed worksheets at least weekly. The textbook also was the primary instructional source of the upper grades, with approximately 70 percent of the eighth graders and 81 percent of the twelfth graders working problems from their textbooks on a daily basis. Worksheets were used in the higher grades with most students but more frequently with lower-performing students. In-class testing occurred regularly in all three grades.

In contrast, approximately one-third of the students across all three grades reported never working in small groups or with manipulatives and tools such as counting blocks, rulers, or geometric shapes. Both teachers and students reported that mathematics reports or projects were done infrequently, if at all.

Students were grouped by ability for mathematics instruction. At the fourth grade level, more than one-half were in classes with students of similar ability. At grade eight, more than two-thirds of the students were grouped into differentiated curricula. At the twelfth grade, 58 percent of students reported being enrolled in an academic high-school program, 34 percent in a general program, and eight percent in a vocational/technical program. For those seniors in an academic high school program, less than three-fourths reported taking Algebra II. Grouping students into tracks determines the level of mathematics coursework to which they can enroll.

The Mathematics Framework for California Public Schools (1991) reports that traditional teaching practices provide for daily assignment of two-page lessons given from the textbook to be done individually and characterized as "drill and practice." These two-page lessons given each day are organized around a specific objective. The lesson typically includes an application example; instruction in how to do a specific procedure, perhaps with some simplified explanation; a few exercises to "check for understanding"; practice exercises; and optional "enrichment problems."

Approximately the first third of the school year is spent reviewing computation and symbolic manipulation topics that have been previously taught. As the school year progresses, new material is presented in upcoming chapters, but it is reviewed infrequently as on semester tests.

James Flanders (1987) analyzed three of the better selling textbooks series in 1987 to see how much new material was introduced in each grade. Credit was given to a book for a "new" page if any new material appeared on that page. The average percentages of "new" pages for the 3 series for each grade from kindergarten through ninth are:

<u>Grade</u>	<u>Percentage</u>
K	100
1	75
2	40
3	60
4	45
5	50
6	38
7	35
8	30
9	90

In summary, mathematics instruction appears to be characterized by classrooms grouped by ability, curriculum dictated by textbooks and standardized tests (Price & Gawronski, 1981), and inconsistent introduction of new material with a heavy emphasis on annual review (Flanders, 1987).

Statement of Need

To function in today's society, literacy in mathematics is as essential as verbal literacy (Cockcroft, 1986). Without the ability to read and understand, no one can become mathematically literate. However, the reverse is also true. Without the

ability to understand basic mathematic concepts, one cannot fully comprehend the implications of many newspaper articles. To sufficiently cope with the demands of society, one must be able to grasp the implications of mathematical ideas. For example, choice, logic, and graphs permeate the daily news and are also a part of routine decision-making.

With so many students experiencing poor performance in mathematics, it is reasonable to expect that teachers will seek assistance for these students through referrals to special programs in the schools. Special education is one such program to which teachers can refer students who demonstrate a severe skill discrepancy.

Students enter the special education assessment process to determine the existence of a suspected disability as a result of a referral by a teacher, parent, or other individual involved with the child. Once students are referred, they are likely to be evaluated and placed into special education programs. In one study, 92% of the students referred were evaluated and 78% were found disabled and declared eligible for special education services (Ysseldyke, 1983).

Although neither national nor state data are available regarding referral for evaluation for special education placement resultant from a mathematics discrepancy in achievement, an informal records review of 107 students placed into special education programs as learning disabled during the 1990-91 school year within Tucson Unified School District found that 12 of these students were placed for an achievement discrepancy in the area of mathematics. Another 19 students who were placed experienced a combination of achievement discrepancies that included mathematics. Thus, 31 of 107 students (29%) placed in special education for a learning disability needed specialized instruction in mathematics.

The Thirteenth Annual Report to Congress on the Implementation of the Individuals with Disabilities Education Act (1991) describes an increase of 160.4% in the population of students identified as Learning Disabled (LD) over a 13 year period (1976-1977 to 1989-1990) In the same report, a comparison of the most recent school years (1988-1989 to 1989-1990) indicates a national increase of 3.45% in the number of students served within the category of LD. The continuing increases in the placement of students identified as LD into special education have correspondingly high costs and administrative implications. A need exists to develop, implement and evaluate alternative assessment procedures and instructional strategies. One assessment and instructional procedure that has gained prominence is curriculum-based assessment (CBA).

Gickling and Havertape (1981) defined curriculum-based assessment (CBA) as "... a procedure for determining the instructional needs of a student based on the student's ongoing performance within existing course content." This definition contains three important points (Tucker, 1991). First, CBA is a procedure for determining the instructional needs of a student. Thus, by inference, assessment data must provide a direct link to instructional strategies in order to be considered CBA.

Secondly, the definition implies frequent assessment when it refers to ongoing performance of the student. Tucker (1991) interprets "frequent" to mean several times each week. Finally, all CBA data are derived from the student's performance within existing course content which refers to the curriculum of the educational program of the school that the child attends.

Several studies have examined the use of the CBA model of assessment and instruction in the area of language arts which includes reading (Russell, 1986;

Truesdell, 1987; Wesson, 1989; Fuchs, 1989; Shinn, 1989; Cawley, 1990; Frank & Gerkin, 1990; Potter & Wamre, 1990; Bean & Lane, 1990; Gable, 1990; and Heshusius, 1991). The CBA Model needs to be evaluated as a method to assess and instruct students without the traditional assessment and placement processes for required for special education services.

Purpose of Study

The purpose of the present study is to determine the effects of curriculum-based assessment and instructional methods (CBA model) on the performance of middle-school students who are at-risk of failing mathematics.

Research Questions

This study addresses the following research questions:

1. Does the use of CBA methods increase mathematical skills of middle school students who are at-risk for failing mathematics?
2. Are differential effects in mathematical achievement observed between two heterogenous groups of middle school students when the Experimental Group instruction uses the CBA and traditional methods, while Control Group has traditional instruction only?
3. Are any gains achieved through the CBA model maintained after the assessment and instruction are terminated for a period of 14 to 18 days?

Significance of the Study

Increased numbers of students continue to be identified as learning disabled placed into special education programs. However, pre-referral strategies that assist teachers in evaluation and instructional remediation of achievement difficulties of students may be sufficient means for addressing some learning problems as they occur within the curriculum. At present, schools, teachers, and administrators are

investigating approaches to evaluating students' learning (Chittenden, 1991). Goals identified for assessment practices include: capitalizing on actual work in the classroom, enhancing teacher and student involvement in evaluation, and meeting accountability concerns of the district.

Traditionally, the academic achievement of students has been assessed by administering a variety of standardized, norm-referenced tests (Gable, et al., 1991). There is a growing concern that these tests fail to provide sufficient clarity regarding the task of teaching (Choate, et al., 1987). In response to the limitations of conventional psychometrics, practitioners are looking increasingly toward the curriculum itself for purposes of assessment. CBA is acknowledged as one method for classroom teachers to link assessment and instruction.

Results of this study should yield important information regarding a) use of the CBA model for evaluation and instruction of middle school students who are at-risk for failure in mathematics, b) effective pre-referral practices as well as alternatives to costly and time-consuming evaluations to determine special education eligibility.

Limitations and Definitions

Limitations

This study is limited to a predetermined number of treatment sessions. Students who may otherwise attain skills to the level of the curriculum in which they are placed if given enough instructional time using the CBA Model may not have the opportunity to do so in this study.

Definitions

Several terms in the study require definition:

1. Mathematics is defined as a science and language of patterns.

2. Mathematical power is the ability to learn to read, write, and speak about mathematics.
3. An assessment is any process or procedure employed in collecting information on performance (Price & Gawronski, 1981).
4. Evaluation involves relating descriptions of behavior to other variables or making comparisons among various descriptions of behavior (Price & Gawsonski, 1981).
5. The Present Performance Level (PPL) is the level at which individuals are presently performing at or near 100 percent (Macht, 1990).
6. Error analysis refers to analyzing the type, quantity, and quality of errors the individual makes (Macht, 1990).

CHAPTER II

REVIEW OF LITERATURE

For the purpose of this study, the literature review addressed the following topics: (1) mathematics performance of middle school students, (2) types of mathematical problems experienced by these students, (3) approaches related to assessment of mathematics performance, and (4) curriculum-based assessment model.

Mathematics Performance of Middle School Students

Overall, the NAEP 1990 Assessment of the Nation found that 98 percent of all eighth graders demonstrated an understanding of third grade material - adding and subtracting whole numbers. Two-thirds of the eighth graders demonstrated that their mathematics understanding included consistent success with multiplication and division of whole numbers, or problems involving more than one step (fifth grade material). Only 14 percent demonstrated successful performance with problems involving fractions, decimals, percents, and simple algebra - topics generally covered by the seventh grade. No eighth graders demonstrated an understanding of mathematics at a level necessary to begin the study of advanced mathematics.

A random sample of eighth grade students in Arizona was selected to participate in the NAEP Assessment. As estimated by the sample, six percent of the eighth grade public school population was classified as Limited English Proficient (LEP), while another seven percent had an individualized education plan (IEP). Schools were permitted to exclude these students from the assessment process.

The results of the Arizona sample reflect the national data provided above. It is interesting to note that within Arizona, 48 percent of the eighth graders are enrolled in an eighth-grade mathematics class, 29 percent are enrolled in pre-algebra, and 18 percent are enrolled in an algebra class.

Within the Tucson Unified School District, Tucson, Arizona, a records review was conducted on 107 students categorized as Learning Disabled who were initially placed into special education during the 1990-91 school year. Eleven percent of these students were found to require academic services only in the area of mathematics and another 18 percent had multiple academic discrepancies which included mathematics (total 29 percent).

To be eligible for placement into special education within Tucson Unified School District as a Learning Disabled student, a 50 percent discrepancy between ability and achievement must exist in the identified academic area. If standard scores are used, a 1 1/2 standard deviation must exist between ability and achievement in the discrepant area. All 29 percent of the students receiving instruction through special education programs for mathematics demonstrated this level of discrepancy on individualized mathematics assessment measures.

Types of Mathematical Problems

Problems experienced by students in mathematics can be dichotomized into problems intrinsic to the students and problems in the environment affecting the student's achievement. This section will examine both, and will review several theories of development (including the medical model as it relates to learning disabilities) and problems effecting mathematics achievement ascribed to environmental factors.

Theories of Development

In this section, three developmental theorists will be discussed: Piaget, Bruner, and Dienes. These theorists were selected for review as their theories discuss the relationship of development and cognitive structures to readiness for skill acquisition in the area of mathematics.

Piaget

Over the past years, considerable attention has been focused on Swiss psychologist, Jean Piaget. His monumental experiments with children, extending over five decades, provide mathematics educators with new insights into how children learn. Since so much of his work has dealt with quantitative concepts of children, his work is most relevant to mathematics.

Piaget asserted throughout his work that cognitive and intellectual changes are the result of a developmental process (Wadsworth, 1984). His basic premise is simply that cognitive development is a coherent process of successive qualitative changes of cognitive structures (schemata), with each structure and its concomitant change deriving logically and inevitably from the preceding one. New schemata do not replace prior ones; they simply incorporate them resulting in a qualitative change.

Piaget conceptualized development as a continuous process with changes in intellectual development being gradual. Schemata (concepts or categories) are constructed and reconstructed (or modified) gradually. For Piaget, development is viewed as a continuum. For purposes of conceptualizing cognitive growth, intellectual development can be divided into four broad stages:

1. Sensori-motor stages (0-2 years). During this stage, behavior is primarily motor. The child does not yet internally represent events and think conceptually.
2. Pre-operational stage (2-7 years). This stage is characterized by the development of language and other forms of representation. It is a time of rapid conceptual development. Reasoning is pre-logical or semi-logical during this period.
3. Concrete Operations stage (7-11 years). During this stage, the child learns to apply logical thought to concrete problems.
4. Formal Operations (11-15 years or older). During these years, the child's cognitive structures reach their greatest level of development, and the child becomes able to supply logical reasoning to all classes of problems (Piaget, 1963).

Piaget perceives development as flowing along in a cumulative manner, each new step in development built upon and becoming integrated with previous steps. Thus, the behaviors described in the four stages are only typical behaviors of a given age, period, or stage.

In a general way, the fact should be emphasized that the behavior patterns characteristic of the different stages do not succeed each other in a linear way (those of a given stage disappearing at the time when those of the following one take form) but in the manner of the layers of a pyramid (upright, or upside down), the new behavior patterns simply being added to the old ones to complete, correct, or combine with them.

Piaget, 1952, p329

The chronological ages during which children can be expected to develop behavior typical of a particular stage are not fixed (Wadsworth, 1984). The age spans suggested by Piaget are normative and denote the times during which an average child can be expected to display the intellectual behaviors characterized by the particular stage.

Piaget's data-gathering procedure included observation and focused conversation with children. Observation used by itself is weak as it is limited by the child's egocentricity, and it is difficult to distinguish the child's play from his beliefs. Focused conversation much like clinical interview allows the child to talk freely while gently being guided toward the target area.

Since Piaget's primary interest is in the child's thought patterns, the research technique must be free enough to make it easy for the child to share them with the examiner. Children's answers were classified into five broad categories:

1. Answer at Random. The child appears uninterested in the question and replies at random, saying the first thing to enter his/her mind.
2. Romancing Answer. The child invents an answer that he really does not believe.
3. Suggested Conviction. The child's answer is designed simply to satisfy the examiner. It is momentary and usually the result of a suggestion by the examiner.
4. Liberated conviction. The child replies after reflection without a suggestive question from the examiner.
5. Spontaneous Conviction. The answer is given immediately because the child already had formulated it since the problem is not new to him.

Even when the solution to a problem is created during the experiment, it implies former intellectual patterns. Children conceive many thoughts they are unable to formulate and verbalize. These thoughts are comprised of a combination of motor schemes and images.

Piaget conceived intelligence as having three components: content, function, and structure (Wadsworth, 1984). Content is what children know about. It refers to observable behaviors - sensori-motor and conceptual - that reflect intellectual activity. Function refers to those characteristics of intellectual activity that are stable and continual throughout cognitive development, e.g. assimilation (integrates) and accommodation (create new concepts or modify existing ones). Structure refers to the inferred organizational properties (schemata or concepts) that explain the occurrence of particular behaviors. For example, if a child is asked to compare a row of 9 checkers to a longer row of 8 checkers and determine which has more checkers, and she responds the row of 8 checkers has more, even though she counts each row, one can infer that she does not have a complete concept of number.

For cognitive development to occur, the child must be active within the environment. The development of cognitive structures is ensured only if the child assimilates and accommodates stimuli in the environment. This can only occur if the child's senses are brought to bear in the environment. When the child is acting in the environment, moving in space, manipulating objects, searching with eyes and ears, or thinking; he/she is taking in the raw ingredients to be assimilated or accommodated. The actions result in the development of concepts.

Actions necessary for cognitive development to occur include more than just physical movement. Actions are behaviors that stimulate the child's thinking

apparatus, and they may or may not be observable. Mental and physical actions in the environment are a necessary but not sufficient condition for cognitive development. The experience alone does not ensure development, but development cannot occur without it. However, also necessary for development are assimilation and accommodation.

For Piaget, all knowledge is a construction resulting from the child's action (Wadsworth, 1979). According to Piaget, three types of knowledge exist: physical knowledge, logical-mathematical knowledge, and social-arbitrary knowledge. Physical knowledge is knowledge of the physical properties of objects and events: size, shape, color, and weight. A child acquires physical knowledge about an object while manipulating it with his/her senses.

Social-arbitrary knowledge is knowledge developed by mankind. It includes knowledge of rules, laws, morals, values, ethics, and language systems. This knowledge evolves within cultures. It is constructed by children through their interactions with other people.

Logical-mathematical knowledge is knowledge derived from thinking about experiences with objects and events (Gallagher & Reed, 1981). Logical-mathematical knowledge can develop only if the child acts on objects. The child creates logical-mathematical knowledge; it is not inherent within objects, as is physical knowledge, but is constructed from the actions of the child on objects. The objects serve as a means of permitting the construction to occur.

Number concepts are examples of logical-mathematical concepts. A child is playing with a set of 12 pennies. She puts them in a row and counts them. There are 12 of them. She puts them in a circle and counts them again. There are still 12 of them. She stacks the pennies, once again, counts them and finds there to be 12 of

them. She finally puts the pennies in a shoebox and shakes them up. Removed from the box and counted, again, there are 12 pennies. Through active experiences like this children construct the concept that the number of objects in a set remains the same regardless of the arrangement of the individual elements. This is a creation of logical-mathematical knowledge. Jerome Bruner (1966), a cognitive psychologist, has provided an alternative way of looking at development.

Bruner

In looking at mathematics instruction, he discovered the children responded to questions with different problem-solving strategies. He proposes three broad categories of strategies: enactive, iconic and symbolic. Using an enactive strategy, the person finds a solution to the problem by acting it out. For example, where to place a couch in a living room can be decided by placing the couch in different positions to see where best it fits.

Before the couch arrives from the furniture store, locating it in a variety of positions within the room can only be accomplished by imagining what it would look like in the different positions. Attempting a solution by forming an image that can be manipulated is referred to as an iconic strategy. Translating the problem into symbols and using rules applied to the symbols to find a solution is known as a symbolic strategy. Using the symbolic strategy to solve the problem with the couch would involve calculations of space and coordination with other furniture items.

The strength of these three strategies and their appropriateness depends on the problem to be solved. Generally, symbolic methods are more powerful than iconic ones, which in turn are more powerful than enactive strategies. Larger numbers of alternative solutions are presented by abstract representation using the symbolic strategy. The developmental trend is to move from enactive attempts at

solutions to iconic and then symbolic using methods which mix the three in differing proportions. It is not necessary to lose the enactive strategy in order to use the symbolic strategy.

Bruner's work suggests development in mathematics as a process of gaining more powerful ways of answering problems. Progress is viewed as the accumulation of ways of thinking that begins with activity, become more abstract by using images which can be changed, and end with the employment of symbols which can be manipulated.

Presented with a problem of sharing 16 sweets equally among four people, a young child may solve this by having 16 sweets and arranging them into four piles. A less young child might solve it without needing to have the sweets as such but by drawing sweets in four columns until there are 16 of them. An older child may adopt a fully symbolic strategy and solve the problem as $16 \div 4 = 4$.

All the children were able to solve the problems presented to them correctly. However, it is not the correctness of the answer, but rather the method of solution that is important. The methods used to solve the problem represent different points along a progression in recognizable development. Learning different and more powerful ways of doing the same thing is part of progress in mathematics. Assessment of progress should include providing children with opportunities to choose how they solve a problem and then observing how they do it as well as if they get it right or wrong. While Bruner discussed three stages of development, Dienes (1971) identified six.

Dienes

Dienes (1971), a cognitive psychologist from England, identified the six stages in the learning of mathematics as follows. In the first stage, in which free

play allows adaptation to occur, the child is encouraged to play freely, but in an environment containing objects which will be used to present and solve problems as a part of later more formal teaching. Thus, the environment is controlled. Playing itself is not advocated but rather playing with things which have important properties influencing the play. For example, there will be objects with properties like different shapes, quantities, or numbers.

The second stage occurs when these properties influence the play by placing constraints on it. Thus, to achieve something, the child must manipulate the objects in certain ways determined by the nature of the objects. For example, an equitable sharing may involve allocating one child 10 small things and another child one large but equivalent thing. The third stage is to recognize that the same sort of restrictions apply in different games and that this is because of an abstract property of the objects - number or size.

In the fourth stage, attention can be focused on the abstraction of a representation found for it. The fifth stage represents stringing together symbols to form a representative language. Finally, the sixth stage involves proving things true or false in the representative language.

Problems Intrinsic to the Student

Cawley (1984) advocates the importance of distinguishing a discrepancy or difficulty from a disorder. A difficulty tends to suggest problems in establishing the rate of performance or growth consistent with others of similar ability. A student may experience difficulty within mathematics. This difficulty may slow the student down or impede progress in developing certain skills or concepts. A disorder on the other hand, implies an aberration or deviance from the norm - something intrinsic to the individual.

Conditions considered intrinsic to the individual and associated with underachievement in students include: mental retardation, sensory impairments (vision and hearing), emotional handicaps, and certain physical and health conditions. Another condition considered by many to be intrinsic to the individual and is particularly relevant to the area of mathematics is that of specific learning disability.

Specific learning disability was defined for educators in Public Law 94-142:

"Specific learning disability" means a disorder in one or more of the basic psychological process involved in understanding or in using language, spoken or written, which may manifest itself in an imperfect ability to listen, think, speak, read, write, spell, or to do mathematical calculations. The term includes such conditions as children who have learning problems which are primarily the result of visual, hearing, or motor handicaps, of mental retardation, or of environmental, cultural, or economic disadvantage.

Education for All Handicapped Children's Act, 1975 (P.L. 94-142)

Some professionals proposed that the definition should state that the disorder is "intrinsic to the individual" (Hammill, Leigh, McNutt, & Larsen, 1981). Professional educator organizations proposed that the definition include either by implication or statement the distinction that an intrinsic psychological or neurological factor has impeded or interfered with the normal development of the child (Kirk & Gallagher, 1986).

Kirk and Chalfant (1984) distinguished two broad categories of learning disabilities: developmental and academic. The major components of developmental learning disabilities are attention, memory, perception, perceptual-motor disorders, and thinking and language disorders. Academic learning disabilities include disabilities in reading, spelling, writing and arithmetic.

Johnson (1979) found deficits in the following areas to be associated with a learning disability in mathematics: memory, visual and auditory discrimination, visual and auditory association, perceptual-motor, spatial awareness and orientation, verbal expression, closure and generalization, and attention. Of course, not all of these were found in each individual identified as having a learning disability in mathematics. Thus, there are some children with discrepancies, for example, in memory who have difficulties in mathematics and there are others who do not. The reverse is also true, there are children with difficulties in mathematics who do not have difficulties with memory.

As Mather and Healey (1990) so aptly noted, some students who experience processing or intracognitive discrepancies may have adequate compensatory strategies to perform successfully within a regular classroom setting. However, other students experiencing the same level of discrepancy may not be able to cope with the regular classroom experience thus possibly requiring special education services. Therefore, the identification of students with learning disabilities requires behavioral observations and experienced professionals to determine that the child is not doing as well as he/she could within the classroom. Such judgments should not be based totally on objective testing resulting in a quantification of abilities. Rather, careful observation and description of performance and presenting symptoms should be significant pieces of information in the process of documenting the disorder and need for special education services.

Problems Extrinsic to the Student

All students who are underachieving or who are performing below grade placement do not have intrinsic conditions. Many factors contribute to underachievement or poor performance. Extrinsic or environmental factors that can lead to poor performance include: lack of opportunity to learn, cultural disadvantage, economic disadvantage, and inadequate instruction.

Error patterns represent another way of looking at extrinsic factors involved in poor performance in mathematics. Englehardt (1977) conducted a study in which 198 third- and sixth grade students were administered an 84-item arithmetic computation test. Those items identified as having incorrect responses were analyzed to infer probable student approaches or misconceptions leading to the responses. According to their commonalities, these inferences were clustered to form error types. This procedure resulted in the identification of eight types of errors. These error types included:

1. Basic fact error. The student responds with a computation involving an error in recalling basic number facts.
 2. Defective algorithm. The student responds by executing a systematic (but erroneous) procedure.
 3. Grouping error. The student's computation is characterized by a lack of attention to the positional nature of our number system.
 4. Inappropriate inversion. The student responds with a computation involving the reversal of some critical aspects of the solution procedure.
 5. Incorrect operation. The student performs an operation other than the appropriate one.
 6. Incomplete algorithm. The student initiates the appropriate computational procedure, but aborts it or omits critical steps.
 7. Identity errors. The student computes problems containing zeros and ones in ways suggesting operational confusion.
 8. Zero errors. The student computes problems containing zeros in ways suggesting difficulty with the concept of zero.
-

The students in this sample committed 2,279 errors. On the average, students attempted 68.7 of the 84 possible computational items and responded correctly to 57.2 while committing an average of 11.5 errors. Interestingly, examination of the distribution of errors revealed that factors of grade level, sex, ethnicity nor urban/rural school district membership suggested the type of errors students would commit.

West (1971) reported in a qualitative study using two students that an examination of their mistakes revealed two types of errors: careless and conceptual. The careless errors were identified by their inconsistent appearance in the students' daily work. They often occurred as a result of the student losing track of where he/she is in the procedure.

Conceptual errors occurred regardless of the number of times the student approached the problem. The student consistently made the error leading to the belief that the child had an incorrect understanding of the procedure. Sometimes, the student made the error only under certain circumstances indicating that he has some grasp of the procedure but not a complete understanding.

Macht (1990) identified three major types of errors: mechanical, conceptual, and motivational. Mechanical errors result from the lack of a minor subskill necessary to complete a task and have little to do with cognitive operations involved in understanding. To rectify the error, the student must learn the specific subskill, often through modeling and practice.

Conceptual errors involve understanding concepts, operations, the purpose of the concepts, and how concepts interact with each other. Sometimes students can perform an operation correctly in isolation or from example, but in another context are unable to apply the concept necessary to identify the operation they know well.

Motivational errors occur when students see no reason to put forth effort. Often the tasks requested hold no personal value for the student.

Approaches Related to the Assessment of Mathematical Performance

Understanding where assessment practices are today requires reviewing a brief history of testing in the United States that emphasizes mathematics and language arts.

The introduction of compulsory education in the United States around the turn of the 20th century led to the development of psychometric tests (Ridgway, 1988). These tests were seen as providing an alternative to subjective forms of student evaluation. The initial impetus for formalized testing generated from concerns that individuals should receive an appropriate education based on their abilities rather than social backgrounds. As a result, the early focus of testing was on the identification of students with special needs. It emphasized different levels of performance as opposed to individual skill analysis for purposes of remediation. Publishers sought to develop tests which accentuated student differences through scores.

The growth of standardized testing for educational purposes was prolific following World War I (Frederiksen, 1984). A 1925 survey of school superintendents in 215 cities of over 10,000 population indicated that classification of students into homogenous groups was a primary use of intelligence tests

(Chapman, 1980). In a follow-up survey during 1926, the same population of superintendents indicated that homogenous ability grouping was employed in 90% of the elementary schools in cities with a population greater than 100,000. By 1930, yearly sales of group intelligence tests reached over 750,000 and those of achievement tests reached 1.5 million (Chapin, 1980).

Standardized testing was institutionalized into civilian life by the outbreak of World War II (Haney, 1984). Approximately 10 million recruits were tested for military assignments using the Army General Classification Test (AGCT). The AGCT was comprised of three types of problems: vocabulary to measure the verbal factor, arithmetic word problems to measure the number and reasoning factors, and block counting to measure the space factor. (Darley, 1953).

During 1965, the U.S. Congress passed a bill that greatly influenced school testing programs. This was the Elementary and Secondary Education Act (ESEA) of 1965. The majority of funds disbursed under ESEA were Title I funds, providing financial assistance to local education agencies for education of children from low-income families. With the federal funding came strings in the form of a provision for the evaluation of the effectiveness of Title I funded programs. Between 1966 and 1981, Title I funding grew from approximately three-quarters of a billion dollars to nearly 4 billion dollars (Grant & Eiden, 1982).

Specifically, the ESEA legislation required that local education agencies (LEAs) could receive funds only if "...effective procedures, including provision for appropriate objective measurements of educational achievements, will be adopted for evaluating at least annually the effectiveness of the programs in meeting the needs of educationally deprived students" (Public Law, 89-10, 89th Congress, H.R. 2362, April 11, 1965, as quoted in Barclay and Mosher, 1968). By 1974, more than

90% of local Title I evaluations consisted of standardized forms of achievement testing (Reisner, Alhen, Boruch, Linn, & Millman, 1982).

Similarly, the Education for All Handicapped Children's Act of 1975 has served as a foothold for the proliferation of and dependence upon standardized diagnostic instruments for evaluation of students for special education programs (Galagan, 1985). Although EHA clearly states that the prerequisite for an evaluation to determine eligibility for placement into special education programs is a suspicion that the student is disabled, Ysseldyke (1983) found that the "suspicion of disability" standard was ignored or subordinated to subjective and chaotic referral methods by individual teachers. Thus, if a teacher referred a student for special education, Ysseldyke (1983) found that a 78% chance of placement existed seemingly as a result of simply being referred.

Mirkin (1980) analyzed the cost of the assessment process for a student referred for special education services. The average referral required 13 to 15 hours of professional time, and the potential cost could be as high as \$1800 for each student.

Currently, it is estimated that 30 million school-age students are administered standardized tests each year (Frederiksen, 1984). The National Center for Fair and Open Testing estimates that each school-aged student living in the United States takes as many as three standardized tests each year (Neill & Medina, 1989). These tests can determine which classes students will take, which schools they will attend, and even which level of academic potential they are expected to achieve. They can also influence decisions about institutional goals, teacher performance, and program funding (Gifford, 1990).

Standardized Achievement Tests

Standardized achievement tests are not the focus of this study. However, in order to understand why methods like curriculum-based assessment are necessary, it is important to understand the general nature of standardized tests.

Definition and Purpose

A standardized test is one in which the procedures for test administration, test content, equipment, and resources to be used, together with a scoring scheme, have been set out in such a way that they are easily replicated by different testers and examiners (Ridgway, 1988). Standardization refers to the process of establishing a baseline of performance on a test, against which student performances can be judged. Thus, for the purpose of this paper, the term standardized test will refer to a test prepared for standardized administration, for which test score norms have been prepared.

Standardized achievement tests are characterized as multiple-choice, norm-referenced, and teacher proof. They are commonly administered in a single test setting with pre-determined time frames on an annual basis. They purport to sample a student's knowledge and recollection of basic facts within a specified subject area.

By design, standardized achievement tests are not intended to represent a curriculum, but rather to sample skills that are included in a variety of common curriculums (Haney, 1991). Within this context, standardized achievement tests are typically given at either end of the learning process, they are given before the student begins learning in order to determine what they don't know, or after they have finished learning, to determine what or how much they have learned (Fessoules & Gardner, 1991).

Standardized tests are also used in the identification process to determine eligibility of students for special education programming. A discrepancy model is frequently employed consisting of a statistical formula for determination of eligibility, particularly in the identification of students for the learning disabilities category.

Criticism has been leveled against standardized tests because they are biased regarding curricular content (Armbruster, Stevens, & Risenshine, 1977; Jenkins & Pany, 1978), technically inadequate for making decisions about individual students (Salvia & Ysseldyke, 1985), and are not useful for making instructional decisions (Salmon-Cox, 1981). For example, while intelligence tests are supposed to diagnose the handicap of a student through the use of standardized procedures and normative comparisons, and achievement tests provide data comparing the student's academic performance with that of a normative reference group, neither can determine a student's appropriate placement or progress within the curriculum, nor are the results relevant to developing academic interventions (Deno, 1985).

Within the norm-referenced model, the learning process is conceptualized as a function of underlying psychological and neurological processes. Therefore, the causes of academic problems are viewed as either faulty information processing or poor skills within the cognitive domain. Either way, academic performance is seen as a skill residing within the learner.

There are problems related to the use of standardized measures for the identification of students for special education programs. First, according to guidelines established by the American Psychological Association, the American Educational Research Association, and the National Council of Measurement in Education (1976), all tests used in education and psychology must be valid, reliable

and have adequate normative data. In a review of assessment procedures and tests used in special education, Salvia and Ysseldyke (1981) listed 26 frequently used tests that were not technically adequate.

A major issue in test selection is the purpose of testing and the way in which testing is to be used. Ridgway (1988) defines the purposes of standardized norm-referenced testing as follows:

1. Comparison with a standardizing sample, for purposes of evaluating school performance with respect to a national and/or state standard.
2. Making comparisons between schools, or between different educational treatments.
3. Comparing the effectiveness of different teachers by examining relative student gains after each year's teaching (or within school years using parallel classes).
4. Rank ordering students to identify students who are low-achievers or high-achievers for purposes of grouping or altering curricula.
5. Following student progress over time.

Ridgway (1988) continued by delineating the purposes of standardized diagnostic tests:

1. Specifying what has, and has not been learned.
2. Identifying particular strengths and weaknesses of individual students.
3. Identifying specific student misconceptions which underlie particular errors.
4. Describing hierarchies of conceptual development which have implications for teaching.
5. Use as the basis for discussions with students, parents, and colleagues.

No single standardized achievement test represents a complete picture of the academic content of a particular school, nor was it intended to by its publishers. As definitions and purposes of tests and testing are defined, it is interesting to look at actual uses and outcomes of these instruments.

Uses and Outcomes of Standardized Tests

In the United States, the standardized achievement test score is accepted by the public and many educators as a valid measure of educational accomplishments (Haetel & Calfee, 1983). In a review of testing practices related to standardized achievement testing, Haladyna, Haas, and Nolen (1989) listed many possible uses of standardized tests. These uses range from policy analysis at the national level to parental review of their child's achievement. For example, test scores are used to rank states by the U.S. Department of Education in its annual "report card", and by legislators and other government officials to assess educational effectiveness of states and school districts. School boards and school district personnel use test scores to determine the effectiveness of their districts and schools within each district. Newspapers report test scores of schools within their distribution areas. Test scores are used by some school district personnel to determine merit pay and to make other personnel decisions. Real estate agents use test scores to rate the quality of neighborhood schools.

Prior to the recent onset of the "age of accountability" and need to evaluate schools at all levels, test scores were used for a rather limited set of purposes (Haladyna, Nolen, & Haas, 1991). Scores were used to group students for instruction, evaluate and modify school district curricula, plan instruction, diagnose achievement deficits, place students into special programs and help parents

understand the general achievement levels of their children. Scores continue to be used for these purposes.

As standardized tests are the key to the lock of special education eligibility, it is helpful to examine the usefulness of these tests. Thurlow and Ysseldyke (1982) examined this question by interviewing 200 school psychologists and teachers. Given a list of commonly administered tests, including the Wechsler Intelligence Scale for Children - Revised (WISC-R), Bender Visual Motor Gestalt (Bender), and the Wide Range Achievement Test (WRAT), 100 psychologists typically rated the measures as useful for instructional planning: 72% reported the WISC-R was helpful, 64% endorsed the Bender, and 80% thought the WRAT assisted in program planning. When the 100 special education resource teachers were asked to evaluate the same tests, a different picture emerged. Only 10% thought the Bender was instructionally relevant, 30% rated the WISC-R as useful during the IEP planning process, and 10% said the WRAT contributed to developing effective instructional plans. It was concluded that teachers are searching for assessment alternatives not provided through traditional forms.

Within education generally, the increased use of standardized achievement tests has brought pressure to raise scores which has led to a variety of outcomes. In a qualitative study using data from teacher and staff interviews and classroom observations on the role of testing in elementary schools, it was concluded that testing programs substantially reduce the time available for instruction, narrow curricular offerings, and potentially reduce the alternatives available to teachers for selecting content, methods, and materials that are incompatible with standardized testing formats (Smith, 1991). An example of this phenomenon was observed during the study where teachers gave up their "Math Their Way" curriculum which is highly

manipulative and hands-on because they felt their students could not make the transition to paper-and-pencil computation quickly enough to prepare for the Iowa Test of Basic Skills.

A national survey recently indicated that 75 percent of teachers report that their school district's curricula has changed "somewhat" or "very much" to match standardized achievement tests used in their districts. Forty-nine percent of the teachers reported that curricula have been changed to match particular questions on tests (Haney, 1991).

Within the context of outcomes unintended by purpose, Dyer (1977) cautions the test-using public concerning the limitations, misconceptions, and consequent misuses and abuses of standardized tests. He cautions:

1. Test norms should never be confused with fixed standards of academic performance.
2. No battery of standardized achievement tests can measure everything one cares about in student learning.
3. Intelligence tests would better be regarded as a general achievement test that is inevitably affected by the student's learning opportunities.
4. The measures gained from educational and psychological tests are inevitably affected by errors of measurement.
5. A machine-scored multiple-choice test cannot pretend to provide anything more than an indirect measure of what children know or how they think.

In summary, a standardized achievement test is a snapshot of a student's ability to recall facts or procedures. By itself, the mathematics section gives only a limited picture of a student's mathematical knowledge - usually focusing on tools,

techniques and memory for facts rather than mathematical thinking, understanding, and skill and communicating (Mathematics Framework for California Public Schools K-12, Version of 20 August 1991).

Standardized testing provides limited information regarding specific student skill performance, while authentic testing looks more specifically at student responses within the natural environment of the classroom on a daily basis.

Authentic Assessment

Definition and Purpose

Authentic assessment is different from methods previously reviewed. It consists of evaluating by asking for the behavior you want to produce. It looks at what students really do and know within the context of their own learning environment. The focus is on student work and action (extrinsic factors) rather than factors intrinsic to the students (Mathematics Framework for California Public Schools K-12, Version of 20 August 1991). Within this work, characteristics of authentic assessment are described as follows:

1. Assessment occurs within the course of normal work. The class does not stop learning to take a test. The work and the assessment are inseparable.
2. The conditions for assessment mirror the conditions for doing mathematics outside of school: students have ample time, they have access to peers, and to tools (books, notes, calculators, pattern blocks). They also have the opportunity to revise their work.
3. The tasks for assessment engage a student's sense of purpose. The selected tasks are multidimensional allowing students to demonstrate thinking, understanding, and communication skills.

4. Feedback from assessment is concrete and specific to the tasks.
5. Students participate in the process of assessment. They assist in creating and applying standards for quality work.

Generally, the purpose for authentic assessment is to improve learning. Stenmark (1989) described four lesser purposes of assessment: keeping track - what are students doing; checking-up - what are students learning; finding out - what's going on; and summing up - what happened (accountability). Stenmark (1989) described areas to be assessed in mathematics under this model:

1. The students' use of mathematics to make sense of complex situations.
2. Students' work on extended investigations.
3. The ability of students to:
 - formulate and refine hypothesis
 - collect and organize information
 - explain a concept orally or in writing
 - work with poorly defined problems or problems with more than one answer similar to those in real life
4. Students' use of mathematical processes, such as computation, in the context of a variety of problems rather than in isolation.
5. The extent of the students' understanding of mathematical concepts.
6. Students ability to define and formulate problems.
7. Whether students' question possible solutions, looking at all possibilities.
8. How a students productive work changes over time.

Authentic assessment is not one single method of assessment (Mitchell, 1989). It includes performance tests, demonstration of student's ability to use

learned skills: observations, open-ended questions with no single answer; exhibitions, students select the products to demonstrate learning; interviews, providing an opportunity for students to reflect on their achievement; and portfolios, collections of student work. The list is limited only by the criterion of authenticity. Schools ask, 'is this what we want students to know and be able to do?'

Uses and Outcomes of Authentic Assessment

In the world of work, people are valued for the tasks or projects they do, their ability to work with others, and their responses to problem situations (Stenmark, 1989). To prepare students for success, the first use of assessment is within the classroom, to provide information for making instructional decisions that will bridge the classroom to the real world.

It has been asserted that if authentic assessment practices were implemented to promote learning, teachers should see students using mathematics with facility to communicate their own thinking about complex situations through pictures, diagrams, graphs, words, symbols or numerical examples (Stenmark, 1989). Students should solve problems using a variety of mathematical tools and models, such as manipulatives, calculators, and computers. Students would plan, invent, design, and evaluate their own mathematical ideas and products. They would be found doing projects and activities. Students would be found working together developing group problem-solving skills. They may even be enjoying mathematics.

If students are to be asked to complete analogies, recall information, or perform basic calculations under time constraints, then schools need curricula in which students have regular opportunities to form analogies, memorize facts, and manipulate equations with speed (Fessoules & Gardner, 1991). By the same token if students are going to be asked to grasp scientific principles, compose a melody, or

write compelling dialogues, then schools need curricula that provide students opportunities to investigate, test, and observe nature; to compose and experiment with melodies; and to craft, rehearse, and revise many scenes many times.

Just as standardized testing has driven curriculum and instruction in schools, so too will the implementation of new measures influence and shape activities in the classroom. For example, currently in many instances teachers teach reading and mathematics using worksheets and practice materials that closely resemble testing materials (Kirst, 1991). Thus, the testing formats and content shape the daily mode of instruction, leading to repeated drill on isolated skills. The tasks and problems used in authentic assessment are complex, integrated, and challenging instructional tasks (Kirst, 1991). They require students to think in order to be able to arrive at answers or explanations. Thus, authentic assessment mirrors instruction, which engages students in thinking from the very beginning. Thus, pedagogical approaches will need to change from current practices to meet the demands of authentic assessment.

Curriculum-Based Assessment

As described by Gickling, Shane, and Croskery (1989), curriculum-based assessment (CBA) is a system for determining the instructional needs of a student based on the student's ongoing performance in existing course content and for delivering instruction as effectively and efficiently as possible to meet those needs. Using this definition, there are three key features of CBA: test stimulus is drawn from the student's curriculum, testing occurs repeatedly across time, and the assessment information is used to formulate instructional decisions (Tucker, 1985). Thus, CBA assists the teacher in determining how the student is performing in the required coursework, whether or not the performance is at the instructional level of

the curriculum in progress, what the instructional level of the student is, and whether or not performance improves with appropriate instructional intervention (Tucker, 1991).

A central premise of the CBA model is that it is necessary for assessment, curriculum, and instruction to be integrated and approached as a unit to promote optimal learning conditions for students and successful instructional environments for teachers (Gickling, 1992). Within this model, data collection, interpretation, and application are interwoven functions whose primary goal is to facilitate the instructional decision-making process. Thus, CBA focuses on each student's entry skills relative to his or her course work, the instructional demands made of each student by various course assignments, and controlling the degree of task difficulty by adapting or modifying the assigned tasks to match the student's skills (Gickling & Thompson, 1985):

1. Curriculum provides the most basic and meaningful form of educational assessment.
2. Curriculum places explicit demands upon the learner.
3. Curriculum must be controlled if academic success is to be achieved by the student.

Most students experience occasional difficulty within a curriculum. Most students share an occasional experience of not understanding what is expected of them, of receiving little or no help, or being unable to do the assigned work (Gickling, 1992). For most students, these negative experiences are offset by sufficient positive ones allowing them to successfully continue within the curriculum. However, there are a substantial number of low-achieving students whose overall school experiences are unsatisfying.

Resulting from the way that schools are organized, these students are expected to proceed up the curricular ladder at the same rate as their peers and to make the same amount of progress given the same amount of instruction and practice. Unable to adapt to instruction and a curriculum that moves too fast and demands too much in relation to their existing skills, they fall further behind and become entrenched in a failure cycle.

These are the students who become academically forgotten, the "curriculum casualties" of school systems (Gickling & Havertape, 1981). Their one basic problem is that their readiness levels and learning rates are not synchronized with the instructional entry skill requirements and within the curriculum making up grade level programs (Hargis, 1982). Thus, an underlying assumption of CBA is that students are basically controlled by the curriculum and that the majority of school related problems are curricula induced rather than student-based.

Uses and Outcomes of Curriculum-based Assessment

Current views of assessment characterize it as a process of data collection for decision-making. Salvia and Ysseldyke (1988) break special educational decision making into five areas: screening, eligibility or identification, instructional planning, student progress, and program evaluation. Instructional planning can be broken down into determining the instructional level of the curriculum and the content of instruction.

Currently, the primary use of CBA is in the instructional planning component of this decision-making process. The central purpose of CBA is to facilitate the delivery of effective and efficient instruction by ensuring that students are placed properly in instructional materials. A basic tenet is that optional learning conditions are those in which instructional tasks contain an appropriate margin of challenge

but are linked to the entry-level skills of the learner to assure the student a high degree of success (Shenn, Knutsen, & Rosenfield, 1989).

Gickling and Havertape (1981) maintained that an instructional match is created when the ratio of known to unknown items is controlled sufficiently to assure a high level of student success. Known items are defined as task items to which the student provides a correct and immediate response. If a student's response to an item is incorrect or hesitant, the item is considered unknown or challenging. The ratios of known to unknown items translate into decisions about the students' instructional, independent, or frustration level (Gickling & Thompson, 1992). The established decision rules for instructional placement in mathematics require that students achieve at least a 70% to 85% correct response rate on drill material and at least an 85% to 100% level in application of mathematical skills in order for the student to be considered "matched" for instructional purposes.

A study conducted by Gickling, Shane, and Croskery (1989) investigated the effects of an intervention incorporating the CBA model on the performance of high school low achievers in general mathematics classes. Subjects were identified as being at-risk for academic failure, placed in matched pairs, and assigned randomly to treatment and control groups.

Treatment group subjects received instructional materials within their regular classrooms that were matched to their instructional levels, developed from on-going assessment of their performance. The control subjects continued to receive regular classroom assignments.

Data analysis was limited to the 17 of 30 subjects completing the study. Independent tests were used to determine changes in scores on mathematics sections of the Stanford Test of Academic Skills and the Nevada State High School

Proficiency Examination (pre-post). Both within and between group comparisons were performed. Within group results showed that both treatment and control subjects made significant gains. However, no significant differences were noted in comparisons between groups.

Multiple regression analysis was used to test the effects of CBA in teacher grades. CBA effects were found to be significantly associated at the .01 level with increased teacher grade scores. A second regression equation tested the effects of CBA on classroom task-related behaviors. Using mean observation scores of on-task, task completion, and comprehension as the criterion variables and the composite mean score plus the presence or absence of CBA as the predictor variables, a multiple regression analysis was conducted. The results were significant. The researchers recommended that the study be replicated, and other studies be conducted.

Research evidence supporting the use of CBA is extremely limited. However, the use of CBA is being viewed as a reasonable procedural process in special education decision-making at the pre-referral level. Galagan (1985) proposed that CBA methods could provide effective screening data on students at the pre-referral stage thus reducing unnecessary referrals for evaluation for special education programming. The use of CBA methods provides systematic observational data on student performance, the degree of performance deviance, and the formulation of alternative instructional methods prior to referral.

The Pennsylvania Department of Education, Bureau of Special Education (1991) issued a paper entitled "Instructional Support". This paper calls for implementation of an instructional support process composed of assessment and intervention procedures used to assure that students receive an effective

instructional program that will meet their learning needs prior to a referral for evaluation to determine eligibility for special education programming. The instructional support process serves as a mechanism for screening students who may be eligible for special education by identifying those students who do not make sufficient progress even after intensive interventions are applied in the regular program.

One component of the Pennsylvania instructional support process is CBA. It has been identified as a best practice in student assessment because it provides classroom teachers and other educators with meaningful and precise information in evaluating students' academic needs.

CHAPTER III

METHODS

Research Design

The present research used a multiple baseline single subject time series design with staggered baselines to study changes in the dependent variable (Barlow & Hersen, 1984). Multiple baselines can be collected across behaviors, conditions or subjects (Tawney & Gast, 1984). In this study, the multiple baseline occurred across subjects.

Staggered baselines (3-5 baseline probes) were used to illuminate internal validity. Baseline data were collected on each subject, however, the number of baseline probes was random across subjects and skills. This assured the researcher that the implementation of treatment was the reason for change in skill acquisition as opposed to other unknown factors.

This study was conducted in three phases (A, B, C). The first phase, Phase A involved a series of baseline probes of the target behavior (percent correct of specified mathematics problems) under study. In Phase B, the independent variable (CBA assessment and instruction) was introduced, and changes in the dependent measure were noted. In Phase C, maintenance of any changes occurring in Phase B were observed through a series of three probes administered 14-18 days after treatment had been discontinued.

Graphic displays assist in organizing data and provide a detailed numerical summary and description of behavior (Tawney & Gast, 1984). Independent analysis of the relationship between variables is a strength of single subject research. The

graph as a reporting format provides an efficient visual summary of data facilitating reliable analysis of data trends both intra- and inter-subjects.

For the purpose of this study, line graphs are used to chart the dependent variable across the three phases (A, B, & C) of the project. The dependent variable, growth-over-time, is plotted on the ordinate line, while the trials and sessions for each phase are plotted on the abscissa. Data paths connect data points within each phase. Phase changes are represented by bold solid vertical lines.

The data, can be analyzed through visual inspection of the dependent variable as it is recorded on growth-over-time line graphs as shown in the examples below:

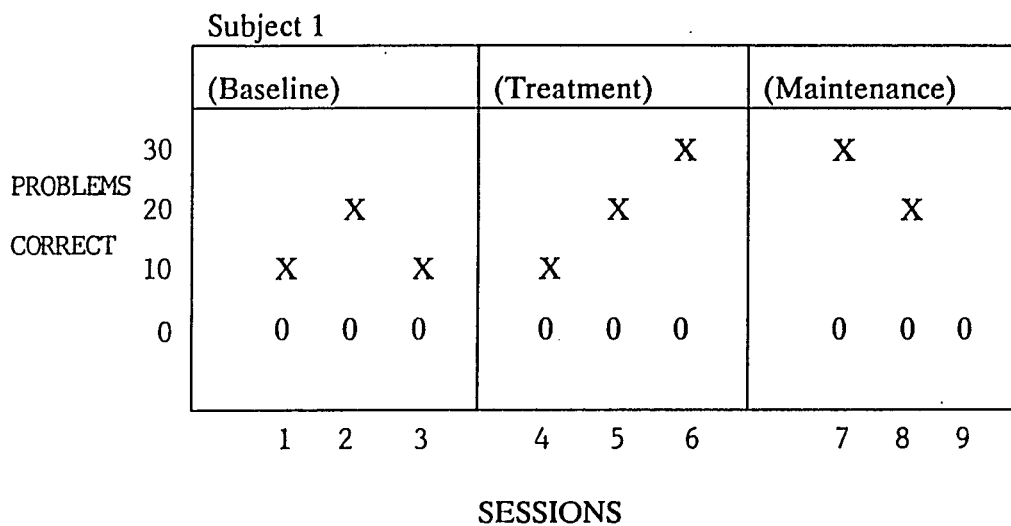


Figure 1. Subject 1 - The effects of Treatment on skill identified

Variables

The dependent variable has been defined as a variable that changes as a function of a change in another variable (Tawney & Gast, 1984). For the present

study, the dependent variable is growth-over-time with respect to the subject area of mathematics with students at the middle school level.

The independent variable is a variable that is manipulated or observed to determine its effect on the dependent variable. Tawney and Gast (1984) describe it as, "... equivalent to the intervention, treatment or experimental conditions under which a behavior is repeatedly measured and evaluated." For the present study, the independent variable is the CBA model which consists of several components: present performance levels (PPLs), repeated probes, error analysis and the development and implementation of instructional strategies within the subject area of mathematics.

Subjects

A total of nine students from one of 19 middle schools in the Tucson Unified School District located in Tucson, Arizona participated in this study. Several factors were important in determining which school was selected. First, the school needed a pool of students who experience failure in math. Secondly, the administrative, support, and teaching staff had to be willing to cooperate with the study. Finally, the student population selected was ethnically diverse. This permits even more generalizability of the results.

All teachers of mathematics within the school were asked to refer students from within their classrooms who had failed two quarters of math. Subjects were drawn from regular education mathematics classes at the seventh and eighth grade levels. They received no special instructional services within the school. Special education students were not selected for this study as the CBA model can be implemented as an effective intervention strategy for students within regular

education thus reducing the need for referral for evaluation for special education programming.

Students were referred by their teachers to participate in the study. From this referral pool of students, the school counselor selected 12 subjects for the study. From the pool of 12 subjects, nine subjects were selected by random drawing to begin the study. From those nine, five subjects were selected through random drawing for placement into Group A. The remaining four subjects comprised Group B. The three subjects who were not selected to begin the study formed the pool to enter the project should one or more of the selected subjects be unable to participate.

The subjects selected for this study and the skills they worked with are described as follows:

1. Subject 1 was a 12 year, 1 month old Hispanic male in the seventh grade. The skill identified to be worked with for this subject was the addition of unlike fractions.
2. Subject 2 was a 14 year, 11month old Hispanic female in the eighth grade. The skills identified for her include: subtraction with regrouping, addition of like fractions, and reducing fractions.
3. Subject 3 was a 14 year, 2 month old Hispanic male in the seventh grade. The skill identified for him was a simple division.
4. Subject 4 is a 13 year, 4 month old Hispanic male in the seventh grade. The skills identified for him include: addition of unlike fractions, changing fractions to percents, and multiplication of fractions.

5. Subject 5 was a 12 year, 9 month old white female in the seventh grade. The skills identified for her include: division with zero and addition of unlike fractions.
6. Subject 6 was a 12 year, 7 month old white male in the seventh grade. The skills identified for him include: addition of unlike fractions, changing fractions to percent, and multiplication of fractions.
7. Subject 7 was a 12 year, 8 month old Hispanic male in the seventh grade. Skills identified for him include: addition of like fractions, addition of unlike fractions, and reducing fractions.
8. Subject 8 was a 14 year, 1 month old Native American female in the eighth grade. Skills identified for her include: changing fractions to percents, multiplication of fractions, and division of fractions.
9. Subject 9 was a 14 year, 8 month old black male in the eighth grade. The skills identified for him were changing fractions to percents and multiplication of mixed numbers.

Instrumentation

The Wide Range Achievement Test - Revised (WRAT-R) was selected as a screening instrument for performance in arithmetic by the subjects. The general purpose of the WRAT-R is to measure the codes needed to learn the basic skills of reading, spelling, and arithmetic (Conoley & Kramer, 1989). The standardization sample included 5,600 persons stratified by age (28 age groups, each with 200 persons), sex, race (white/non-white), geographic region, and community residence (metropolitan/non-metropolitan).

The test exhibits internal consistency, has a developmental progression of raw scores, offers subtests with moderately high correlations with Woodcock-

Johnson achievement subtests, and the format of administration is similar to classroom activities (Conoley & Kramer, 1989). The instructions for administering and scoring individual subtests are clearly presented, as are the procedures for converting raw scores to standard scores and grade equivalents. The resultant standard scores have a mean of 100 and standard deviation of 15.

Test users are cautioned that more information is needed on reliability (the test-retest sample was small; 81 persons on Level 1 and 67 on Level 2). Validity is questionable as it is unclear what the test is measuring. So without further information in these areas, test users should be cautious when using the test to assist in the diagnosis of learning disabilities, help determine personality structure, or check school achievement for vocational assessment, job placement or testing.

However, for the purposes of this study, these concerns are not relevant as scores from the WRAT-R were neither used nor reported. The sole purpose for using the WRAT-R in this study was to screen for the subject's present performance level in arithmetic operations. At the point where subjects are determined to be unsuccessful, informal probes developed by the Experimenter were used to analyze their errors. Thus, the WRAT-R served only as a screening device.

Treatment Groups

Subjects were divided into two groups. Group A (Experimental Group) was comprised of five subjects; while Group B (control group) was comprised of four subjects. Each subject received six treatment sessions in math skill areas specific to their needs as determined by CBA during a two week period. While Group A received treatment, Group B received regular classroom instruction.

Group B received internal probes during the treatment phase of Group A. The internal probes consisted of a total of two probes for each subject over the two

week treatment period. The internal probes were conducted to assure that external factors to the study did not account for gains in identified baseline skills of subjects. Group B received the experimental treatment following Group A's treatment period for the purpose of replicating the experiment.

Experimental Treatment

This project involves three phases in the experimental process: baseline, treatment and maintenance. This section delineates the steps used in each phase with both groups (A & B).

Group A

Baseline (Phase A)

1. Each subject was called individually from class to the office to meet the experimenter.
2. The school counselor introduced the experimenter to the subject by saying, "This is Mrs. Bornfield, she is here to do some math activities with you to see if some special instruction will help you. She will meet with you often over the next few weeks."
3. The Experimenter took the subject into a small room where they sat across a cleared table from each other. The experimenter had the WRAT-R with a math problem sheet located next to her on the table. The WRAT-R had been selected as a screening instrument for performance in arithmetic by the subjects as it reflects the sequence of the district implemented mathematics curriculum. Scores from the WRAT-R were not reported in this study.

4. The Experimenter gave the following instructions, "I will give you a sheet with a number of math problems on it. I want you to do the best you can. Some of the problems will be easy for you and some will be hard. Again, do the best you can, remembering that it is okay to make mistakes. Do you have any questions? (Experimenter responded to questions). Please begin."
5. The subject had been given a Wide Range Achievement Test - Revised (WRAT-R). During the administration of this measure, the Experimenter took notes on the subject's performance as a part of the information gathering process to assist in conducting the error analysis. The notes described from the Experimenter's perspective the process used by the subject to solve the math problems.
6. At the completion of this task, the subject was asked to return the protocol to the Experimenter. Upon reviewing the protocol and notes, Experimenter interviewed the subject regarding the method he/she used to solve particular problems in order to evaluate the PPL and gather information to conduct the error analysis. Student responses were recorded. Experimenter said, "Tell me what you did to get this answer?"
7. Upon completion of the interview, a specific skill selected through error analysis of a series of PPLs and a 10 problem set was developed by the Experimenter to give to the subject. In developing the 10 problem set, the Experimenter considered the complexity of the identified skill. For example, in learning the addition of unlike fractions, subjects must be able to initially identify a denominator in the presented problem that may serve as the common denominator, as in $1/10 + 2/5 = \underline{\quad}$. As it

becomes more complex, the subjects may need to multiply the denominators to determine the common denominator as in $1/2 + 1/3 = \underline{\quad}$. This problem set served as the first data point for the baseline phase.

8. To establish a baseline, the subject was asked to complete the probe described as the 10 problem set. Experimenter said, "Now, I would like you to do these problems for me. Once again, do the best you can and don't worry about making mistakes. Do you have any questions? (Experimenter responded to questions). Please begin."
9. As the subject finished and returned the problems to Experimenter, an interview regarding the methods used by the subject to solve the problems was conducted. Experimenter stated, "Please share with me the method you used to solve this problem." If this proved unsuccessful in soliciting a response, Experimenter asked, "What did you do first?" If no response, Experimenter asked, "What is the number?". This interview soliciting the method used to solve the math problems by the student was important in establishing the PPL.
10. At the conclusion of the interview, the subject was given a 5 minute break.
11. At the conclusion of the break, the subject was asked to complete 10 similar problems using the same procedure as above. This was replicated 3-5 times to establish a baseline.
12. At the conclusion of the last set of problems, the subject was returned to class.

Treatment Phase (Phase B)

1. Group A subjects began Phase B within 3 days of baseline completion for a total of 6 sessions of treatment.
2. Experimenter had performed an error analysis, selected skills to be taught, and developed an instructional strategy prior to meeting with the subject for the first treatment session. Selected math teaching methods proved to work given that instruction was linked directly to the error analysis.
3. For treatment sessions, the subject was called from class and directed to meet Experimenter in the small resource room. Sessions lasted approximately 20 minutes.
4. Experimenter greeted subject by name upon arrival and interacted briefly in a chit-chat manner on subjects such as weather, sports, or special activities.
5. Seated at the table, Experimenter discussed a math subskill, identified through error analysis as a problem for the student, with the subject and implemented the instructional strategy.
6. At the conclusion of the intervention, a PPL was given, consisting of 10 math problems in the area of identified skill deficit. Experimenter said, "Please complete these problems for me. Do you have any questions: (Experimenter responded to questions). Please begin."
7. When the subject completed the task, Experimenter thanked him/her and sends him/her back to class.
8. The above procedure was repeated for all subjects(with one exception to be explained) in Group A over 6 sessions.

9. At the completion of the 6th session, the subjects were thanked for their cooperation and told that they would meet one more time following a 14-18 day break period. This additional session was held for the purpose of determining if maintenance of performance level after treatment of the dependent variable occurred. Also, treatment results were shared with each subject regarding their own performance at this final session.

Maintenance Phase (Phase C)

1. Between 14 and 18 days after the termination of the treatment phase, each subject was called individually from class and asked to meet with Experimenter in the small resource room.
2. Experimenter welcomed the subject by name and engaged in idle chit-chat regarding weather or sports for a few moments re-establishing rapport.
3. Experimenter said, "I want to see how well you're doing in math now. So, I'm going to give you a series of 3 sets of 10 problems each to perform, with a 5 minute break between sets. Once again, I want you to do the very best that you can. Do you have any questions? (Experimenter responded to questions). Please begin."
4. At the completion of the 3 sets, the subject was thanked and sent back to the classroom.

Group B

1. Group B subjects participated in Baseline activities with Group A. (See "Baseline (Phase A)" under the heading of "Group A" for delineation of procedures).

2. At the conclusion of the baseline phase, Group B subjects received 2 probes at intervals of one week during Group A's treatment phase. Limited probes were used during this period due to the risk of subject burn-out (Tawney & Gast, 1984).
3. Probes consisted of 1 set of 10 problems similar in type and format to sets administered during the baseline phase.
4. For each probe, the subject was called from the class and asked to meet examiner in the small resource room.
5. Upon arrival at each probe, the subject was greeted by name and engaged in idle chit-chat (e.g. weather, sports) for a few moments to re-establish rapport.
6. Examiner stated, "Once again, I would like you to complete a set of 10 problems in math. Please do the best you can. Don't worry about mistakes. Do you have any questions? (Experimenter responded to questions). Please begin."
7. Subject completed the task and returned the product to examiner. Experimenter dismissed the subject back to class.
8. This procedure was used for both probes with each of the Group B subjects.
9. At the conclusion of the treatment phase (Phase B) of Group A, Group B entered the treatment phase. All procedures identified for use with Group A during the treatment phase were replicated with Group B. For these procedures, please refer to subsection entitled "Treatment Phase (Phase B)."
10. Maintenance phase procedures developed for Group A were replicated for Group B. Please refer to the subsection entitled "Maintenance Phase (Phase C)" for these procedures.

CHAPTER IV

RESULTS

The purpose of this research was to study the efficacy of the curriculum-based assessment (CBA) model on math achievement of middle school students. The researcher's model of CBA consisted of the following components: (a) assessment of student's present performance levels (PPLs); b) error analysis; and (c) intervention based on error analysis.

The design of this study was composed of: a) baseline probes to establish PPLs; b) treatment followed by a probe of performance; and c) maintenance probes to determine if regression in the skills attained during treatment occurred. The researcher used nine students to investigate the effects of the intervention on subject performance.

Subjects were divided into two groups. Group A was designated as the experimental group and consisted of five subjects numbered one through five. Group B was designated as the control group and consisted of four subjects numbered five through nine. A multiple baseline single subject design was used for the study.

This chapter presents the findings for each of the nine subjects. Data are presented using graphic displays for each skill across conditions of baseline, interval probes (Group B only), treatment and maintenance. The individual research questions were not addressed in Chapter IV so that all findings across conditions could be presented in a comprehensive protocol for each subject. Findings and conclusions for each of the three research questions are summarized in Chapter V.

1. Does the use of CBA methods increase mathematical skills of middle school students who are at risk for failing mathematics?
2. Are differential effects in mathematical achievement observed between two heterogeneous groups of middle school students when the Experimental Group instruction uses the CBA and traditional methods, while the Control Group has traditional instruction only?
3. Are any gains achieved through the CBA model maintained after the assessment and instruction are terminated for a period of 14 to 18 days?

For both groups, baseline probes were taken on February 14, 1992. Group A began treatment sessions on February 17th continuing through February 27th. Maintenance probes were administered to Group A between March 12th and March 16th with the exception of Subject 1 who received these probes on March 6th.

For Group B, interval probes were taken between February 17th and February 27th. This was a non-treatment period for Group B occurring during the treatment period of Group A. Treatment began on March 2nd and extended through March 12th. Maintenance probes for Group B were taken between March 26th and March 30th.

Although the original study plan called for six treatment sessions of 20 minutes over a two-week period for each identified skill, the researcher quickly found the subjects capable of learning more than one skill per session. To accommodate subject needs and motivation, additional new skills were introduced, which resulted in fewer than six treatment sessions available for presenting and learning the new skills.

Thus the data reflect baseline, treatment sessions ranging from three to six sessions, and maintenance probes.

During the treatment period (two weeks beginning February 17,1992) for Group A, subjects in Group B were probed twice to detect any change in the skills identified during the initial baseline period. The results of baseline and non-treatment probes (interval probes) of subjects in Group B are presented in graphic form following each subject's narrative.

Each student's results are presented on growth-over-time graphs. Each graphic horizontal axis (abscissa) represents sessions with the subjects in the study. Each graph's vertical axis (ordinate) represents a subject's performance in the mathematical skill in which he/she was experiencing difficulty. The specific mathematical skills, which constituted the independent variable, were specific to the needs of each subject and tended to vary with subjects.

Subject 1 (S-1)

S-1 was a 12 year, 1 month old Hispanic male in the seventh grade. The dependent variable, growth-over-time in a skill area of mathematics, was measured in assessment sessions held over a three week period between February 14th and March 6th. S-1 contracted chicken pox at the end of the third treatment session and was unable to continue with the treatment sessions. However, maintenance probes were conducted with S-1 on March 6th. This was 14 days after his third treatment session.

The results of the PPL assessment sessions under the conditions of baseline, treatment, and maintenance are graphically displayed in Figure 1. The dependent variable was the percentage of problems correct in the area of addition of unlike fractions. Over the three baseline sessions,

S-1's scores reflect a flat profile of zero. During treatment his scores range from a low of 90 percent correct to a high of 100 percent correct. His maintenance scores reflect a flat profile of 100 percent correct.

SUBJECT 1

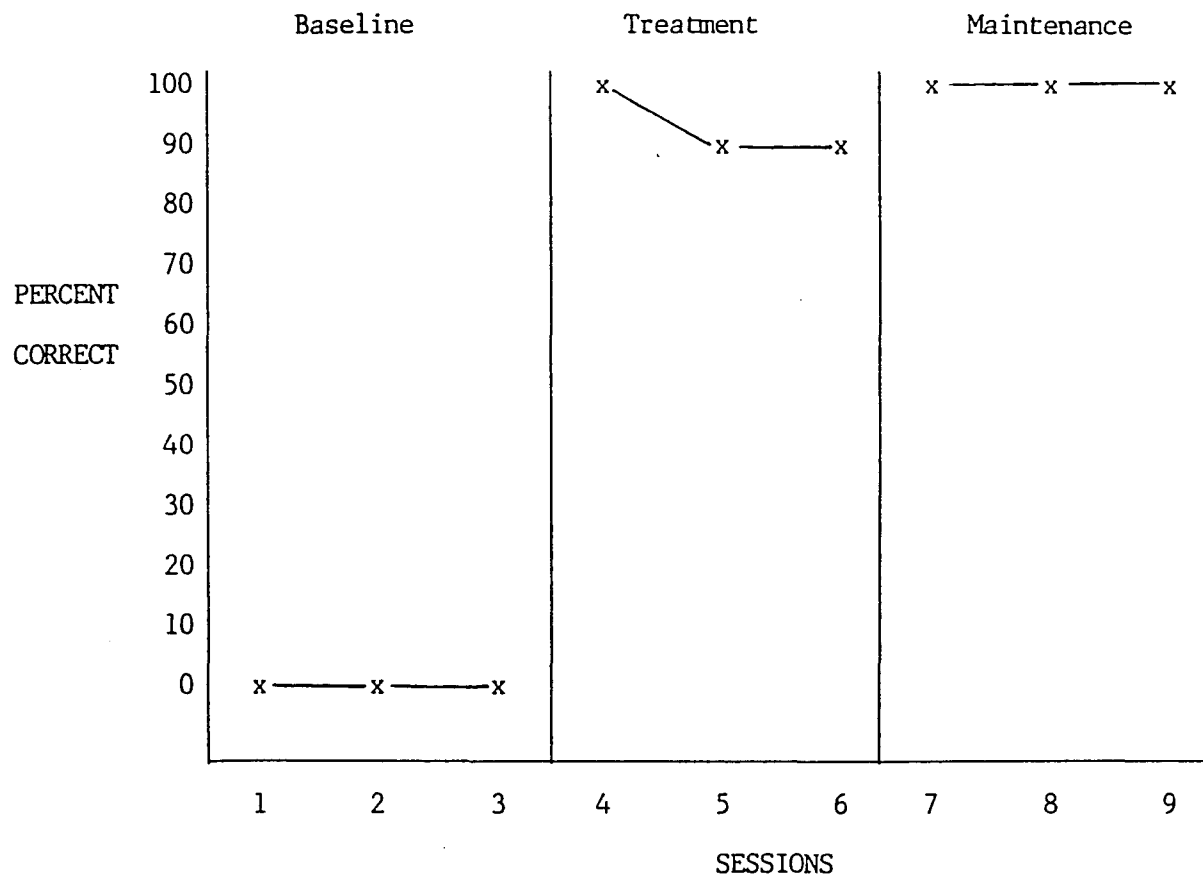


Figure 1. Subject 1 - The Effects of Treatment on Addition of Unlike Fractions

Subject 2 (S-2)

S-2 was a 14 year, 11 month old Hispanic female in the eighth grade. The dependent variable, growth-over-time in skill areas of mathematics, was measured in assessment sessions held over a four week period between February 14th and March 12th.

The results of the assessment sessions under the conditions of baseline, treatment and maintenance are graphically displayed in Figures 2 through 4. As shown in Figure 2, the dependent variable was the percent correct in the area of subtraction with regrouping. Over the baseline sessions S-2's scores reflect a positively accelerated slope with a progressive range of scores from 50 percent correct to 90 percent correct. The baseline slope will be discussed in Chapter 5. During treatment, her scores range from a low of 80-percent to a high of 100 percent correct. Her maintenance scores reflect a downward turn from a high of 90 percent correct to a low of 80 percent correct.

SUBJECT 2

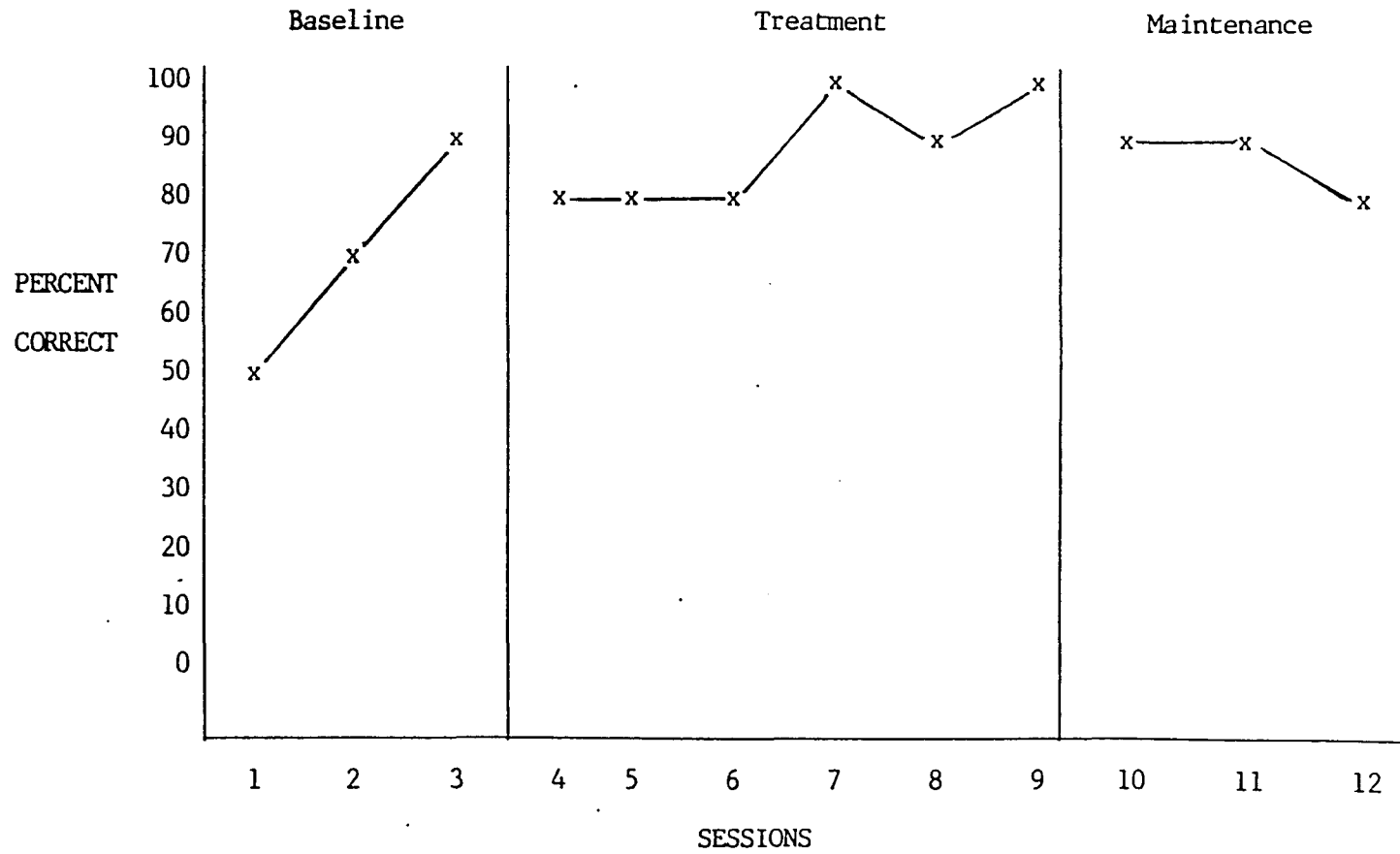


Figure 2. Subject 2 - The Effects of Treatment on Subtraction with Regrouping

As shown in Figure 3, the dependent variable was the percent correct in the area of addition of like fractions. Over the baseline sessions, S-2's scores reflect a stable profile of zero percent correct. During treatment, her scores also reflect a stable profile but of 100 percent correct. Her maintenance scores reflect a range of a low of 90 percent to a high of 100 percent correct.

SUBJECT 2

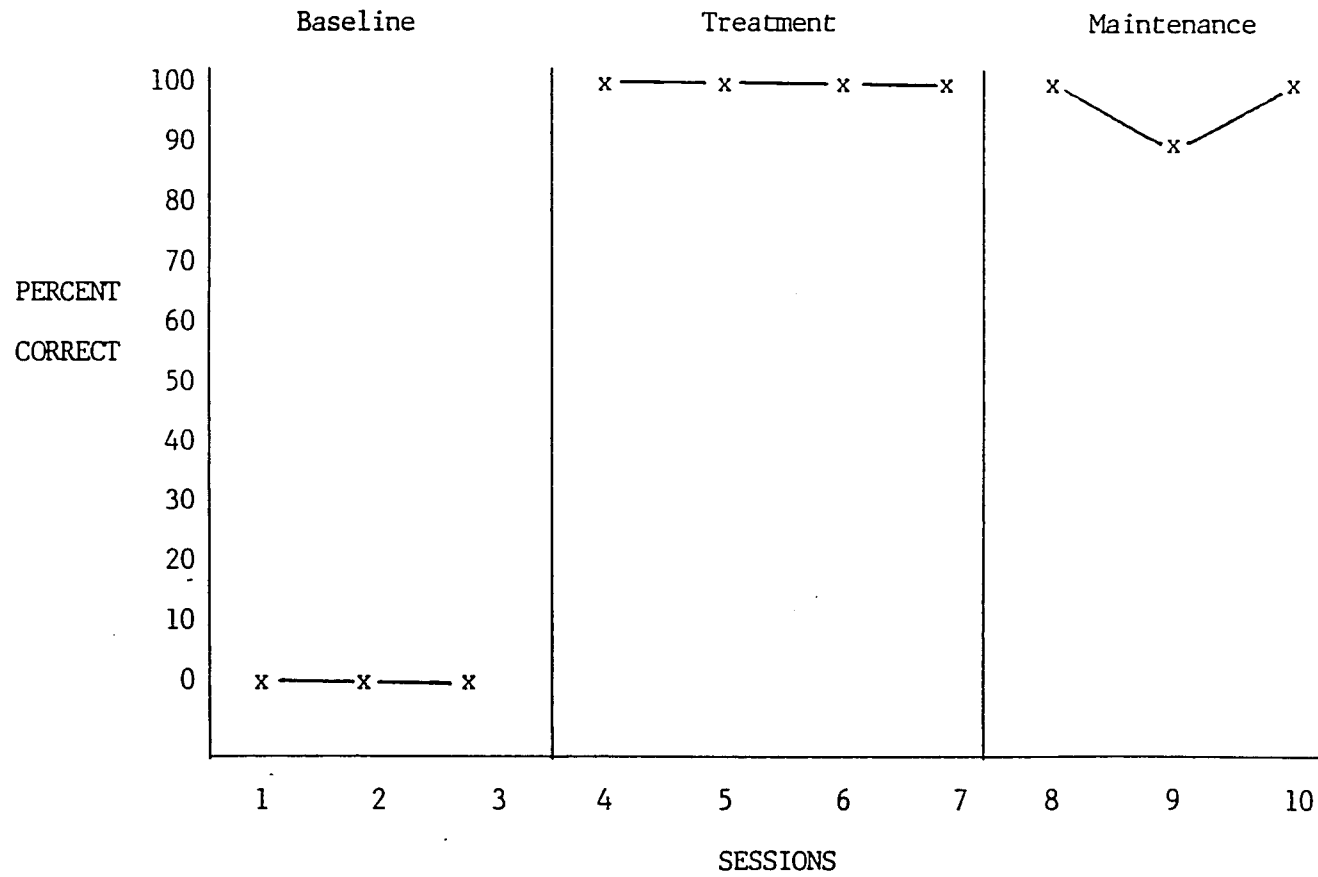


Figure 3. Subject 2 - The Effects of Treatment on Addition of Like Fractions

As represented in Figure 4, the dependent variable was the percent correct in the area of reducing fractions. Over the baseline sessions S-2's scores reflect a stable profile of zero percent correct. During treatment, her scores ranged from a low of 60 percent correct to a progressive high of 100 percent correct. However, maintenance probes found her range of a low of 20 percent correct and a high of 30 percent correct.

SUBJECT 2

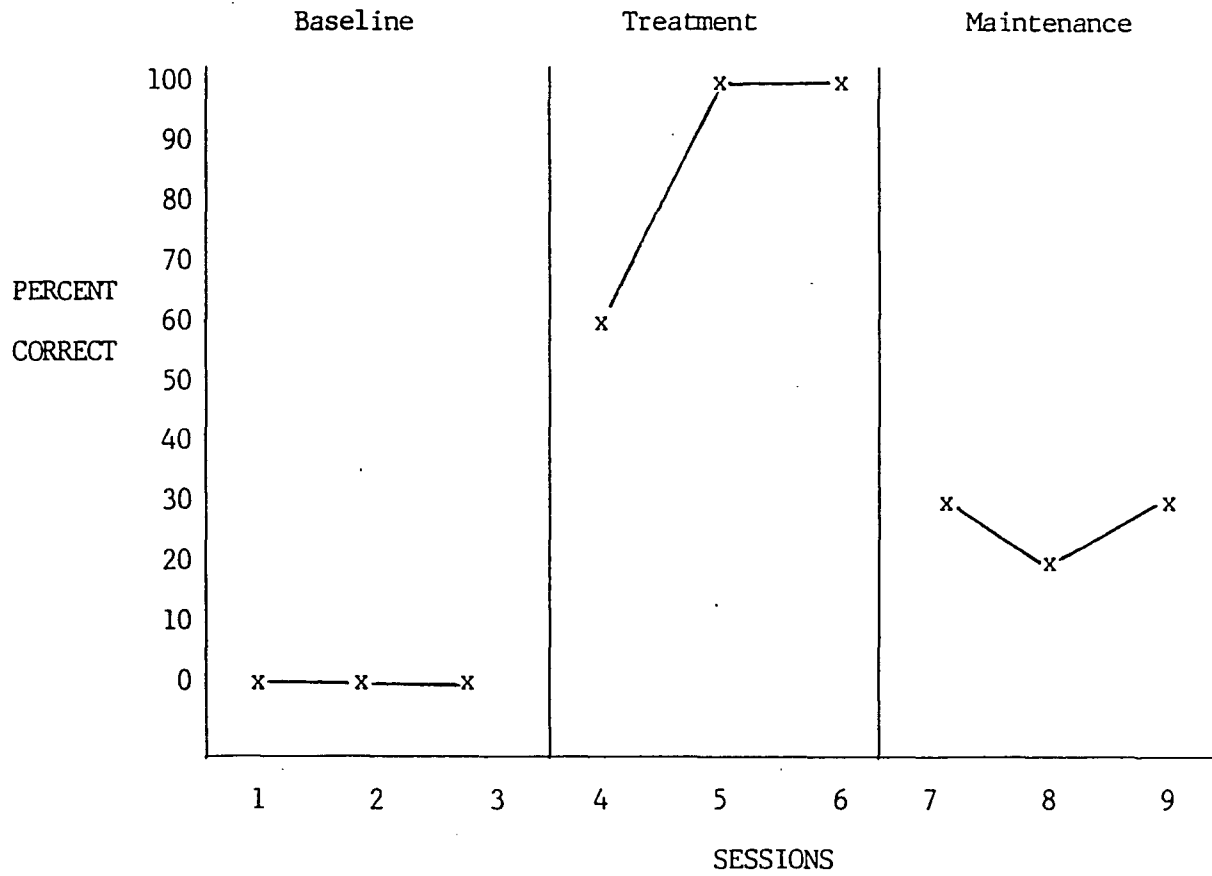


Figure 4. Subject 2 - The Effects of Treatment on Reducing Fractions

Subject 3 (S-3)

S-3 was a 14 year, 2 month old Hispanic male in the seventh grade. S-3 has a gran mal seizure disorder for which he takes Tegretol. The dependent variable, growth-over-time in a skill area of mathematics, was measured in assessment sessions held over an approximate four week period between February 14th and March 12th.

The results of the assessment sessions under the conditions of baseline, treatment, and maintenance are graphically displayed in Figure 5. Shown in Figure 5, the dependent variable was the percent correct in the area of simple division. Over the baseline sessions, S-3's scores reflect a stable, profile of zero percent correct. During treatment, his scores fluctuated with a low of 60 percent correct. S-3 had a medical appointment for an electroencephalogram (EEG) during the morning of the lowest scored probe. This was a part of a routine assessment process for his seizure disorder. However, he was clearly affected by the test as observed by his need to discuss it at some length. During maintenance probes, S-3 profiled a positively accelerated slope with a beginning of 70 percent correct and a high of 100 percent correct.

SUBJECT 3

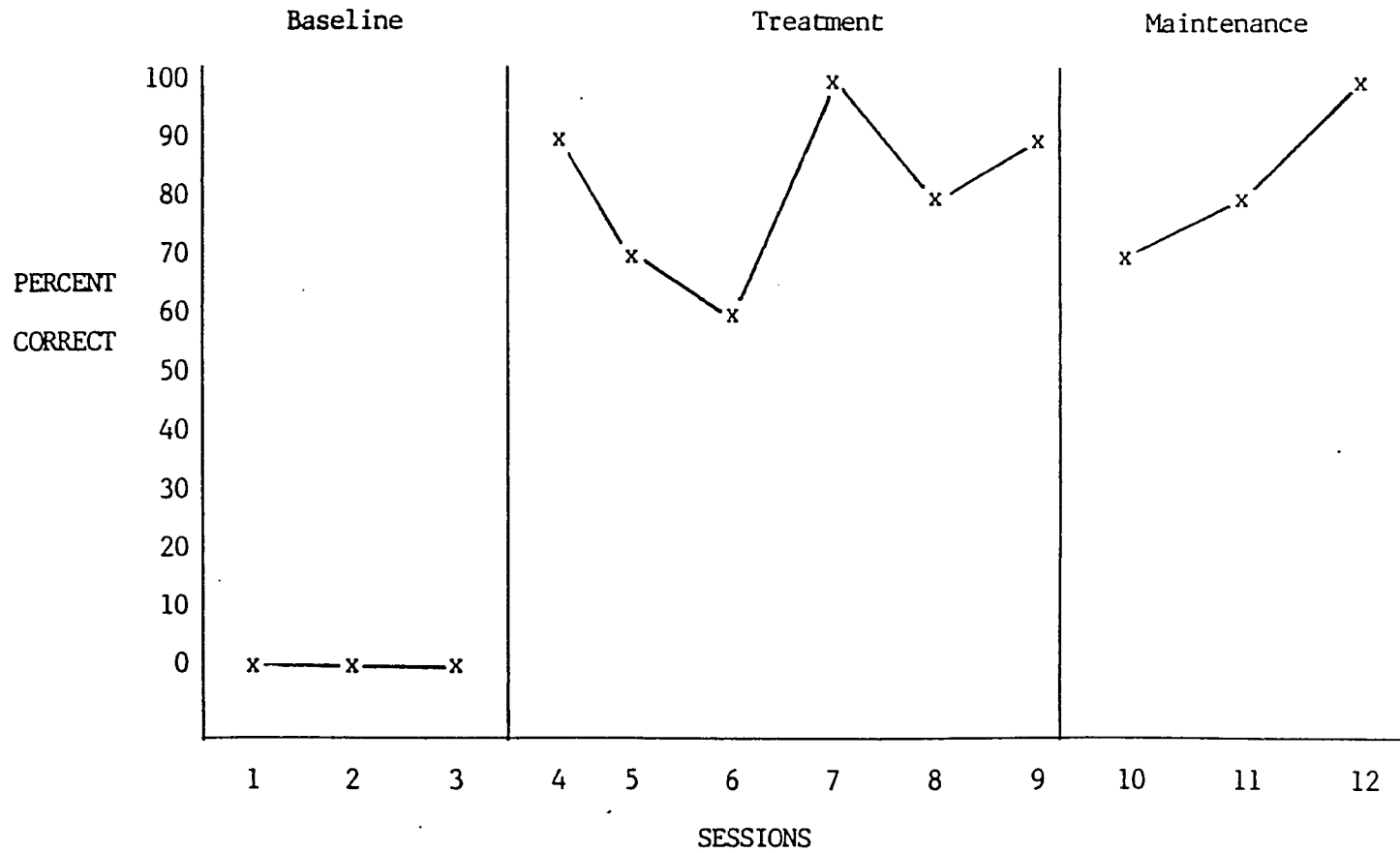


Figure 5. Subject 3 - The Effects of Treatment on Simple Division

Subject 4 (S-4)

S-4 is a 13 year, 4 month old Hispanic male in the seventh grade. The dependent variable, growth-over-time in a skill area of mathematics, was measured in assessment sessions held over a four week period from February 14th through March 12th.

The results of the assessment sessions under the conditions of baseline, treatment, and maintenance are graphically displayed in Figures 6 through 8. As shown in Figure 6, the dependent variable was the percent correct in the area of addition of unlike fractions. Over the baseline sessions, S-4 reflected a stable profile of zero percent correct. During treatment, his scores ranged from a low of 80 percent correct to a high of 100 percent correct. His maintenance profile is very stable at a 100 percent correct.

SUBJECT 4

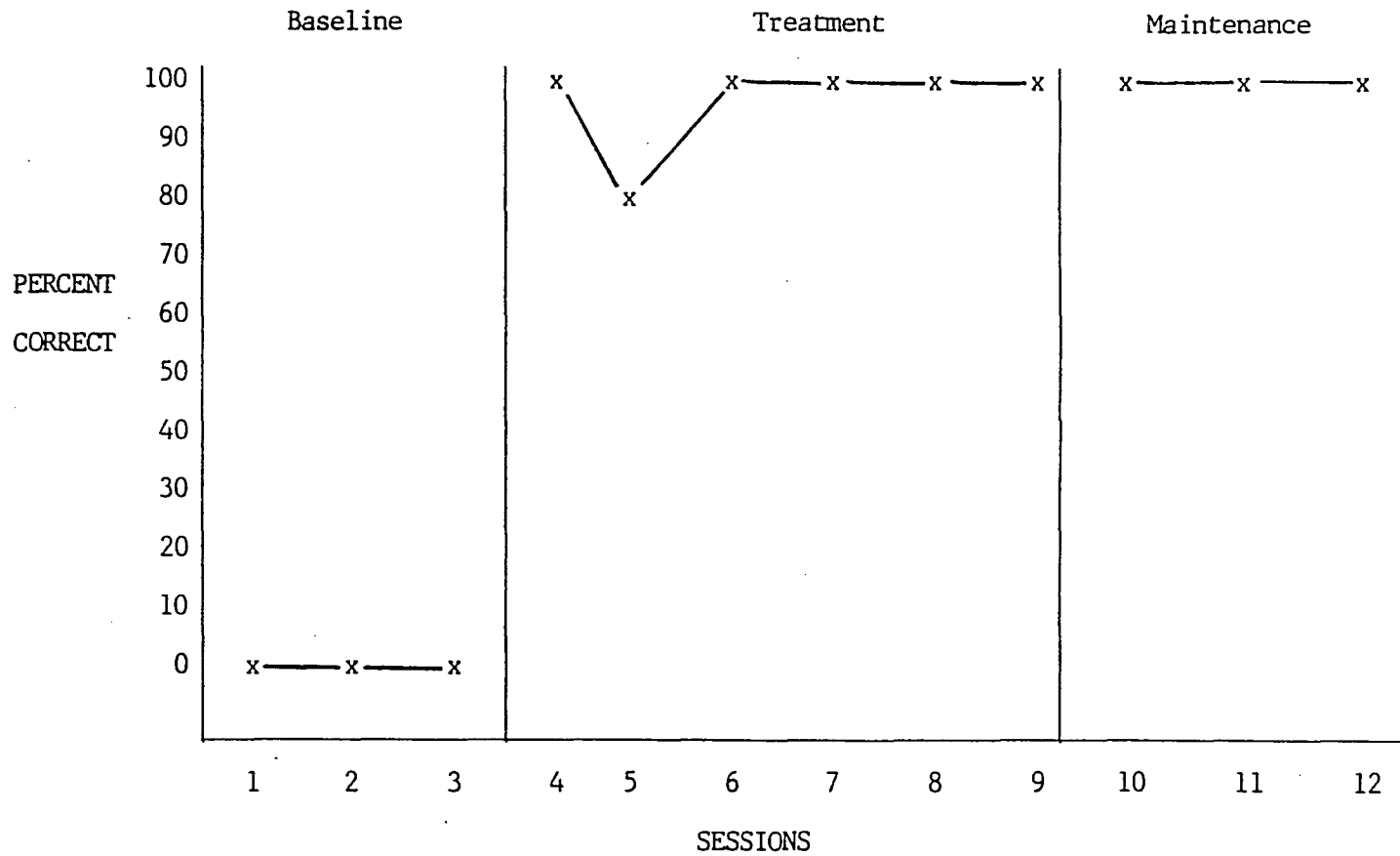


Figure 6. Subject 4 - The Effects of Treatment on Addition of Unlike Fractions

As shown in Figure 7, the dependent variable was the percent correct in the area of changing fractions to percents. Over the baseline sessions, S-4's scores reflect a stable profile of zero percent correct. Treatment probes reflect a positively accelerated slope with a downward turn at the end. Scores ranged from a low of 60 percent correct to a high of 100 percent correct. His maintenance profile also reflects a downward turn for the final session. His scores range from a high of 100 percent correct to a low of 90 percent correct.

SUBJECT 4

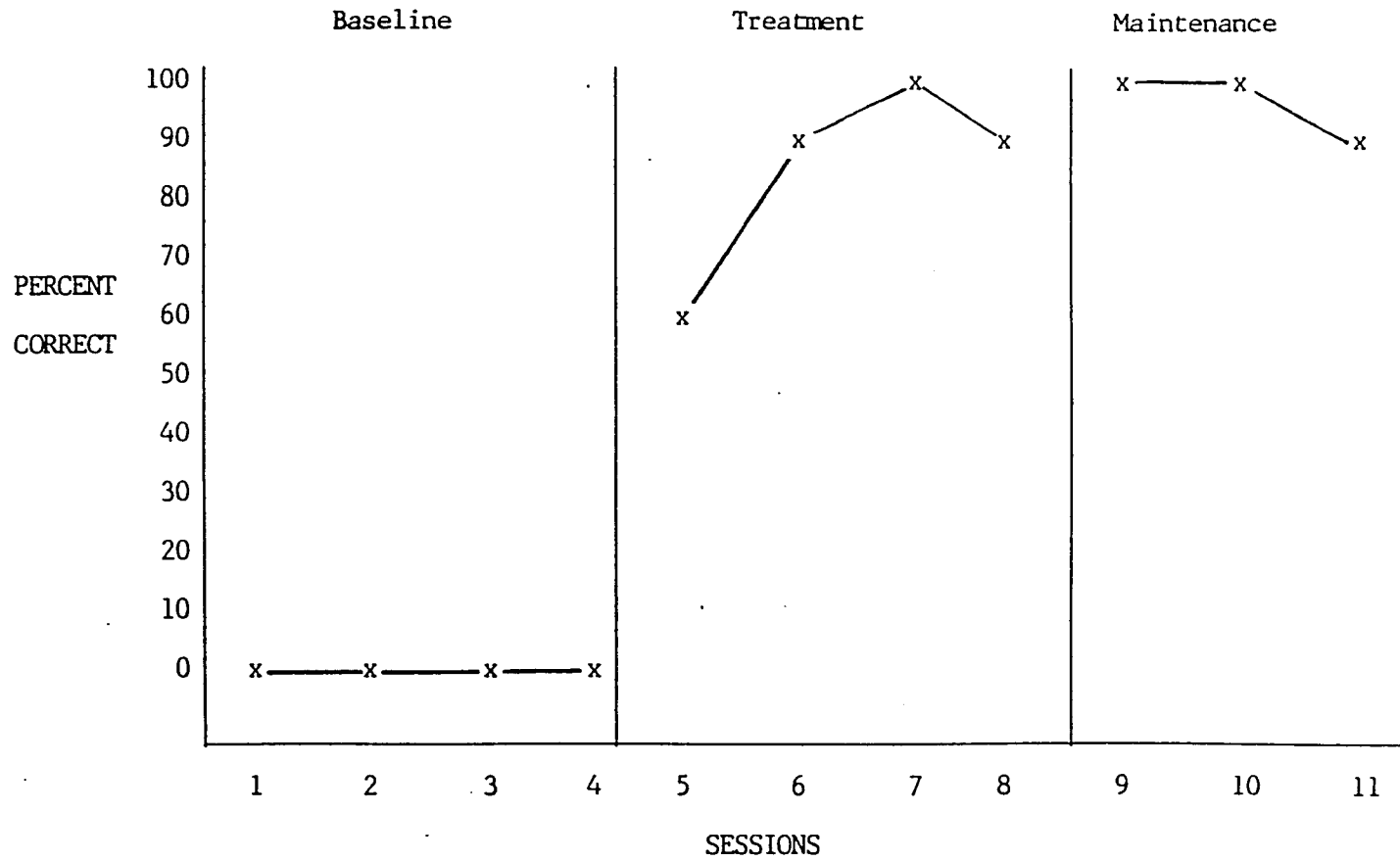


Figure 7. Subject 4 - The Effects of Treatment on Changing Fractions to Percents

As shown in Figure 8, the dependent variable was the percent correct in the area of multiplication of fractions. Over the baseline sessions, S-4's scores reflect a stable profile of zero percent correct. Treatment probes reflect a positively accelerated slope with a progressive range of a low score of 50 percent correct to a high score of 100 percent correct. Maintenance scores fluctuate beginning with a 70 percent correct moving to 100 percent correct and falling to 90 percent correct.

SUBJECT 4

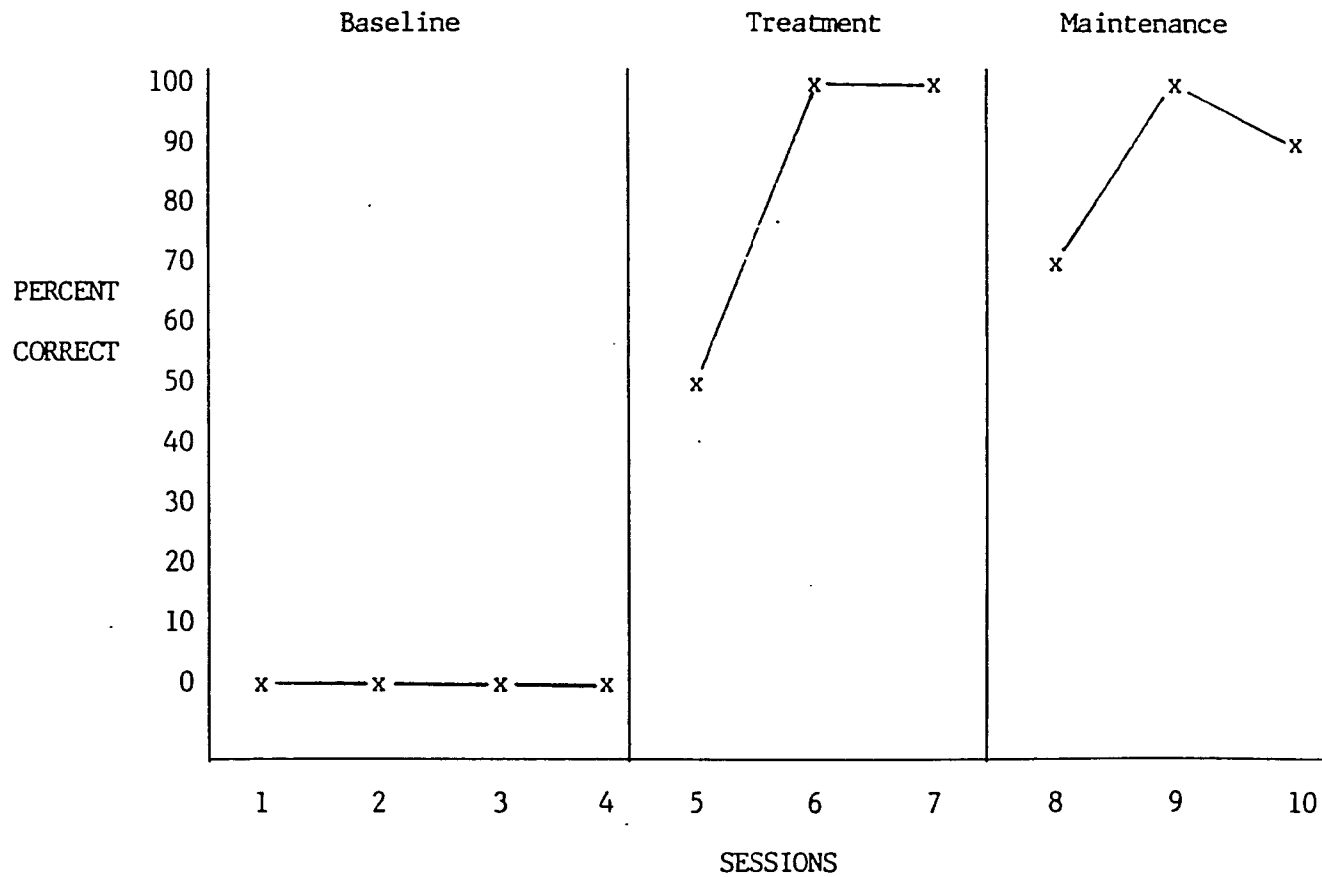


Figure 8. Subject 4 - The Effects of Treatment on Multiplication of Fractions

Subject 5 (S-5)

S-5 was a 12 year, 9 month old white female in the seventh grade. The dependent variable, growth-over-time in a skill area of mathematics, was measured in assessment sessions held over a four week period from February 14th through March 12th.

The results of the assessment sessions under the conditions of baseline, treatment, and maintenance are graphically displayed in Figures 9 and 10. As shown in Figure 9, the dependent variable was the percent correct in the area of division with zeros. Over the baseline sessions, S-5 reflected a stable profile of zero percent correct. During treatment, her scores ranged from a low of 90 percent correct to a high of 100 percent correct. Maintenance probes reflected the same pattern with a low percent correct and a high of 100 percent correct.

SUBJECT 5

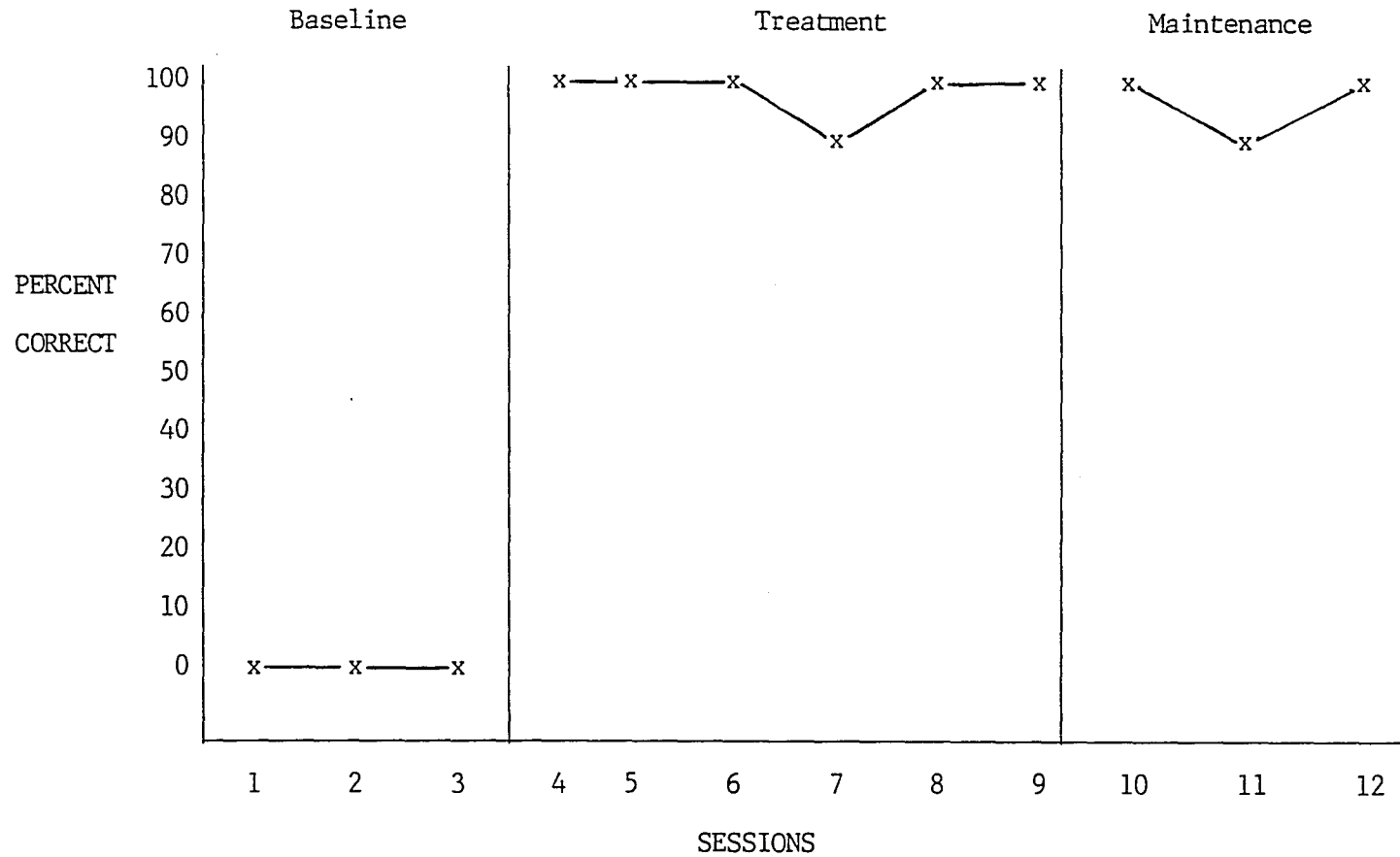


Figure 9. Subject 5 - The Effects of Treatment on Division with Zeros

As shown in Figure 10, the dependent variable was the percent correct in the area of addition of unlike fractions. Over the baseline sessions, S-5 reflected a stable profile of zero percent correct. During treatment, she fluctuated with a range of a low of 70 percent correct to a high of 100 percent correct. Maintenance probes reflected a low score of 90 percent correct and a high score of 100 percent correct.

SUBJECT 5

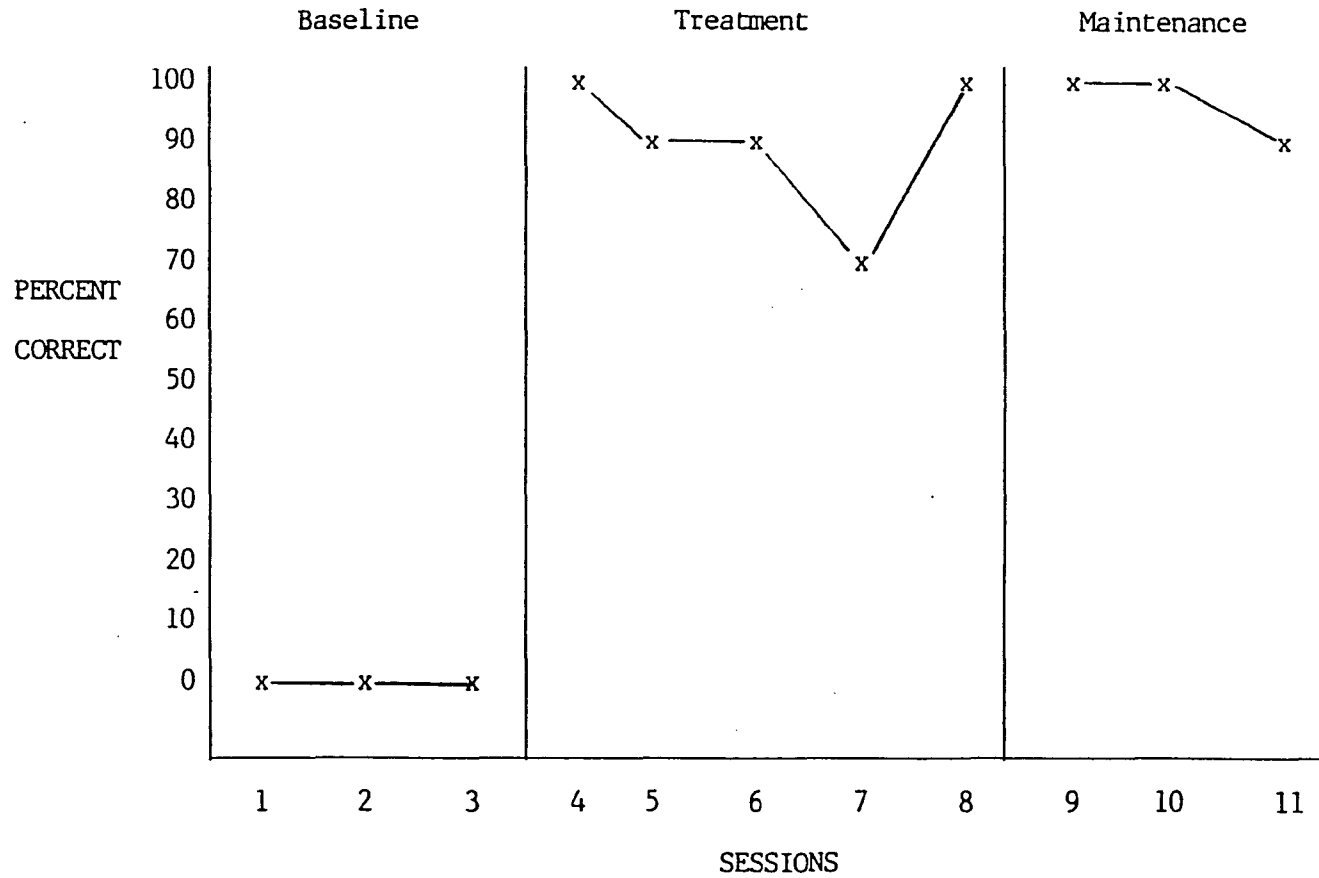


Figure 10. Subject 5 - The Effects of Treatment on Addition of Unlike Fractions

Subject 6 (S-6)

S-6 was a 12 year, 7 month old white male in the seventh grade. The dependent variable, growth-over-time in a skill area of mathematics, was measured in assessment sessions held over a period of approximately six weeks from February 14th through March 26th.

S-6 is in Group B (control group) so assessment included the added condition of interval probes. These occurred during the treatment phase of Group A at which time Group B was receiving no intervention.

The results of the assessment sessions under conditions of baseline, interval, treatment, and maintenance are graphically displayed in Figures 11 through 13. As shown in Figure 11, the dependent variable was the percent correct in the addition of unlike fractions. During baseline, S-6 reflected a stable profile of zero percent correct. However, during the interval with no interventions provided through the study, S-6 reflects a 100 percent correct response on both probes. During treatment, his scores fluctuate with a low of 70 percent correct and a high of 100 percent correct. Maintenance is a stable profile at 100 percent correct.

SUBJECT 6

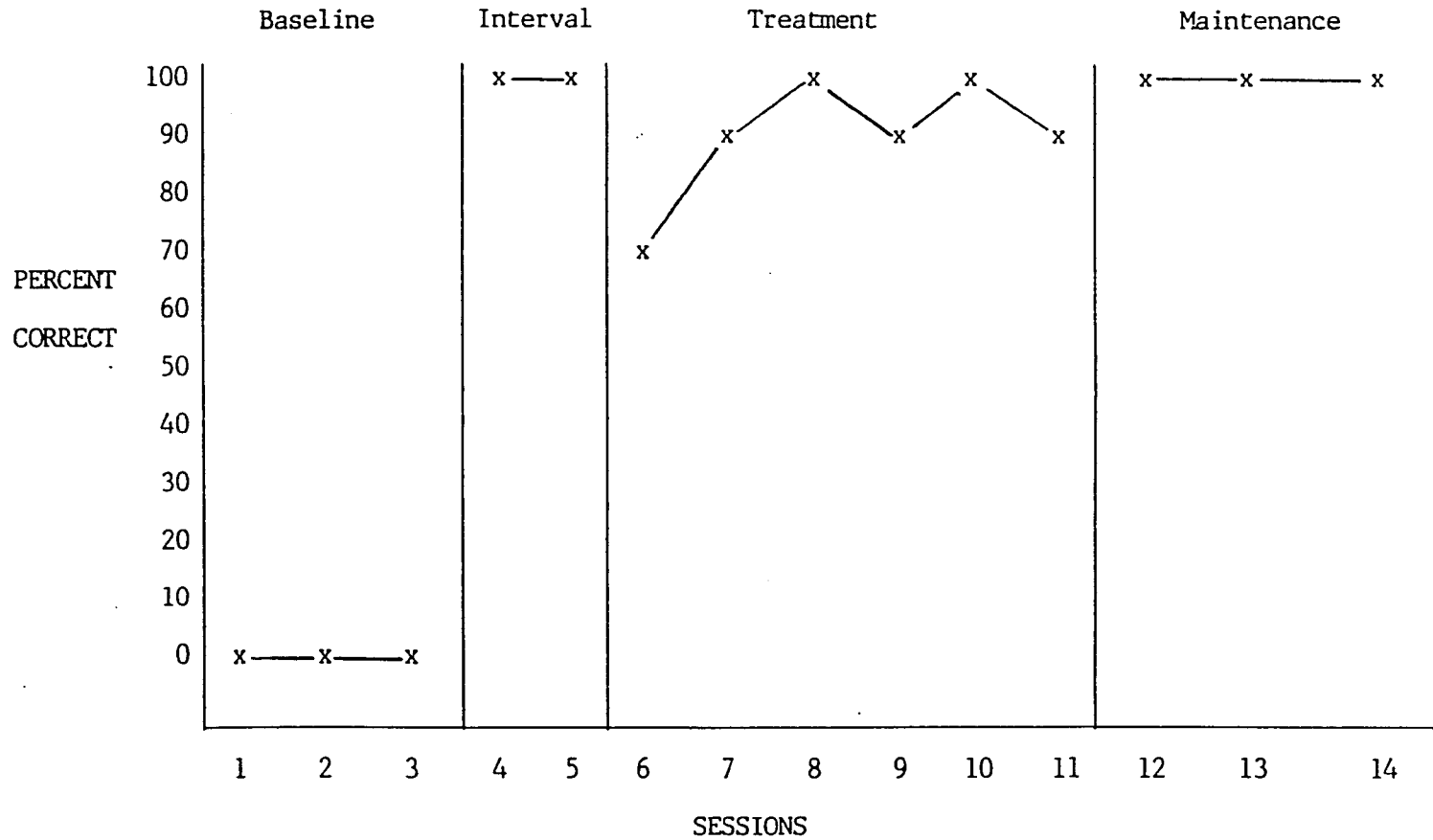


Figure 11. Subject 6 - The Effects of Treatment on Addition of Unlike Fractions

As shown in Figure 12, the dependent variable was the percent correct in the changing of fractions to decimals. During baseline, the subject initially presented a negatively accelerated slope recovering during probe 4 to create a positively accelerated slope. Scores ranged from a low of 10 percent correct to a high of 50 percent correct. During treatment, S-6 began the first session at 70 percent correct plummeting to zero percent correct during probe 2 followed by 4 consecutive probes at 100 percent correct. Maintenance scores range from a low of 90 percent correct to a high of 100 percent correct.

SUBJECT 6

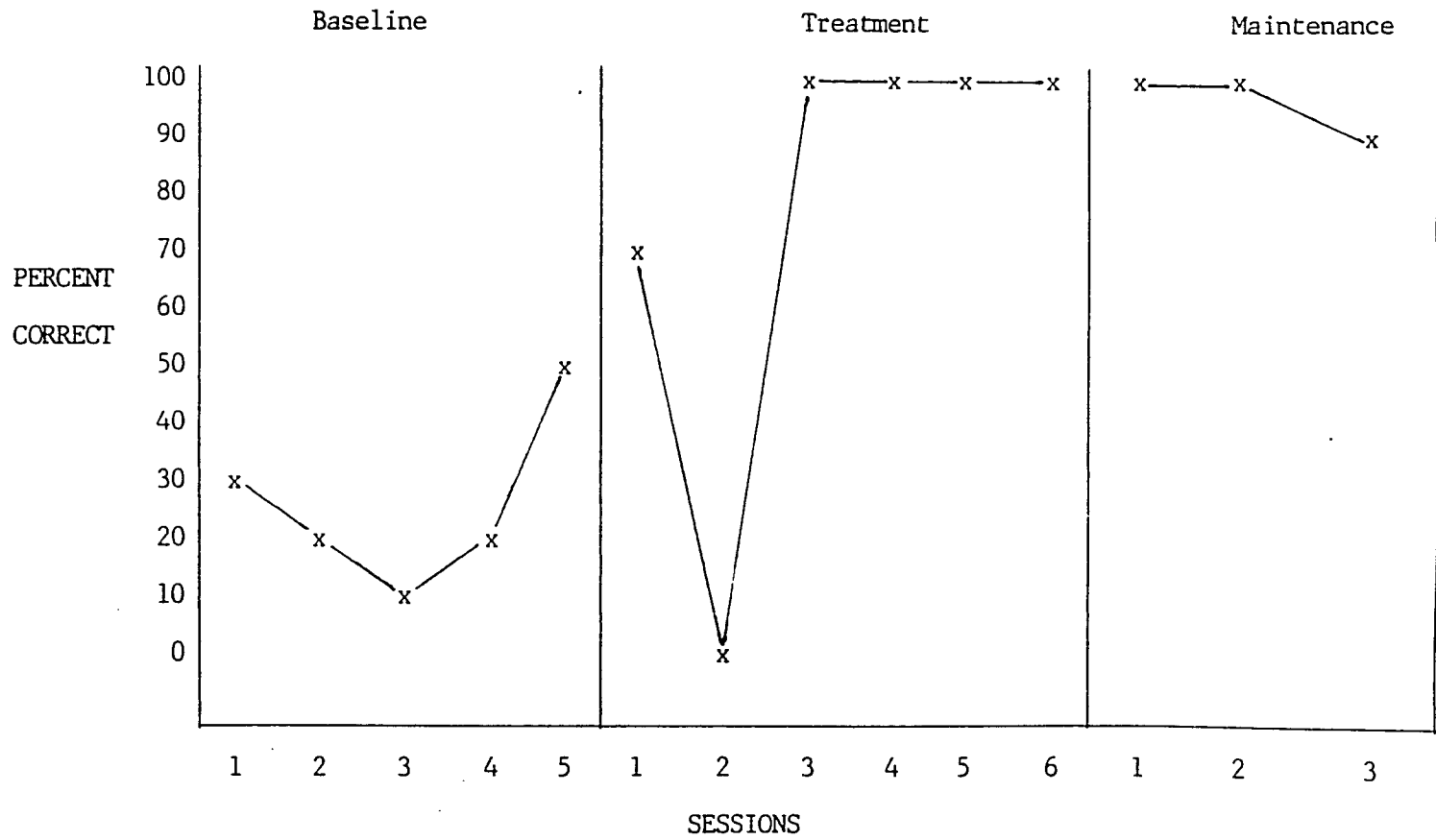


Figure 12. Subject 6 - The Effects of Treatment on Changing Fractions to Percents

As shown in Figure 13, the dependent variable was the percent correct in the multiplication of fractions. During baseline, S-6 presented a stable profile of zero percent correct. During treatment, he attained a low score of 70 percent correct and a high score of 100 percent correct. Maintenance probes presented a stable profile of 100 percent correct.

SUBJECT 6

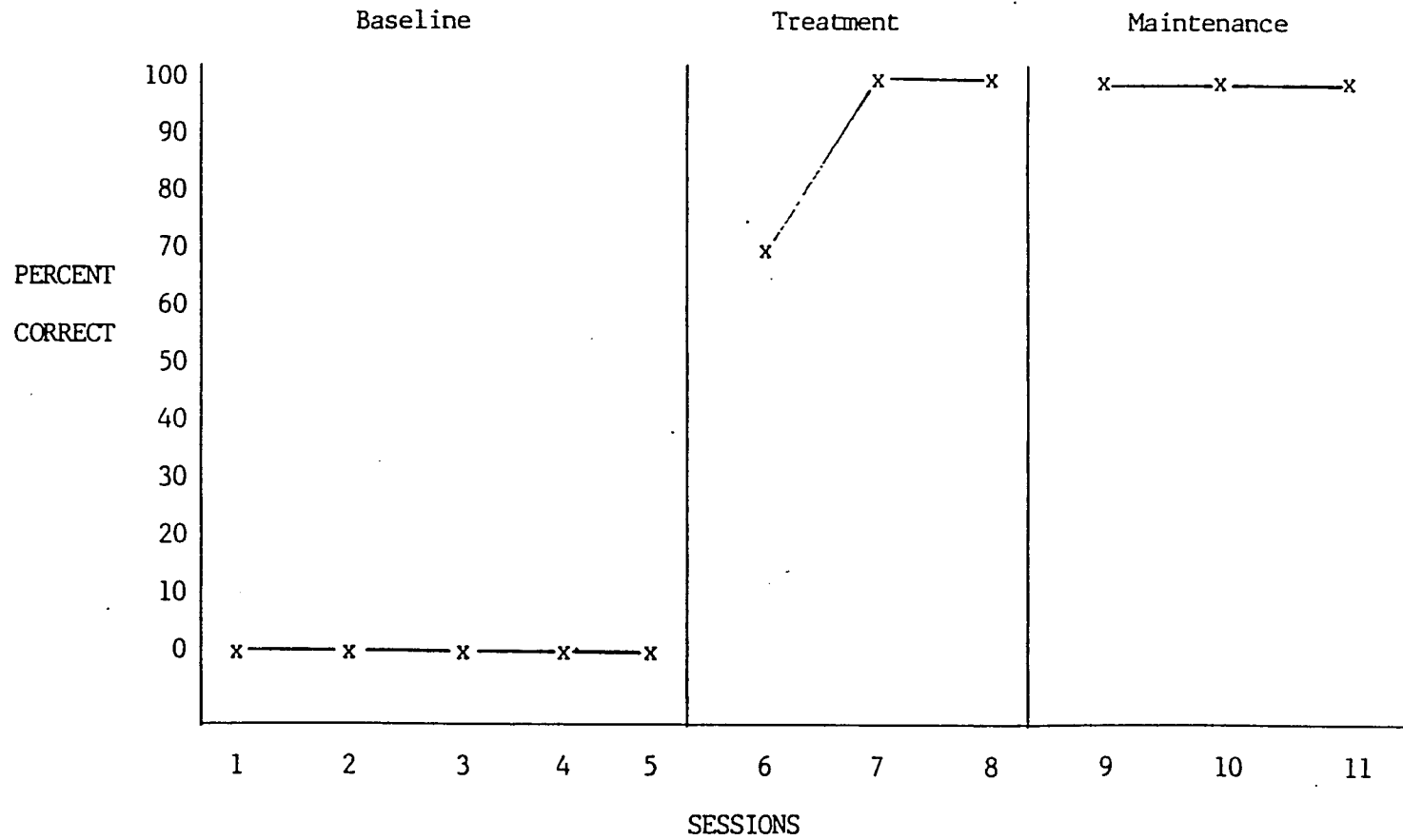


Figure 13. Subject 6 - The Effects of Treatment on Multiplication of Fractions

Subject 7 (S-7)

S-7 was a 12 year, 8 month old Hispanic male in the seventh grade. The dependent variable, growth-over-time in a skill area of mathematics, was measured in assessment sessions held over a period of approximately six weeks from February 14th through March 26th.

The results of the assessment sessions under conditions of baseline interval, treatment, and maintenance are graphically displayed in Figures 14 through 16. As shown in Figure 14, the dependent variable was the addition of like fractions. During baseline and interval conditions, S-7's scores reflect a stable profile of zero percent correct. Given treatment, he maintained a 100 percent correct across all sessions. During maintenance, he displays a positively accelerated slope with a range of scores from 60 percent correct to 100 percent correct.

SUBJECT 7

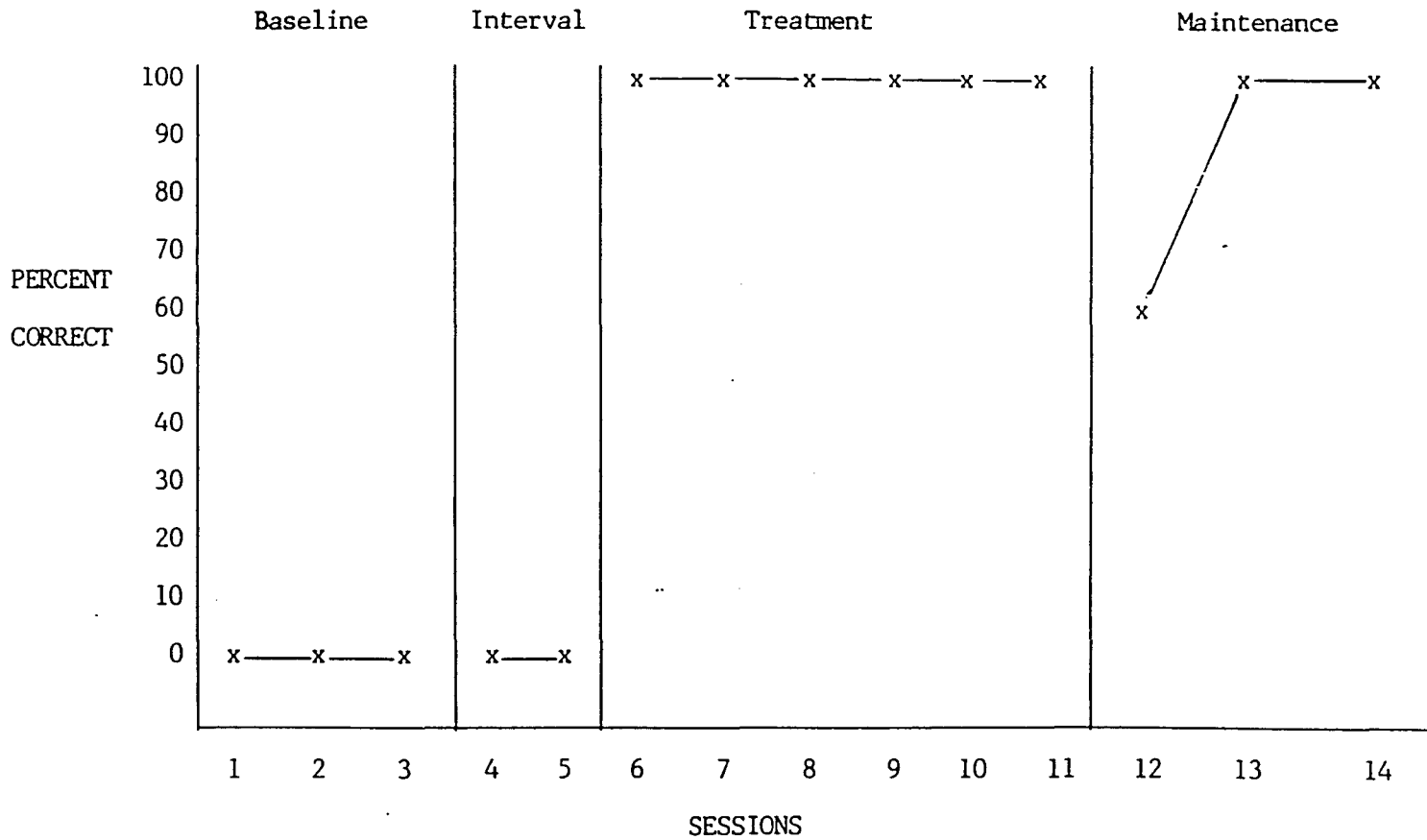


Figure 14. Subject 7 - The Effects of Treatment on the Addition of Like Fractions

As shown in Figure 15, the dependent variable was the addition of unlike fractions. During baseline, the subject's scores reflect a stable profile at zero percent correct. During the treatment sessions, S-7's scores range from a low score of 80 percent correct to a high score of 90 percent correct. During maintenance, S-7's scores slope slightly downward with a high of 100 percent correct moving to 90 percent correct.

SUBJECT 7

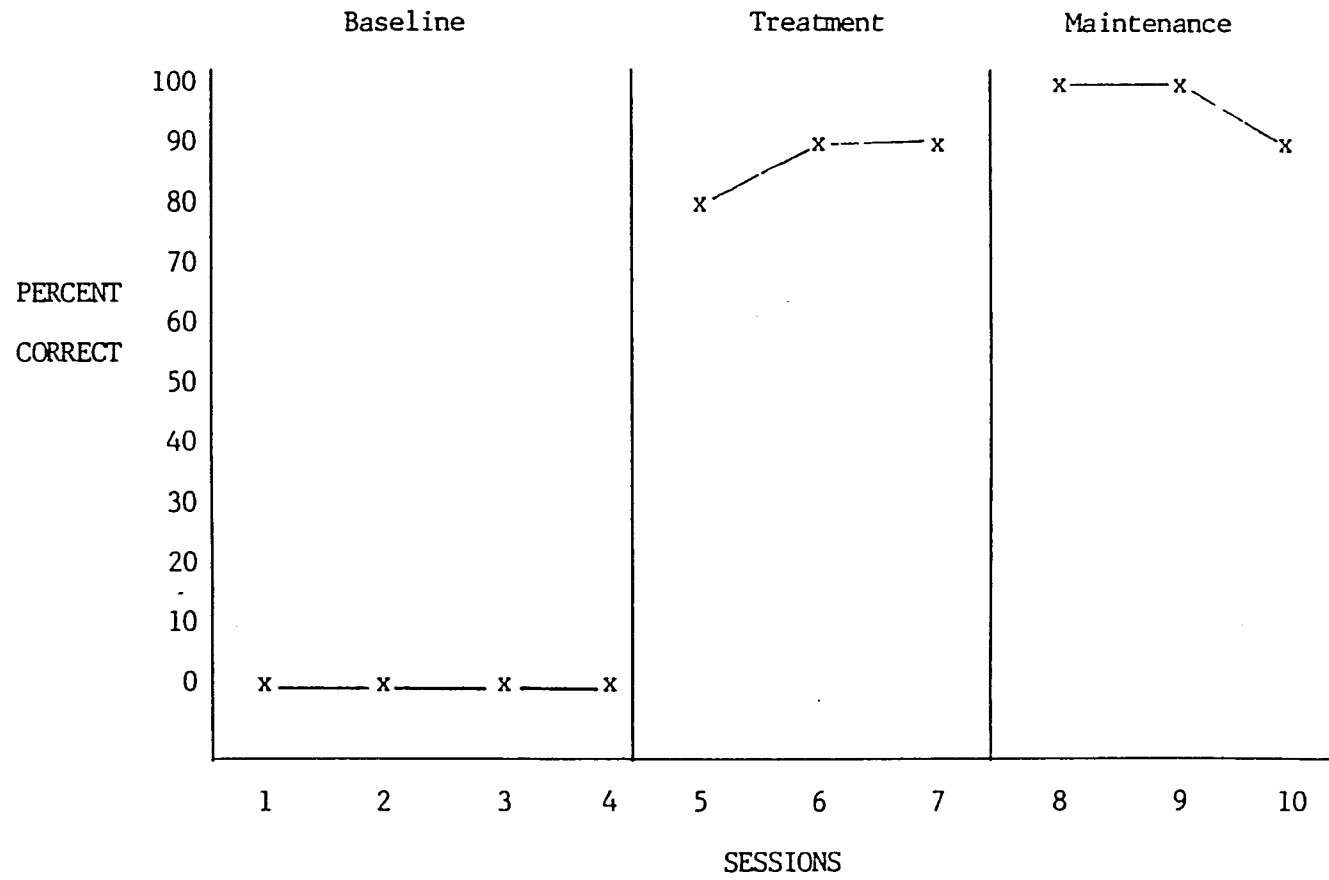


Figure 15. Subject 7 - The Effects of Treatment on the Addition of Unlike Fractions

As shown in Figure 16, the dependent variable was the percent correct of reducing fractions. During baseline, the subject's (S-8) scores reflect a stable profile at zero percent correct. During treatment, scores reflected a range of 80 percent correct to 100 percent correct. Maintenance scores reflect a drop from 100 percent to 70 percent correct.

SUBJECT 7

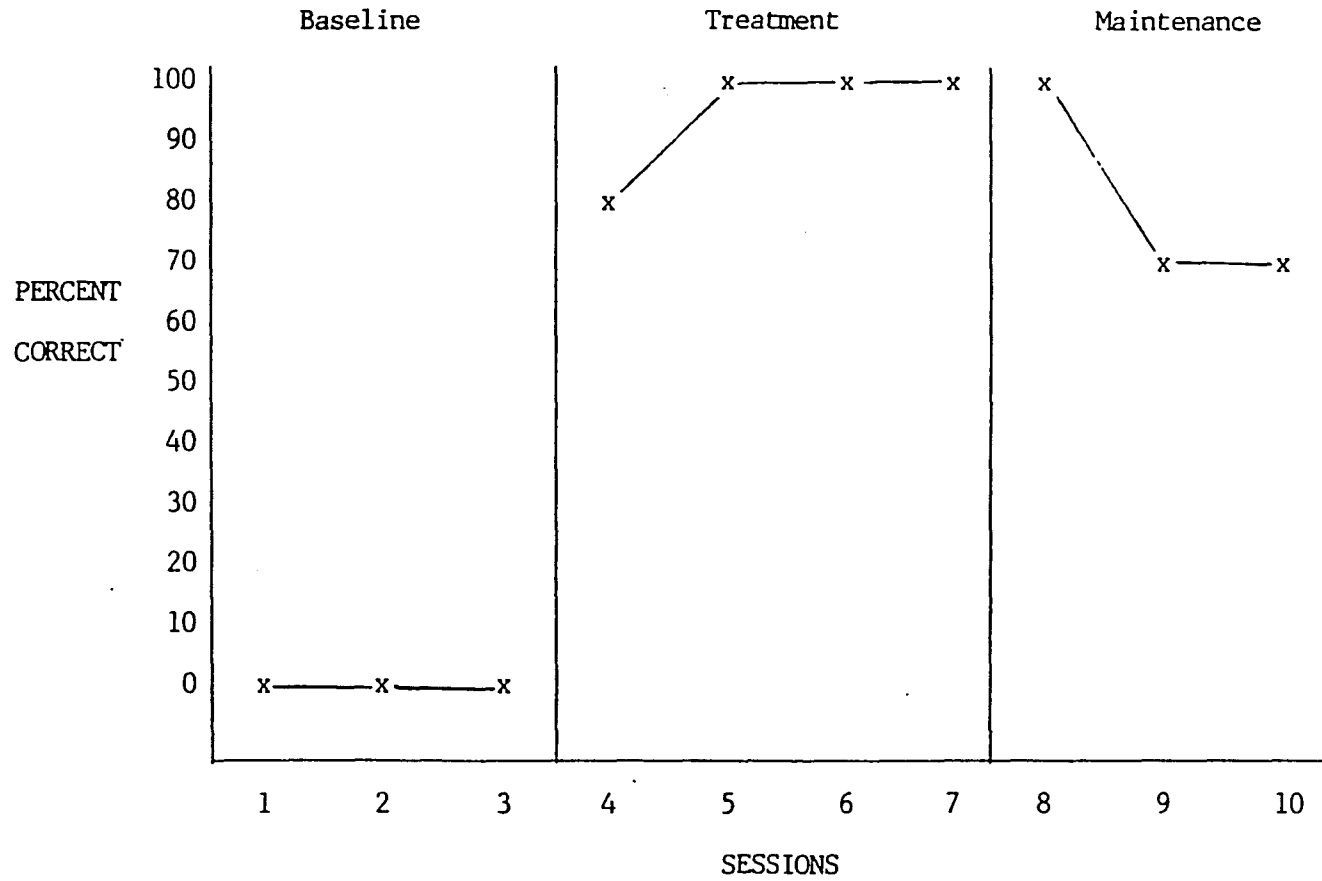


Figure 16. Subject 7 - The Effects of Treatment on Reducing Fractions

Subject 8 (S-8)

S-8 was a 14 year, 1 month old Native American female in the eighth grade. The dependent variable, growth-over-time in a skill area of mathematics, was measured in assessment sessions held over a period of approximately six weeks from February 14th through March 26th.

The results of the assessment sessions under conditions of baseline, interval, treatment, and maintenance are graphically displayed in Figures 17 through 19. As shown in Figure 17, the dependent variable was the changing of fractions to percents. During baseline, S-8's scores range from a low of zero percent correct to a high of 10 percent correct. This occurs in the same profile during the interval sessions. Treatment scores range from a low of 90 percent correct to a high of 100 percent correct. However, maintenance probes reflect a stable profile of zero percent correct.

SUBJECT 8

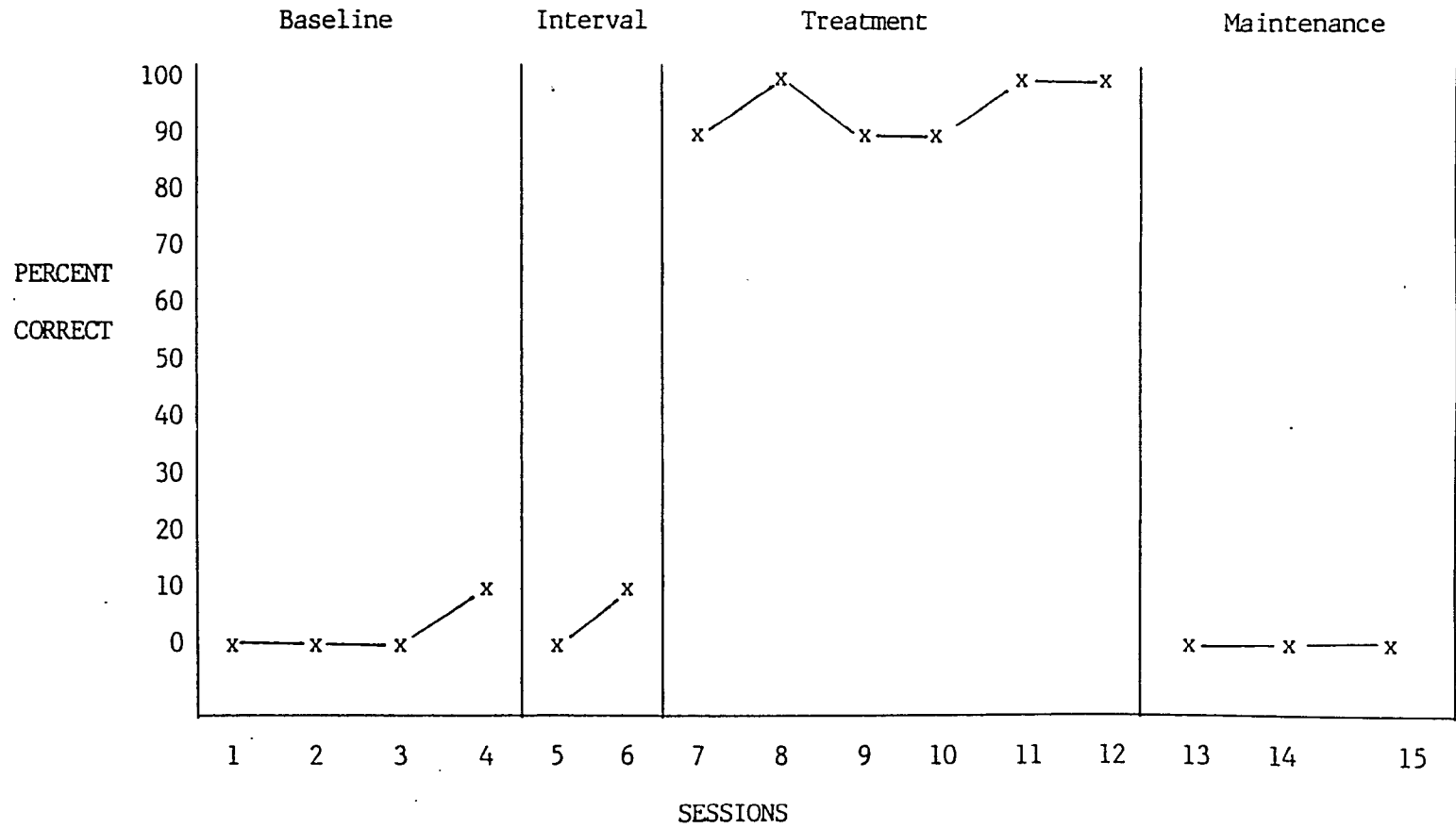


Figure 17. Subject 8 - The Effects of Treatment on Changing Fractions to Percents

As shown in Figure 18, the dependent variable was multiplication of fractions. During baseline, S-8's scores reflect a stable profile of zero percent correct. During treatment the scores ranged from a low of 90 percent correct to a high of 100 percent correct. Maintenance scores reflect a negative slope of a high of 100 percent dropping to 90 percent correct.

SUBJECT 8

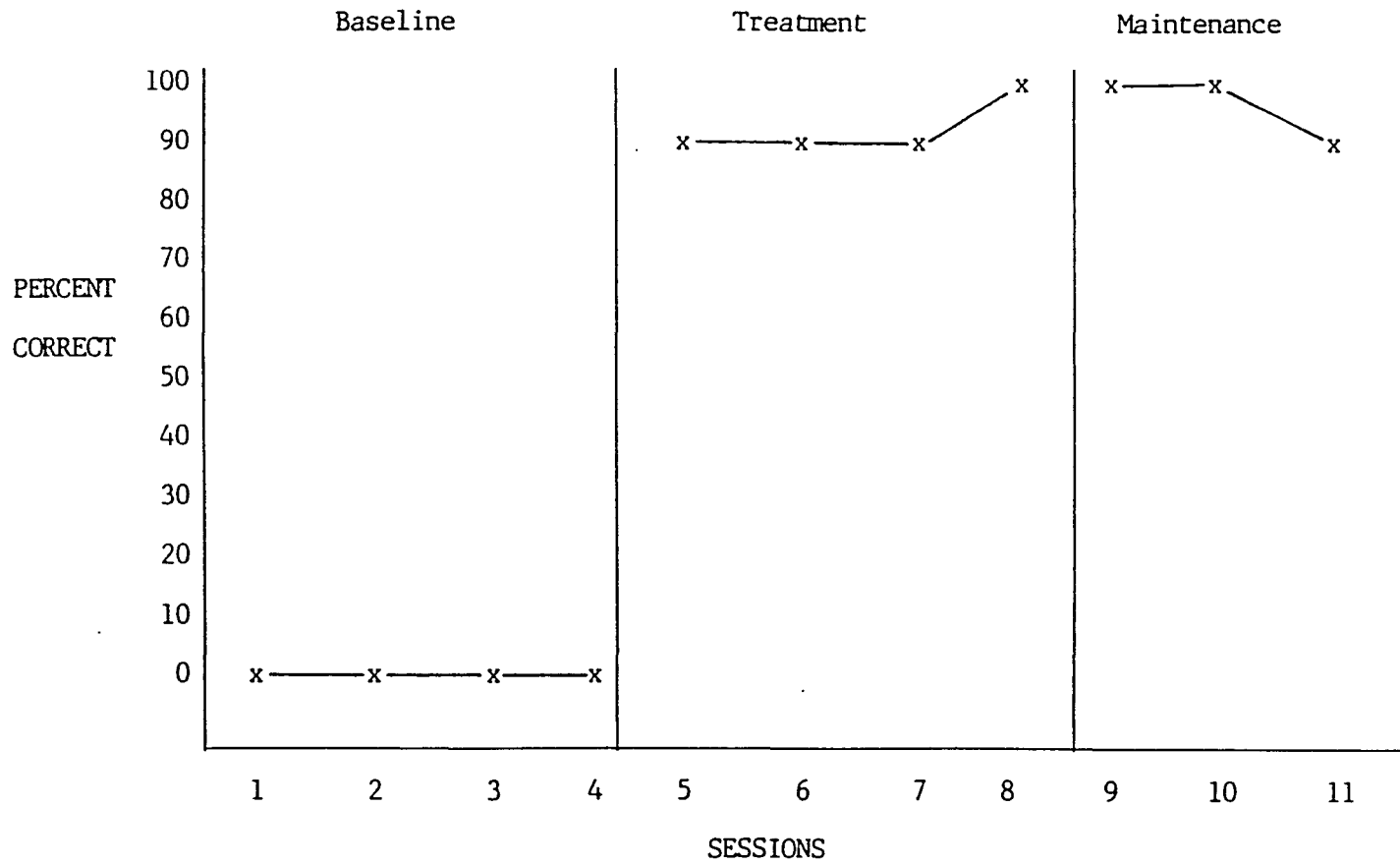


Figure 18. Subject 8 - The Effects of Treatment on Multiplication of Fractions

As shown in Figure 19, the dependent variable was the division of fractions. During baseline, S-8's scores reflect a stable profile of zero percent correct. During treatment the scores range from a low of 90 percent correct to a high of 100 percent correct. Maintenance scores reflect a stable profile of 100 percent correct.

SUBJECT 8

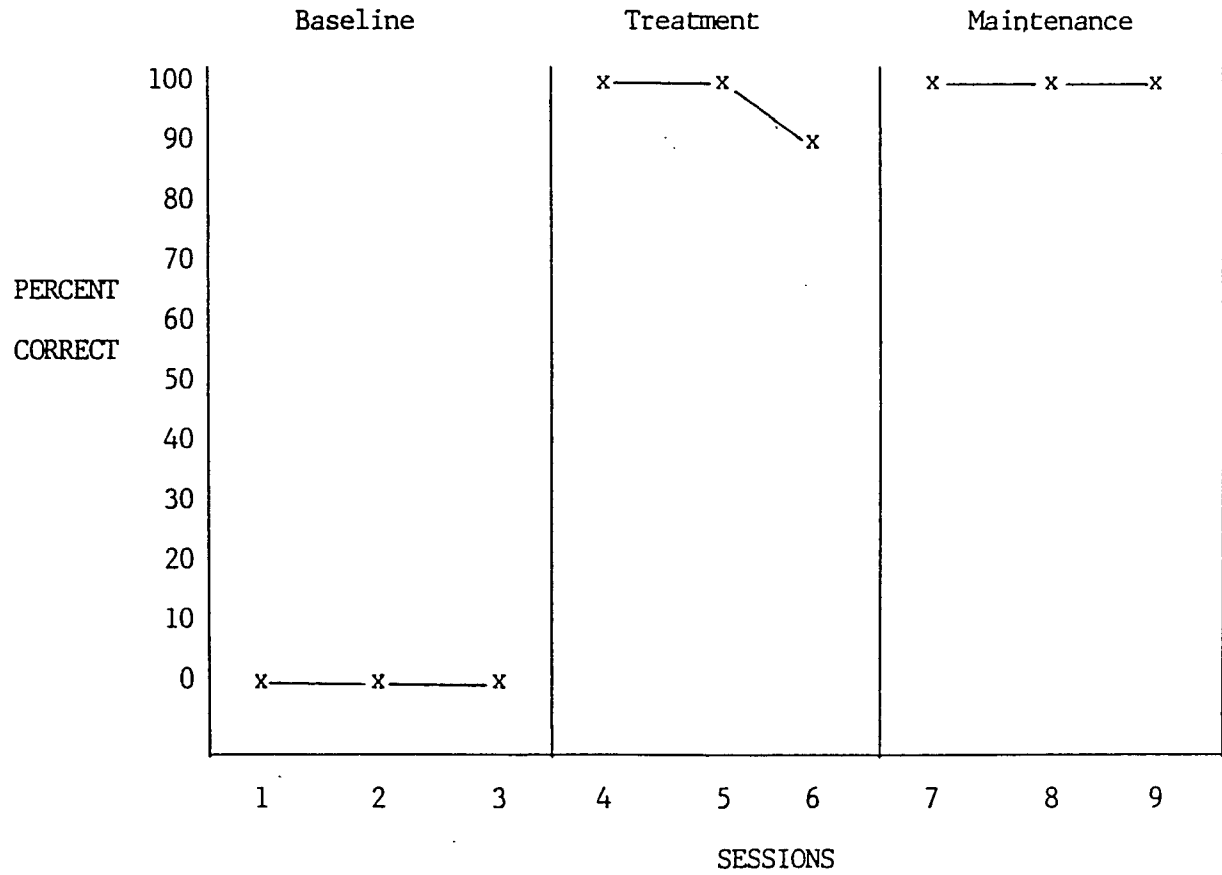


Figure 19. Subject 8 - The Effects of Treatment on Division of Fractions

Subject 9 (S-9)

S-9 was a 14 year, 8 month old black male in the eighth grade. The dependent variable, growth-over-time in a skill area in mathematics, was measured in assessment sessions held over a period of approximately six weeks from February 14th through March 26th.

The results of the assessment sessions under conditions of baseline, interval, treatment, and maintenance are graphically displayed in Figures 20, the dependent variable was changing fractions to percents. During baseline and interval, S-9's scores reflect a stable profile of zero percent correct. Treatment scores show a progression upwards with a drop during session two. Scores range from 50 percent correct to 100 percent correct. Maintenance scores range from a low of 90 percent correct to 100 percent correct.

SUBJECT 9

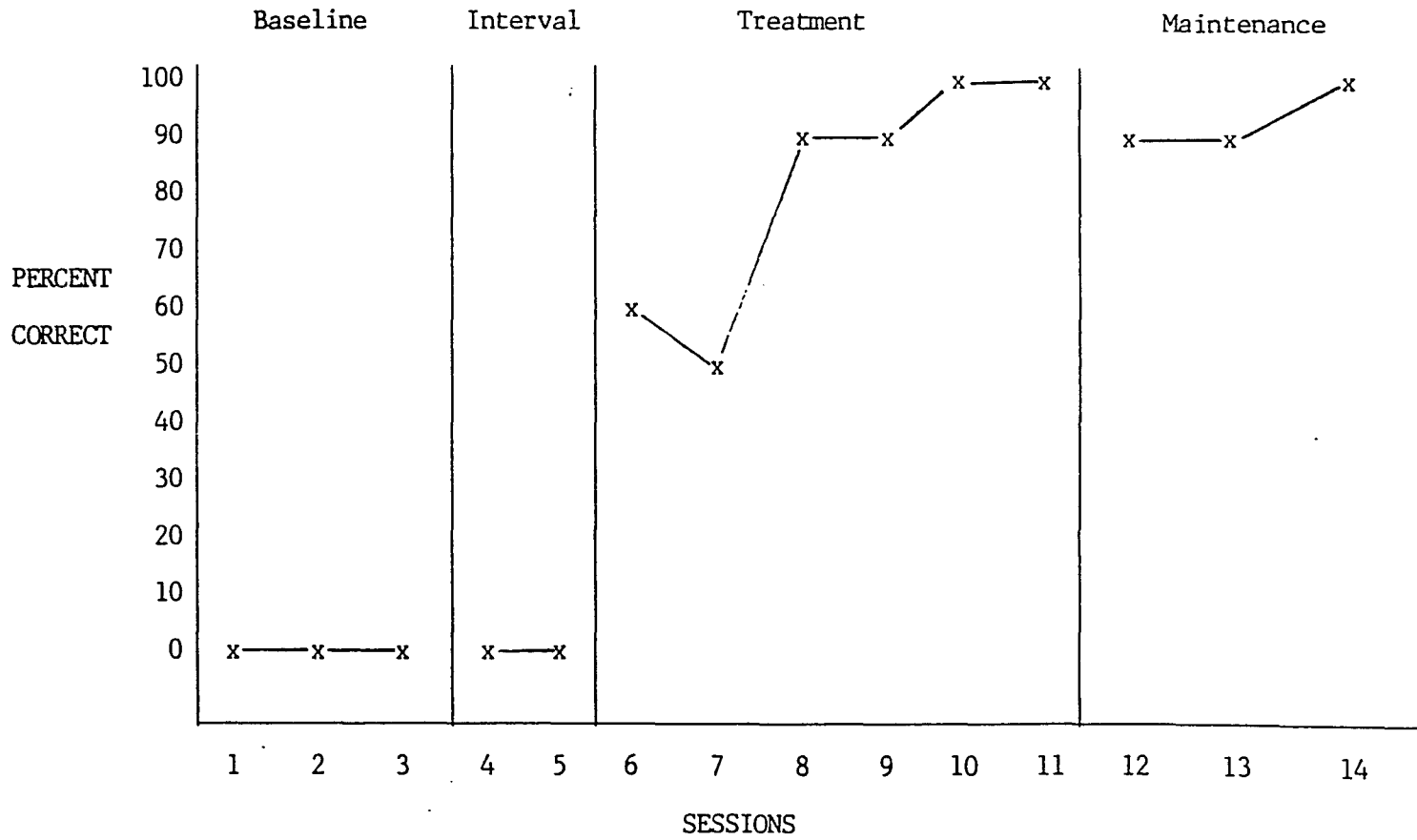


Figure 20. Subject 9 - The Effects of Treatment on Changing Fractions to Percents

As shown in Figure 21, the dependent variable was multiplication of mixed numbers. During baseline, S-9's scores reflect a stable profile of zero percent correct. During treatment, there is a dramatic drop in scores during session 6. Scores range from 40 percent correct to 100 percent correct. During maintenance, a negatively accelerated slope is evident.

SUBJECT 9

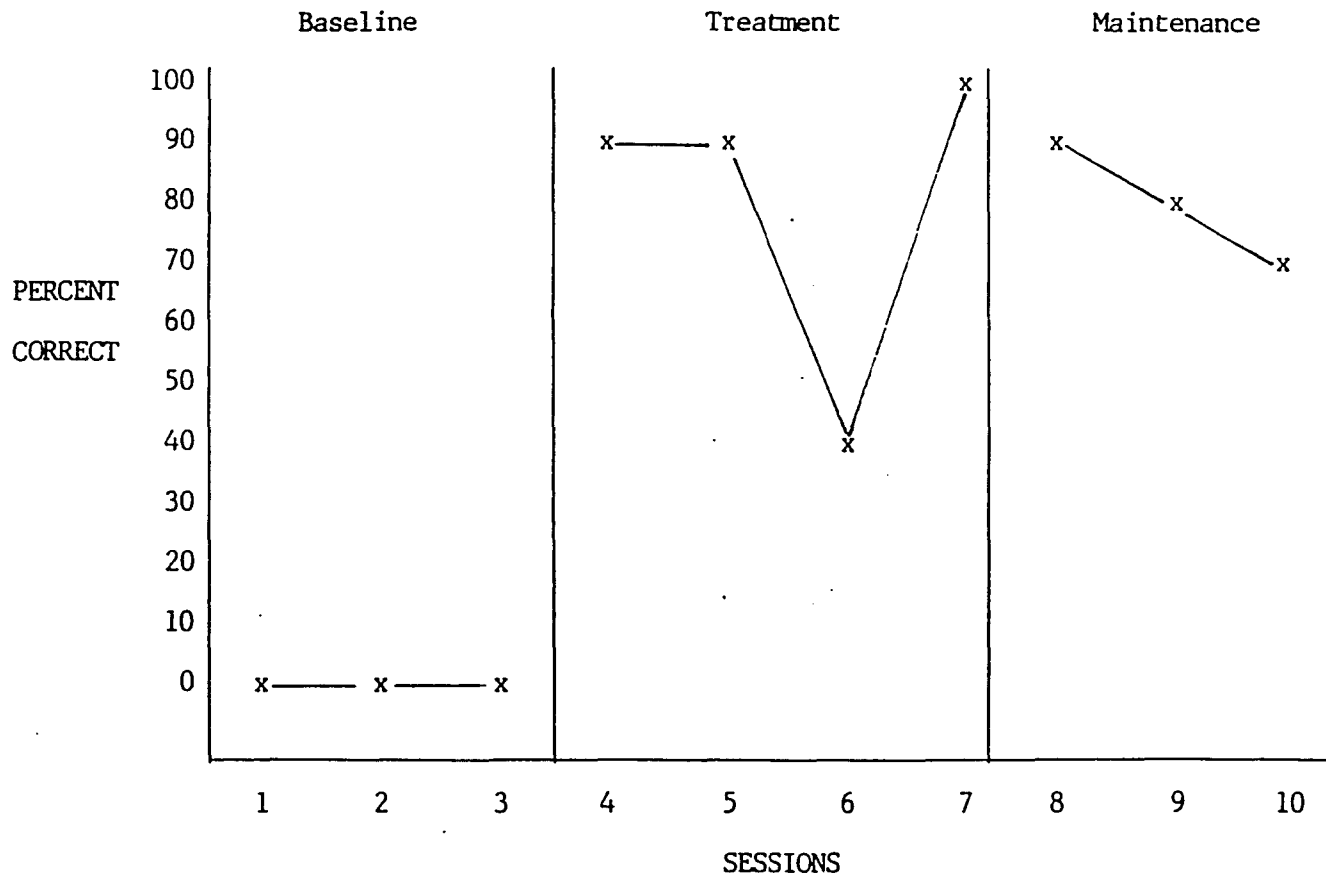


Figure 21. Subject 9 - The Effects of Treatment on the Multiplication of Mixed Numbers

Summary

The combined results of the nine subjects progressing through the conditions of baseline, interval probes, treatment and maintenance are presented in Table 1.

Table 1

Percentage Correct for Each Skill by Subject Under Each ConditionGroup A

	Baseline	Probes	Treatment	Maintenance
Subject 1				
addition of unlike fractions	0	---	93	100
Subject 2				
subtraction with regrouping	70	---	89	87
addition of like fractions	0	---	100	97
reducing fractions	0	---	87	27
Subject 3				
simple division	0	---	82	83
Subject 4				
addition of unlike fractions	0	---	98	100
fractions to percents	0	---	85	97
multiplication of fractions	0	---	83	87
Subject 5				
long division	0	---	98	97
addition of unlike fractions	0	---	90	97

Group B

Subject 6				
addition of unlike fractions	0	100	90	100
fractions to percents	26	---	78	96
multiplication of fractions	0	---	90	100
Subject 7				
addition of like fractions	0	0	100	87
addition of unlike fractions	0	---	87	96
reducing fractions	0	---	95	80
Subject 8				
fractions to percents	3	5	95	0
multiplication of fractions	0	---	93	97
division of fractions	0	---	97	100
Subject 9				
fractions to percents	0	0	82	93
multiplication of mixed numbers	0	---	80	80

CHAPTER V

CONCLUSIONS AND FUTURE RESEARCH IMPLICATIONS

The purpose of the present study was to determine the effects of curriculum-based assessment and instructional methods (CBA Model) on the performance of middle school students who are at-risk of failing mathematics. Current literature regarding assessment stresses a great need within education generally to link assessment more closely to curriculum and instruction. The CBA Model attempts to do this through the use of frequent testing probes, error analysis, and instruction based on results of the curriculum-based error analysis.

This study used a multiple baseline single subject time series design with staggered baselines to study changes in student performance. The study used 10 problem set probes which reflect on the graphs as a 10 percent drop if one problem is missed. The study was conducted in three phases. After administering the WRAT, the first phase involved a series of baseline probes of the mathematical skills of each subject. The second phase introduced treatment in the form of the CBA Model (assessment and instruction). The final phase, maintenance, use probes to determine if performances during the treatment phase were maintained after a period of 14-18 days subsequent to the cessation of treatment.

Middle school subjects were selected who were at-risk for failing mathematics and the study examined three questions. **Research Question 1: Does the use of CBA methods increase mathematical skills of middle school students who are at-risk of failing mathematics?**

To answer this question, the researcher measured skill performance during the baseline condition followed by intervention based on error analysis of the

information gathered in baseline. The results were measured over three to six treatment sessions.

All subjects' response data to baseline and treatment conditions with the exception of Subject 2 in the skill area of subtraction with regrouping and Subject 6 in the skill area of addition of unlike fractions showed a strong functionally positive relationship between the use of CBA methods and increased competency in selected mathematic skills.

Subject 2's baseline data reflected a positively accelerated slope beginning at 50 percent correct moving to 70 percent correct on the second probe and reaching 90 percent correct on the final baseline probe. Baseline was not established and skill acquisition improved progressively. Therefore, the researcher chose to move into treatment for the skill of subtraction with regrouping. She did demonstrate a percentage increase from an average of 70 percent correct during baseline to 89 percent correct during treatment.

Subject 6 was in the control group (Group B) and as such experienced a non-treatment interval period of two weeks between baseline and treatment conditions. In the identified skill area of addition of unlike fractions, the subject attained mastery (100 percent correct) of this skill during the non-treatment period. When queried regarding this situation, he responded by sharing with the researcher that his regular classroom teacher had taught him how to perform the skill accurately after he had experienced difficulty and requested assistance. However he was continued in treatment as as a reinforcer because this was not the first time this subject had been introduced to this particular skill. He attained 100 percent correct on two probes over a period of two weeks (February 17th through March 6th). On the first treatment session, he attained only 70 percent correct on this skill. He did

fluctuate during the remainder of the treatment sessions between 90 percent and 100 percent correct. Maintenance probes indicated retention of the skill at a stable rate of 100 percent correct across the three probes.

Subject 3's treatment performance is erratic beginning at 90 percent correct in session one, dropping to 70 percent correct in session two and reaching a low of 60 percent correct in session three, then progressing upward during the next three sessions. This subject had a routine electroencephalogram (EEG) on the morning of the third session, but this may have affected his performance as observed by his need to discuss the procedure in length.

In summary, the data reflect a strong positive relationship between the treatment effects of the CBA Model and increased skill attainment by subjects. The interventions appeared to facilitate skill development for subjects very quickly. This result supports Hargis' (1982) assumption that the majority of school related problems are curriculum induced rather than student based. Simply put, the origin of many academic deficiencies in mathematics appear extrinsic to the student and can be treated easily by finding the correct instructional level of the student using a CBA Model followed by appropriate interventions.

Research Question 2: Are differential effects in mathematical achievement observed between two heterogenous groups of middle school students when the Experimental Group instruction uses the CBA and traditional methods, while the Control Group has traditional instruction only?

To answer this question, the researcher followed procedures used with Group A to measure skill performance during the baseline condition followed by two interval probes (non-treatment) of Group B subjects over a period of two weeks.

Subjects 7 and 9 responses to baseline and non-treatment interval probes showed no change from zero correct performance in skill acquisition between the two conditions. Thus, without the intervention of the CBA method, no change was noted. Data for these subjects following treatment are presented below.

Subject 6's response to interval probes indicated mastery (100 percent correct) of the identified baseline skill of addition of unlike fractions even though baseline probes had shown zero percent correct performance. The unusual performance profile for S-6 was explained earlier.

Although Subject 6 made dramatic gains in the identified skill area without the benefit of the CBA method, it is interesting to note that his gains were made only after his classroom assignments had been performed incorrectly and he initiated a request for assistance from his teacher, which included analysis of his errors and intervention based on those errors. Thus, when the teacher employed the vital components of the CBA method, (i.e. present performance levels, identification of error analysis, and practice, the student learned the skills.

Subject 8's response at baseline on the identified skill of changing fractions to percents indicated an increase to 10 percent correct on the fourth baseline probe which followed three probes at zero percent correct. In analyzing the data, it was found that the subject was responding correctly to the equivalency of one-half to 50 percent. This is a well known equivalency and for this subject required no computation to determine the answer. The same profile and same equivalency appeared in the interval period probes with the first interval probe at zero percent correct moving to 10 percent correct on the second probe. As a result, the researcher concluded that this slight increase in performance was the result of an

artifact of the assessment process as opposed to a real achievement gain of a new skill on the part of the subject.

In summary, the data for research question 2 reflect a consistent pattern. Three of the four subjects' interval data remained virtually unchanged from the baseline condition and all showed skill acquisition with treatment. However, one subject experienced a dramatic gain in the area of the identified skill during the interval phase. According to the subject, this was the direct result of teacher intervention. All five subjects in the treatment condition experienced gains in their identified skill areas. Use of the CBA Model made a difference in skill acquisition for subjects in their identified skill areas. This result supports the CBA premise that by controlling task difficulty, focusing on entry skills, adapting tasks, and providing intervention to match the student's skill levels, students can be successful (Gickling & Thompson, 1985).

Research Question 3: Are gains achieved through the CBA Model maintained after the assessment and instruction are terminated for a period of 14 to 18 days?

To answer this question, the researcher measured skill performance during the treatment and maintenance conditions. Three maintenance probes were given for each identified skill.

The data for all subjects, with the exception of Subject 2 (reducing fractions) and Subject 8 (changing fractions to percents) indicated a strong positive functional relationship between treatment and maintenance of skills using CBA methods.

Subject 2's treatment data in the area of reducing fractions reflected three sessions with the first session at 60 percent correct followed by two sessions at 100 percent correct. Maintenance data reflected attainment levels of 30 percent correct

and 20 percent correct indicating a regression in the level of skill attainment. This was the last skill in a series of three introduced to the subject. She had only three treatment sessions to master the skill and had obvious difficulty during the first session. The researcher concludes that this subject needed more reinforcement of this skill in order to maintain it after the cessation of treatment.

Subject 8's treatment data in the area of changing fractions to percents reflected attainment at a level of 90 to 100 percent correct over a period of six sessions. However, maintenance data reflected a stable profile of zero percent correct across the three probes. Analysis of the data indicated that she made consistent errors in the procedure. Instead of dividing the denominator into the numerator, she reversed the procedure on every single problem. The researcher felt that she gained a skill during treatment which she never conceptually understood. Thus, the wrong procedure did not trigger any type of recognition that she was making incorrect responses to the problems presented.

In summary, the data reflected a strong positive functional relationship between skill acquisition during treatment and maintenance of that skill. All nine subjects maintained at least one of their identified skills at a mastery level. This result supports Galagan's (1985) premise that the use of the CBA Model provides systematic observational data on student performance, the degree of performance deviance, and allows for the formulation of alternative instructional methods that can effectively remediate the problem allowing the student to progress in the curriculum.

Problems Encountered in This Study

Several problematic issues were involved with this study. The first was the researcher's failure to extend baseline assessment, in three instances, to determine if

a more accurate representation of the skill levels of the subjects could have been obtained. This occurred with Subject 2 in the area of subtraction with regrouping, Subject 6 in the area of changing fractions to percents, and Subject 8 in the same skill area as Subject 6. In the case of Subject 8, an explanation was available related to an artifact of the assessment process. However, in the case of Subjects 2 and 6, this resulted in the inability to accurately determine whether or not the increased skill development could be attributed to treatment effects of the CBA Model. However, the strong positive functional relationship between baseline and treatment conditions across the other seven subjects and other skills for the subjects in question provide a high probability that the effects of treatment contributed positively to the skill gains of the subjects where the baselines are in question.

A second issue that needs consideration by future researchers was the decision to use 10-question sets for probes. There were times when this proved extremely tedious for subjects and required extra effort by the researcher to motivate subjects to continue. The use 5-question sets for probes should be sufficient.

Implications

Based on the results of the present study, CBA has implications for use within education at several levels.

1. CBA could be used very effectively as a pre-referral assessment and intervention strategy.
2. CBA could serve a very useful purpose in the formal evaluation process used to determine eligibility for special education.
3. CBA's role could be used to determine the in-class curricula match to the students' needs and assessments used.

4. IEP objectives could be developed from the CBA assessment and instructional outcomes be used to perform ongoing evaluation of the objectives.
5. CBA could assist in moving students toward re-entry into the regular mathematics curriculum if assessment is based on curricular content.

The composition of the study groups varied, which suggests that many students from ethnic and culturally diverse backgrounds when also identified as "at-risk" or in the "gray area" might be served using CBA methods as a very effective intervention strategy.

Future Research

- 1, Future research in the use of CBA methods should be completed with students in secondary and elementary settings. If used in a well-designed model, CBA may not only reduce referrals for evaluation for special education programming, but it could also be effective in preventing "curricular casualties".
2. The results of this study indicate that CBA is a very effective pull-out strategy for use with middle school students who are at-risk for failing mathematics. However, study is needed to determine if use of the CBA method can be implemented broadly in a cost effective manner. It certainly has implications for prevention of academic problems across at-risk populations, including the culturally diverse, economically disadvantaged, and mildly disabled. If these students were given assistance through use of the CBA Model when problems within the curriculum began to arise, it is possible that their need for long-term

supplementary programs would be greatly reduced. Several options in using the CBA Model should be studied:

- a. classroom teachers,
 - b. resource teachers,
 - c. teacher assistants working under the direction of a teacher,
 - d. roving regular classroom teachers,
 - e. consultant teachers, and
 - f. resource-type teachers providing blocked times for reviews of particular skills to which up to 5 students could attend. For example, this teacher would offer a 30 minute block on Tuesday at 10:00 a.m. to review addition with regrouping. Classroom teachers would have been asked to have pre-registered students who would benefit from this review.
3. Data-based research is needed in how to involve regular classroom teachers through training and implementation of the CBA Model as a routine tool to be used with low-achieving students. To facilitate this process, creative scheduling or financing could be looked at by the individual school sites. For example, site administrators may choose to spend a designated amount of their discretionary funds to pay teachers to attend training.
 4. The use of CBA methods by child study teams as a pre-referral assessment and intervention strategy may serve to significantly reduce referrals for evaluation for special education programming.

- 5 University teacher preparation programs should explore the efficacy of incorporating the CBA method into competencies required of all teacher candidates.
6. Since annual standardized testing is mandated in most states, it would be interesting to do a long-term (must have time to replicate the study with the controls) study on the results of the CBA Model compared to a control group of elementary students in areas of reading and mathematics to determine long-term effects on achievement. The standardized test results would serve as pre- post measures and comparison of assessed achievement levels.

APPENDIX A

Human Subject Committee



1690 N. Warren (Bldg. 526B)
 Tucson, Arizona 85724
 (602) 626-6721 or 626-7575

Gail Bornfield, M.A.
 c/o Candace L. Bos, Ph.D.
 Department of Special Education & Rehabilitation
 Education 429
 Main Campus

May 20, 1992

RE: HSC A92.71 EFFECTS OF A CBA MODEL ON MIDDLE SCHOOL STUDENTS IN
 MATH ACHIEVEMENT

Dear Ms. Bornfield:

We received your 24 April 1992 project approval form and 19 May 1992 consent forms for your above referenced project. The procedures to be followed in this study pose no more than minimal risk to participating subjects. Regulations issued by the U.S. Department of Health and Human Services [45 CFR Part 46.110(b)] authorize approval of this type project through the expedited review procedures, with the condition(s) that subjects' anonymity be maintained. Although full Committee review is not required, a brief summary of the project procedures is submitted to the Committee for their endorsement and/or comment, if any, after administrative approval is granted. This project is approved effective 20 May 1992 for a period of one year.

The Human Subjects Committee (Institutional Review Board) of the University of Arizona has a current assurance of compliance, number M-1233, which is on file with the Department of Health and Human Services and covers this activity.

Approval is granted with the understanding that no further changes or additions will be made either to the procedures followed or to the consent form(s) used (copies of which we have on file) without the knowledge and approval of the Human Subjects Committee and your College or Departmental Review Committee. Any research related physical or psychological harm to any subject must also be reported to each committee.

A university policy requires that all signed subject consent forms be kept in a permanent file in an area designated for that purpose by the Department Head or comparable authority. This will assure their accessibility in the event that university officials require the information and the principal investigator is unavailable for some reason.

Sincerely yours,

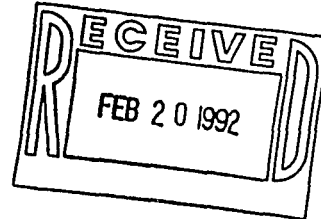
William F. Denny
 William F. Denny, M.D.
 Chairman, Human Subjects Committee

WFD:rs

cc: Departmental/College Review Committee

TUCSON UNIFIED SCHOOL DISTRICT

P.O. BOX 40400
1010 EAST TENTH STREET
TUCSON, ARIZONA 85717-0400



February 19, 1992

Gail Bornfield
7441 N. Rae Ave.
Tucson, AZ 85741

Dear Ms. Bornfield:

We are pleased to inform you that your request to do research in the Tucson Unified School District has been approved by: Carol Fillenberg, Assistant Superintendent - Middle School Region.

PROJECT TITLE: Effects of CBA Model on Middle School Students in Math Achievement

REFERENCE NUMBER: 1457

It will be your responsibility to contact the administrator(s) of the site(s) and/or departments participating in your study to secure their approval. They are responsible for accepting or rejecting your offer to participate. Therefore, you must:

1. show them a copy of this letter, and
2. obtain the signature of (the) participating site/department administrator on the Request To Do Research In A TUSD School Or Department form (see attached), and
3. return the completed form(s) to the Department of Testing and External Research.

Please remember that building principals have administrative responsibility and control of conduct for your study in their area.

Finally, please provide this Department with one copy of the final report of the completed study.

Sincerely,

Pamela Brown Clarridge, PhD
Assessment Specialist
Testing and External Research

PBC/clz
Attachment

REQUEST TO DO RESEARCH IN A TUSD SCHOOL OR DEPARTMENT

Please indicate whether you approve or disapprove this study being conducted in your school or department and return to Testing and External Research.

Gail Bonfield is (circle one)

GRANTED NOT GRANTED permission to conduct a research study entitled

Effects of a CBA Model on Middle School
(name of study)

Students in Math Achievement

in Vail Middle School
(school or department)

from Feb / 14 / 1991 to March / 31 / 1992.
month day year month day year

Kelly Langford
Principal or Department Head Signature

2/20/92
Date

REFERENCES

- American Psychological Association, American Educational Research Association, and National Council on Measurement in Education (1976). Standards for Educational and Psychological Tests. Washington D.C.: American Psychological Association.
- Armbruster, B., Stevens, R., & Rosenshine, B. (1977). Analyzing content coverage and emphasis: A study of three curricula and two tests (Report No. 26). Urbana - Champaign: Center for the Study of Reading, University of Illinois.
- Bailey, S. & Mosher, E. (1968). ESEA: The Office of Education Administrators and Law. Syracuse: Syracuse University Press.
- Barlow, D. & Hersen, M. (1984). Single Case Experimental Designs Strategies for Studying Behavior Change. New York: Pergamon Press
- Bean, R. & Lane, S. (1990). Implementing curriculum-based measures of reading in an adult literacy program. Remedial and Special Education (RASE), 11, 5, 39-46.
- Bruner, J. (1966). Studies in Cognitive Growth. New York: John Wiley & Sons, Inc.
- Capps, L. & Cox, L. (1991). Improving the mathematics in our schools. Focus on Exceptional Children, 23, 1-8.
- Cawley, J. (1984). Developmental Teaching of Mathematics for the Learning Disabled. Rockville, MD: Aspen Publications.
- Cawley, J. (1990). An examination of reading performance of students with mild educational handicaps or learning disabilities. Journal of Learning Disabilities, 23, 3, 5, 284-290.
- Chapman, P. (1980). Schools as sorters: Lewis M. Terman and the intelligence testing movement, 1890-1930. Doctoral dissertation, Stanford University (University Microfilms No. 80-11, 615).
- Chittenden, E. (1991). Authentic assessment, evaluation, and documentation of student performance. In V. Perrone (E.) Expanding Student Assessment (pp 22-31).
- Choate, J. Bennet, T., Enright, B., Miller, L., Poteet, J., & Rakes, T. (1987). Assessing and Programming Basic Curriculum Skills. Boston: Allyn and Bacon.
- Cockcroft, W. (1986). Mathematics Counts. London: Her Majesty's Stationery Office.

- Colburn, W. (1825) s in arithmetic. In J. Bidwell & R. Clason (Eds: the History of Mathematics Education (pp.14-38. .: National Council of Teachers of Mathematics.
- Collison, J. (1 performance assessment to determine mathematics Arithmetic Teacher, , 40-48.
- Conoley, J. & Kram The Tenth Mental Measurements Yearbook. Lincoln: Unaska Press.
- Dailey, J. (1953) my general classification test, first civilian edK. Buros (Ed.), The Fourth Mental Measurementsland Park, NJ: Gryphon Press.
- Deno, S. (1985). based measurement: The emerging alternative. ildren, S2, 3, 219-232.
- Dienes, Z. (1971) Mathematics. New York: Hutchinson Educational
- Dyer, H. (1977). testing: How mean is the median? Measurement 1-10.
- Engelhardt, T. (19 f children's computational errors: A qualitative ish Journal of Educational Psychology, 47, 149-154.
- Flanders, J. (1987) he content in mathematics textbooks is new? Arithm, 18-23.
- Frank, A. & Gerk Case studies in curriculum-based measurement. Training in Mental Retardation, 25, 2, 113-119.
- Frederiksen, N. (1 test bias:; Influences of testing on teaching merican Psychologist, 39, 3, 193-202.
- Fuchs, L. (1989). and validity of skills analysis within curriculum-ba: Diagnostique, 14, 4, 203-221.
- Gable, R. (1990). l measurement in oral reading: Linking assessment a Preventing School Failure, 35, 1, 37-42.
- Galagan, G. (1985) nal testing: Turn out the lights, the party is ove children, S2, 3, 288-299.
- Gallagher, J. &). Learning Theory of Piaget and Inhelder. Mbooks-Cole.

- Gickling, E. (1992, April). Curriculum-based assessment: A naturalistic guide to reading and mathematics instruction. Paper presented at the Annual Meeting of the Council for Exceptional Children, Baltimore, MD.
- Gickling, E. & Havertape, J. (1981). Curriculum-based assessment. In J. A. Tucker (Ed.), Non-test Based Assessment. Minneapolis: The National School Psychology Inservice Network, University of Minnesota.
- Gickling, E., Shane, R., & Croskery, K. (1989). Developing mathematics skills in low-achieving high school students through curriculum-based assessment. School Psychology Review, 18, 3, 344-353.
- Gickling, E. & Thompson, V. (1985). A personal view of curriculum-based assessment. Exceptional Children, 52, 3, 205-218.
- Gifford, B. (1990). From Gatekeeper to Gateway: Report on the National Commission on Testing and Public Policy. Chestnut Hill, MA: Boston College Press.
- Grant, W. & Eiden, L. (1982). Digest of Educational Statistics. Washington D.C.: U.S. Government Printing Office.
- Haertel, E. & Calfee, R. (1983). School achievement: Thinking about what to test. Journal of Educational Measurement, 20, 119-132.
- Haladyna, T., Haas, N., & Nolen, S. (1989). Test score pollution (Tech Rep. No. 1). Phoenix: Arizona State University West Campus.
- Haladyna, T., Nolen, S., & Haas, N. (1991). Raising standardized achievement test scores and the origin of test score pollution. Educational Researcher, 20, 5, 2-8.
- Hamill, D., Leigh, J., McNutt, G., & Larson, S. (1981). A new definition of learning disabilities. Learning Disability Quarterly, 4, 336-343.
- Haney, W. (1984). Testing reasoning and reasoning about testing. Review of Educational Research 54, 4, 597-654.
- Haney, W. (1991). We must take care fitting assessments to functions. In V. Perrone (Ed.), Expanding Student Assessment. (pp. 143-163).
- Hargis, C. (1982). Teaching Reading to Handicapped Children. Denver: Love Publishing.
- Heshusius, L. (1991). Curriculum-based assessment and direct instruction: Critical reflections on fundamental assumptions. Exceptional Children, 57, 4, 315-328.

- Jenkins, J. & Pany, D. (1978). Standardized achievement tests: How useful for special education? Exceptional Children, 44, 448-453.
- Johnson, W. & Packer, A. (1987). Workforce 2000: Work and Workers for the Twenty-first Century. Indianapolis: Hudson Institute.
- Kirk, S. & Chalfant, J. (1984). Academic and Developmental Learning Disabilities. Denver: Love Publishing.
- Kirk, S. & Gallagher, J. (1986). Educating Exceptional Children. Boston: Houghton Mifflin Company.
- Kirst, M. (1991). Interview on assessment issues with Lorrie Sheppard. Educational Research, 20, 2, 21-23.
- Lapointe, A., Mead, N., & Phillips, G. (1989). A World of Differences: A International Assessment of Science and Mathematics. Princeton: Educational Testing Service.
- Macht, J. (1990). Managing Classroom Behavior: An Ecological Approach to Academic and Social Learning. New York: Longman.
- Mathematical Sciences Education Board, National Research Council (1990). Reshaping School Mathematics: A Philosophy and Framework for Curriculum. Washington, D.C.: National Academy Press.
- Mathematics Framework for California Public Schools: Kindergarten through Grade Twelve. Version of 20 August 1991.
- Mather, N. & Healey, W. (1990). Depositing aptitude - achievement discrepancy as the imperial criterion for learning disabilities. Learning Disabilities: A Multidisciplinary Journal, 1, 2, 40-48.
- McKnight, C., Crosswhite, F., Dossey, J., Kifer, E., Swafford, J., Travers, K., & Cooney, T. (1987). The Underachieving Curriculum: Assessing U.S. School Mathematics from an International Perspective
- Mirkin, P. (1980). Conclusions. In J. Ysseldyke & M. Thurbow (Ed.), The Special Education Assessment and Decision-Making Process: Seven Case Studies. Minneapolis: University Minnesota Institute for Research on Learning Disabilities.
- Mitchell, R. (1989). What is "authentic assessment"? Portfolio, 1, 5, 5-6.
- Mullis, I., Dossey, J., Owen, & Phillips, G. (1991). The State of Mathematics Achievement: Executive Summary. Washington, D.C.: Educational Testing Service.

- Myers, M. & Burton, G. (1989). Yes, you can....plan appropriate instruction for learning disabled students. Arithmetic Teacher, 36, 7-9.
- National Council of Teachers of Mathematics (1989). Curriculum and Evaluation Standards for School Mathematics. Reston: A Report to the Nation on the Future of Mathematics Education. Washington, D.C.: National Academy Press.
- National Research Council (1989). Everybody Counts: A Report to the Nation on the Future of Mathematics Education. Washington, D.C.: National Academy Press.
- Neill, M. & Medina, N. (1989). Standardized testing: Harmful to educational health. Phi Delta Kappas, 70, 9, 688-697.
- Piaget, J. (1952c). The Origins of Intelligence in Children. New York: International Universities Press.
- Piaget, J. (1936b). The Psychology of Intelligence. Patterson, N.J.: Littlefield & Adams.
- Potter, M. & Wamre, H. (1990). Curriculum-based measurement and developmental reading models: Opportunities for cross validation. Exceptional Children, S7, 1, 16-25.
- Price, J. & Gawronski, J. (Eds.). (1991). Changing School Mathematics: A Responsive Process. Reston: National Council of Teachers of Mathematics.
- Reisner, E., Alkin, M., Bornch, R., Linn, R., & Millman, J. (1982). Assessment of the Title I Evaluation and Reporting System. Washington D.C.: U.S. Department of Education.
- Ridgway, J. (1988). Assessing Mathematical Attainment. Windsor: NFER-Nelson Publishing Company, Ltd.
- Rosenfield, S. & Shinn, M. (Eds.). (1989). Mini-series on CBA [Special Issue]. School Psychology Review, 18, 3, 297-370.
- Russell, J. (1986). Characteristics of time-series data collected through curriculum-based reading measurement. Diagnostique, 12, 1, 3-15.
- Salmon-Cox, L. (1981). Teachers and standardized achievement tests: What's really happening? Phi Delta Kappas. 62. 631-634.
- Salvia, J. & Ysseldyke, J. (1981). Assessment in Special and Remedial Education (2nd ed.). Boston: Houghton Mifflin.

- Shinn, M., Knutson, N., & Rosenfield, S. (1989). Curriculum-based assessment: A comparison of models. School Psychology Review, 18, 3, 299-316.
- Smith, M. (1991). Put to the test: The effects of external testing on teachers. Educational Researcher, 20, 5, 8-12.
- Steen, L. (1986). A time of transition: Mathematics for the middle grades. In R. Lodhulz (Ed.) A Change in Emphasis. Parkway, Mo: Parkway School District.
- Stenmark, J. (1989). Assessment Alternatives in Mathematics: An Overview of Assessment Techniques that Promote Learning. Berkeley: Lawrence Hall of Science.
- Stevenson, H., Lee, S., & Stigler, J. (1986). Mathematics achievement of Chinese, Japanese, and American Children. Science, 231, 693-699.
- Stigler, J. & Perry, M. (1988). Cross-cultural studies of mathematics teaching and learning: Recent findings and new directions. In D. Grouws, T. Cooney, & D. Jones (Eds.). Perspectives on Research on Effective Mathematics Teaching. Reston: National Council of Teachers of Mathematics.
- Tawney, J. & Gast, D. (1984). Single Subject Research in Special Education. Columbus: Charles E. Merrill Publishing Company.
- Thurlow, M. & Ysseldyke, J. (1982). Instructional planning: Information collected by school psychologists vs. information considered useful by teachers. Journal of School Psychology, 20, 1, 3-10.
- Truesdell, L. (1987). Curriculum-based assessment: Process and application. Journal of Reading, Writing and Learning Disabilities International, 3, 4, 281-289.
- Tucker, J. (1985). Curriculum-based assessment: An introduction. Exceptional Children, 52, 3, 199-204.
- Tucker, J. (1991, July). Assessment and evaluation for specially-designed instruction. Paper presented at a "Principal's Academy", Minneapolis, MN.
- U.S. Department of Education (1991). Thirteenth Annual Report to Congress in the Implementation of the Education for the Handicapped Act. Washington, D.C.: U.S. Government Printing Office.
- Wadsworth, B. (1979). Piaget's Theory of Cognitive Development. New York: Longman, Inc.
- Wadsworth, B. (1984). Piaget's Theory of Cognitive and Affective Development. New York: Longman, Inc.

- Wesson, C. (1989). The discriminate validity of curriculum-based measures for establishing reading groups. Reading Research and Instruction, 29, 1, 23-32.
- West, T. (1971). Diagnosing pupil errors: Looking for patterns. The Arithmetic Teacher, , 467-469.
- Willoughby, S. (1990). Mathematics Education for a Changing World. Alexandria: Association for Supervision and Curriculum Development.
- Zessoules, R. & Gardner, H. (1991). Authentic assessment: Beyond the buzzword and into the classroom. In V. Perrone (Ed.), Expanding Student Assessment (pp. 47-71).