ISSN 0883-9530

JOURNAL FOR RESEARCH IN MATHEMATICS EDUCATION

MONOGRAPH NUMBER 2

LEARNING TO ADD AND SUBTRACT

National Council of Teachers of Mathematics

Printed in the United States of America

ACKNOWLEDGEMENTS

The set of studies reported in this monograph could not have been carried out without the help and cooperation of several persons. First, we thank the children who were willing to be tested, interviewed and observed. They were patient with us and our questions and in every way helpful. Second, we thank the teachers and staffs of Sandy Bay Infant School and Waimea Heights Primary School for their enthusiasm and support. Next, we are indebted to the staffs of the Center for Education at the University of Tasmania and St. John Fischer College for their hospitality and help during 1979-80 when the study was conducted.

In Wisconsin many project staff members helped clean, organize, and summarize the information after it had been collected. We are particularly indepted to Anne Buchanan, Jacob Evanson, and Martha Romberg for their conscientious work. For typing and retyping drafts of the manuscripts, we thank Dorothy Egener, Teri Frailey, Donna Mlsna, and Kay Schultz. We also must thank both Doug Grouws and Deborah Stewart for editing this volume with care.

Table of Contents

List of Tables

List of Figures

DIFFERENT WAYS CHILDREN LEARN TO ADD AND SUBTRACT

Thomas A. Romberg
Wisconsin Center for Education Research
University of Wisconsin-Madison

Kevin F. Collis
University of Tasmania

Edited by
Douglas A. Grouws
University of Missouri-Columbia

In 1979 the Research Committee of the Graduate School at the University of Wisconsin-Madison, the Wisconsin Center for Education Research, and the University of Tasmania jointly funded the principal investigators to carry out the series of studies relating children's cognitive capacity to their performance and to the strategies they used when working addition and subtraction problems. The Wisconsin Center for Education Research was supported in part by a grant from the National Institute of Education (Grant No. NIE-G-81-0009). The opinions expressed in this paper do not necessarily reflect the position, policy, or endorsement of the National Institute of Education.

ABSTRACT

This monograph summarizes the findings from five related studies carried out by the authors in Sandy Bay, Tasmania, Australia, in 1979-80. The overall purpose of the studies was to examine whether children in grades 1-3 who differed in cognitive capacity learned to add and subtract in different ways.

The first study was a cross-sectional survey designed to determine the memory capacity of a population of children. The second study was designed to portray performance differences on a variety of mathematically related developmental tasks for the same population of children. Data from these two studies were used to form groups of children who differed in cognitive capacity. Six groups were formed via cluster analysis, with memory capacity being the primary distinguishing characteristic.

The third, fourth, and fifth studies each used a sample of students from the six cluster groups across grades. The third study examined both the performance and the strategies these children used to solve a structured set of addition and subtraction word problems. The fourth study involved repeated assessment of the children's performance on items measuring objectives related to addition and subtraction. In the last study these children and their teachers were observed during classroom instruction in mathematics to see how addition and subtraction were taught and whether or not instruction was related to the children's cognitive capacity.

The results show that children's differences in capacity were reflected in their performance on both verbal and standard problems and in the strategies they used to solve problems. However, instruction did

not vary for these children within classrooms. The picture that emerges is one of children struggling to learn a variety of important concepts and skills. Some children were limited by their capacity to process information. Most were able to solve a variety of problems by using invented strategies, those that had not been taught. They dismissed or failed to see the value of the taught procedures in solving these problems. Finally, the capacity of children to process information, the procedures students invented to solve a variety of problems, and the way in which instruction was carried out in schools did not seem related to each other.

Chapter 1

INTRODUCTION

> For several centuries being able to find "one's sums and differences" has been considered one mark of a schooled person. Although today we may have expanded our expectations about what constitutes literacy, we still expect all children to efficiently carry out operations on whole numbers. Yet, in spite of these expectations about the skills of addition and substraction, there has been little consensus about how such skills develop. (Romberg, 1982, p. 1)

The basic question under investigation was, Do children who differ in cognitive-processing capacity learn to add and subtract differently? In raising this question, it was assumed that the evolution of children's performance on mathematical tasks (such as addition and subtraction) must be related both to their developing cognitive abilities and to related instruction they receive. To examine this question five related studies were conducted in Sandy Bay, Tasmania, Australia, in 1979-80. This monograph summarizes the findings from those studies.

The rationale for these studies is detailed in a conceptual paper (Romberg, Carpenter, & Moser, 1978). In that paper the authors describe how, for nearly a decade (1968-1976), the Studies in Mathematics project at the Wisconsin Center for Education Research had concentrated its efforts on the relationship between instructional processes, methods, and materials and the acquisition of mathematical skills associated with mathematical learning. The work in that project led to the development of a complete elementary mathematics program, <u>Developing Mathematical Processes</u> (DMP) (Romberg, Harvey, Moser, & Montgomery, 1974, 1975, 1976). Although DMP was based on empirical evidence and theories of learning, development, and instruction (see Romberg, 1977), a number of questions were raised as the program was being developed.

In particular, it became clear that a complete picture of mathematics instruction was lacking. What was needed was a characterization of the mathematical content to be learned, a description of children's cognitive capacity with mathematical material, and an identification of the features of classroom instruction such as how children perform on learning tasks, teacher presentations of mathematical material, pupil engagement, and teacher-pupil interactions during lessons. Thus, past work indicated that the interactions between content, cognitive capacity, and instruction needed to be carefully examined. The following sections describe the three areas involved in this investigation: content (addition and subtraction), cognitive capacity, and classroom instruction.

Addition and Subtraction

We chose as the vehicle for this investigation children's early work in addition and subtraction. There were several reasons for this choice. First, this area represents the first attempt that schools make to teach what might be recognized as formal mathematics. By this we mean learning to symbolically represent a problem situation (often via word problems), operate on the symbols, and interpret the result. Second, considerable work had been done at the Wisconsin Center for Education Research on logically analyzing the semantic-syntactic relationship for these mathematical skills as they apply at the early elementary school level (e.g., Carpenter & Moser, 1983; Moser, 1979).

Third, in the 1970s the staff at the Wisconsin Center for Education Research had developed instructional materials to teach addition and subtraction. However, children in classrooms using those materials were only moderately successful in learning to solve problems using those operations. Fourth, various researchers had identified several strategies young children use to solve elementary addition and subtraction problems (see Carpenter, Moser, & Romberg, 1982). Finally, a clinical observation schedule for assessing performance on some addition and subtraction tasks had been developed (Carpenter & Moser, 1979).

Word problems. To solve a typical addition and subtraction word problem, one first must understand its implied semantic meaning. Quantifying the element of the problem comes next (e.g., choosing a unit and counting how many). Then, the implied semantics of the problem must be expressed in the syntax of addition and subtraction. Next the child must be able to carry out the procedural (algorithmic) steps of adding and subtracting. Finally, the results of these operations must be expressed. Most children bring to such problems well developed counting procedures, some knowledge of numbers, and some understanding of physical operations on sets of objects such as "joining" and "separating." Thus, from this context researchers have a unique opportunity to examine variations in how children process information prior to, during, and after formal instruction when they attempt to solve word problems.

Semantics. Not all word problems involving addition and subtraction have the same semantic structure. In fact, most current work uses four broad classes of addition and subtraction problems:

Change, Combine, Compare, and Equalize (Carpenter & Moser, 1983). There are two basic types of change problems, both of which involve action. In change-join problems, there is an initial quantity and a direct or implied action that causes an increase in that quantity. For change-separate problems, a subset is removed from a given set. In both classes of problems, the change occurs over time. Within both the join and separate classes, there are three distinct types of problems depending upon which quantity is unknown (see Table 1). Both combine and compare problems involve static relationships for which there is no action. Combine problems involve the relationship existing among a particular set and its two disjoint subsets. Two problem types exist: the two subsets are given and one is asked to find the size of their union, or one of the subsets and the union are given and the solver is asked to find the size of the other subset. Compare problems involve the comparison of two distinct, disjoint sets. Because one set is compared to the other, it is possible to label one set the referent set and the other the compared set. The third entity in these problems is the difference, or the amount by which the larger set exceeds the other. In this class of problems, any one of the three entities could be the unknown--the difference, the referent set, or the compared set. The larger set can be either the referent set or the compared set. Thus, there exist six different types of compare problems.

The final class of problems, equalize problems, are a hybrid of compare and change problems. There is the same sort of action as found in the change problems, but it is based on the comparison of two disjoint sets. As in the compare problems, two disjoint sets are compared; then the question is posed, What could be done to one of the

Table 1
Semantic Classification of Word Problems
(Carpenter & Moser, 1983)

Join	Separate
Change	
1. Connie had 5 marbles. Jim gave her 8 more marbles. How many marbles does Connie have altogther?	2. Connie had 13 marbles. She gave 5 marbles to Jim. How many marbles does she have left?
3. Connie has 5 marbles. How many more marbles does she need to have 13 marbles altogether?	4. Connie had 13 marbles. She gave some to Jim. Now she has 8 marbles left. How many marbles did Connie give to Jim?
5. Connie had some marbles. Jim gave her 5 more marbles. Now she has 13 marbles. How many marbles did Connie have to start with?	6. Connie had some marbles. gave 5 to Jim. Now she has 8 marbles left. How many marbles did Connie have to start with?
Combine	
7. Connie has 5 red marbles and 8 blue marbles. How many marbles does she have?	8. Connie has 13 marbles. Five are red and the rest are blue. How many blue marbles does Connie have?
Compare	
9. Connie has 13 marbles. Jim has 5 marbles. How many more marbles does Connie have than Jim?	10. Connie has 13 marbles. Jim has 5 marbles. How many fewer marbles does Jim have than Connie?
11. Jim has 5 marbles. Connie has 8 more than Jim. How many marbles does Connie have?	12. Jim has five marbles. He has 8 fewer marbles than Connie. How many marbles does Connie have?
13. Connie has 13 marbles. She has 5 more marbles than Jim. How many marbles does Jim have?	14. Connie has 13 marbles. Jim has 5 fewer marbles than Connie. How many marbles does Jim have?
Equalize	
15. Connie has 13 marbles. Jim has 5 marbles. How many marbles does Jim have to win to have as many marbles as Connie?	16. Connie has 13 marbles. Jim has 5 marbles. How many marbles does Connie have to lose to have as many marbles as Jim?
17. Jim has 5 marbles. If he wins 8 marbles, he will have the same number of marbles as Connie. How many marbles does Connie have?	18. Jim has five marbles. If Connie loses 8 marbles, she will have the same number of marbles as Jim. How many marbles does Connie have?
19. Connie has 13 marbles. If Jim wins 5 marbles, he will have the same number of marbles as Connie. How many marbles does Jim have?	20. Connie has 13 marbles. If she loses 5 marbles she will have the same number of marbles as Jim. How many marbles does Jim have?

sets to make it equal to the other? If the action to be performed is on the smaller of the two sets, then it becomes an equalize-join problem. On the other hand, if the action to be performed is on the larger set, then an equalize-separate problem results. As with compare problems, the unknown can be varied to produce three distinct equalize problems of each type.

To build the connection between semantic forms and relevant symbolism, one form is usually used as a model to introduce the symbolism. Because there are many semantic forms for which the same symbolic sentence is appropriate, the pedagogical problem is how to relate the symbolism to all the semantic problems. Traditionally, the symbolism has been taught independently of word problems. The symbolic procedures were taught, and some word problems were assigned so that students could apply their symbolic procedures. No serious consideration was given to the semantic structure of the problems. In fact, it is now clear that in many texts only a few of the semantic forms are ever included (see DeCorte, Verschaffel, Janssens, & Joillet, 1984). It is no surprise, then, that students have found little connection between different types of problems and the symbolic procedures they are taught (e.g., Vergnaud, 1982).

Development of instructional materials. During the early 1970s, the staff of the Wisconsin Center for Education Research produced the DMP curriculum for grades K-6 (Romberg et al., 1974, 1975, 1976). In creating this program the problem of connecting word problems and symbolic procedures had been recognized. For addition and subtraction, it was decided to use one semantic context to introduce and to give meaning to the symbolism and then to relate the symbolism to other semantic situations. In the initial version of DMP, equalizing was used. This context proved to be difficult for both teachers and students when they examined other semantic forms. A revised set of materials was later developed in which part-part-whole was used as the basic context for initial instruction (Kouba & Moser, 1979, 1980).

Strategies for solving word problems. In order to solve the variety of addition and subtraction word problems children use numerous strategies (see Carpenter et al., 1982; Carpenter & Moser, 1983). For addition and subtraction three basic strategy levels have been identified: strategies based on direct modeling with fingers or physical objects, strategies based on the use of counting sequences, and strategies based on recalled number facts. In addition, in the most basic strategy (counting all with models), children use physical objects or fingers to represent each of the addends, and then the union of the two sets is counted.

There are three distinct strategies involving counting sequences for addition problems. In the most elementary strategy, the counting sequence begins with one and continues until the answer is reached. This strategy is similar to the counting all with models strategy except that children do not use physical objects or fingers to represent the addends. However, this strategy and the two following counting strategies require some method of keeping track of the number of counting steps that represent the second addend in order to know when to stop counting. Most children use their fingers to keep track of the number of counts, but a substantial number give no evidence of any physical action accompanying their counting. When fingers are used, they appear to play a very different role than in the direct modeling strategy. In this case, the fingers do not represent the second addend per se, but are used to keep track of the number of steps incremented in the counting sequence. When using fingers, children often do not appear to have to count their fingers, but can immediately tell when they have put up a certain number of fingers.

The other two counting strategies are more efficient and imply a less mechanical application of counting. In applying these strategies, a child recognizes that it is not necessary to reconstruct the entire counting sequence. In the counting on from first strategy, a child begins counting forward with the first addend in the problem. The counting on from larger strategy is identical except that the child begins counting forward with the larger of the two addends.

Although learning of basic number facts appears to occur over a protracted span of time, most children ultimately solve simple addition and subtraction problems by recall of number combinations rather than by using counting or modeling strategies.

For subtraction a number of distinct classes of subtraction strategies have been observed at the direct modeling and counting levels. One of the basic strategies involves a subtractive action. In this case, the larger quantity in the subtraction is initially represented and the smaller quantity is subsequently removed from it. When concrete objects are used, the strategy is called separating from. The child constructs the larger given set and then takes away or separates, one at a time, a number of objects equal to the number given in the problem. Counting the set of remaining objects yields the answer. There is also a parallel strategy based on counting called counting down from. A child initiates a backward counting sequence beginning with the given larger number. The backward counting sequence contains as many counting number words as the given smaller number. The last number uttered in the counting sequence is the answer.

The separating to strategy is similar to the separating from strategy except that elements are removed from the larger set until the number of objects remaining is equal to the smaller number given in the problem. Counting the number of objects removed provides the answer. Similarly, the backward counting sequence in the counting down to strategy continues until the smaller number is reached and the number of words in the counting sequence is the solution of the problem.

The third pair of strategies involves an additive action. In an additive solution, the child starts with the smaller quantity and constructs the larger. With concrete objects (adding on), the child sets out a number of objects equal to the smaller given number (an addend). The child then adds objects to that set one at a time until the new collection is equal to the larger given number. Counting the number of objects added on gives the answer. In the parallel counting strategy (counting up from given), a child initiates a forward counting strategy beginning with the smaller given number. The sequence ends with the larger given number. Again, by keeping track of the number of counting words uttered in the sequence, the child determines the answer.

The fourth basic strategy is called matching. Matching is only feasible when concrete objects are available. The child puts out two sets of cubes, each set standing for one of the given numbers. The sets are then matched one-to-one. Counting the unmatched cubes gives the answer. A fifth strategy (choice) involves a combination of counting down from and counting up from given, depending on which is the most efficient. In this case, a child decides which strategy requires the fewest number of counts and solves the problem accordingly.

As with addition, modeling and counting strategies eventually give way to the use of recalled number facts or derived facts. Children's explanations of their solutions suggest that the number combinations they are calling upon are often addition combinations. Of significant interest to researchers and teachers must be the link, if any, between the logical analysis of the semantic forms of problems and the strategies children actually use to solve such problems.

Summary. Several points about addition and subtraction were noted at the outset of this project. First, word problems that can be solved by addition and subtraction differ in semantic form. Second, children have developed "primitive" or "child" strategies to solve addition and subtraction word problems prior to school learning experiences or at least prior to formal instruction on consolidated "efficient" methods of solution. Third, differences in the semantic form of word problems elicit different strategies from children. Finally, a logical analysis of the operations and related word problems seems to imply that these initial strategies should become more and more inefficient as the number of semantic forms is increased, or the numbers become larger, or the number of steps necessary for solution increases.

It therefore seems to be a reasonable goal of mathematics instruction to teach more formal, generalizable algorithmic procedures for solving the variety of addition and subtraction word problems. However, little is known about several aspects of this process and a number of questions arise. How will learning of the mathematical procedures be affected by the number, type, and success of the preexisting problem-solving strategies an individual child possesses? How do children who are successful problem solvers combine their existing strategies with formal mathematical modes of presentation? How should teachers adapt instruction to take account of a child's demonstrated level of functioning in this area? Raising these questions leads to a consideration of the relationship of general cognitive functioning to performance by children on addition and subtraction word problems.

Cognitive Capacity

Concern for cognitive abilities is well entrenched in research in mathematics education. The approach adopted in this project was based on claims from two sources: differential abilities and cognitive development.

Differential abilities. Based on the extensive work of a number of educational psychologists in the Thurstone tradition of distinct mental abilities, we decided to attempt to measure the ability of students to solve addition and subtraction problems. The procedure in this approach is to use test scores and psychometric analyses to identify differential abilities, traits, aptitudes, styles, and so forth. For example, such characteristics as intelligence, rate of learning, field independence/dependence, and spatial ability have been identified, and samples of students have been ordered from high to low on those traits. These traits are assumed to be fixed, stable characteristics, largely biological in origin, which describe intellectual differences between individuals in the same way as height, weight, stature, and so forth describe physical characteristics. Although we did not utilize tests developed from this perspective, we used the psychometric strategy of administering to each student a number of tests, scoring the tests, relating the scores, and classifying students based on their scores.

Our initial task was to find and administer measures of cognitive functioning that appeared logically related to the learning of mathematical material. However, we decided to use a battery of tests that seemed to be related to the children's level of cognitive development. Only instruments that could be shown prima facie to contain tasks related to early mathematical learning such as number conservation and counting were selected.

Cognitive Development. We chose the measures to be used in the study from work in cognitive development. This perspective is based on the notion that individuals adaptively interact with the environment and gradually evolve intellectual processes through discontinuous stages. Rather than being fixed, differences between individuals are viewed as a function of growth. Children in the primary grades, for example, usually are at a concrete operations stage, think in terms of themselves (are egocentric), and think of concrete referents near at hand. Hence, they should not be expected to reason about hypothetical, external situations.

The choice of tests from this perspective grew out of work on children's understanding of mathematics. This research gained impetus following the failure of the "new math" programs to live up to early expectations. Psychologists interested in mathematics learning began to investigate developmental and learning phenomena by using elementary mathematical material (e.g., Collis, 1975). These investigators used the clinical interview as a technique for studying the mathematical concepts that children had formed. Much of the work was stimulated by the notions of Jean Piaget (Inhelder & Piaget, 1958). Later interest was related to the work on memory capacity by Pascual-Leone (1976) and Case (1972). This view of cognitive functioning enabled psychologists to turn from the mere <u>description</u> of stages of development of mathematical thinking to an <u>explanation</u> of the phenomena that kept appearing in their work with individual children.

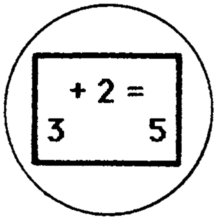
This evolution can be traced through the work of Collis (1971, 1974a, 1974b, 1975, 1976, 1978, 1980a, 1980b, 1982; Collis & Biggs, 1979; Biggs & Collis, 1982). The earlier papers use mathematical items to describe and, to some extent, to modify Piaget's stage theory (Inhelder & Piaget, 1958). The later papers, after about 1976, provide tentative explanations of the developmental phenomena found earlier in terms of Case's information-processing theory (Case, 1975). The most recent papers (e.g., Biggs & Collis, 1982) describe an intellectual skills model which, although it allows for the stage phenomenon, places the emphasis on the increasing complexity of responses within a given stage.

At the time this project began, a number of theorists were in various stages of refining and generalizing theoretical systems that included both structural and process components and that were of significance in relation to a broad range of developmental tasks. Since then most theorists have published their theoretical positions (e.g., Case, 1985; Fischer, 1980; Halford, 1980; Klahr, 1984; Pascual-Leone, 1984; Seigler, 1981; Sternberg, 1984). The investigators in this project selected the Case model for two reasons. First, it seemed at that time to be the most applicable to the content area and the methodology that we envisaged using. Second, it had tests available that could be utilized in the project.

M-Space. Central to Case's theory and crucial to this project is the concept of the short-term memory capacity (M-space). This basic variable we believed was central to a child's ability to process the mathematical material presented. Thus, the first set of tests we used was to measure M-space. The M-space construct has a long history in psychological theory, going back as far as Baldwin's conception of attention span (Baldwin, 1895). Basically it refers to the number of mental elements that an individual can attend to at any one time. In this project we have adopted Case's proposal (Case, 1985) that the overall processing space available is constant and is shared between two mental activities, the execution of an ongoing mental operation and the retention and retrieval of the product of such an operation. If the available processing space is exceeded, the individual exhibits symptoms of cognitive overload and is unable to solve the given problem. An early study by Collis (1973) elicited this phenomenon in relation to mathematical exercises. A simple example quoted by Collis in a later publication (Collis, 1980b) may help to explain cognitive overload.

> A child at the early concrete operational stage (circa 6 or 7 years) is asked to find the value of the statement 3+2+4; a typical interview goes as follows:
>
> Tester: What number does 3+2+4 equal?
> Child: 3+2=5 (pause) what was the other number?
> Tester: I said, "What number does 3+2+4 equal?"
> Child: Ah yes. Now, 3 plus (pause) what is the sum again? (p. 87)

What appears to be happening may be explained by using a diagram. Let us suppose that, in the diagram below, the rectangle represents the space available for processing data. At the early concrete operations stage it can be demonstrated that the space is sufficient to take in two elements and one operation and to perform the necessary calculation meaningfully (Collis, 1975). The processing space is, however, fully occupied. If one now attempts to introduce another operation and



element, the space available is exceeded and overflow results; part of the data necessary for a successful outcome is forced out of consideration. As the subject realizes the situation and retrieves one piece of data, another piece is forced out of the space and so on. Hence, in these circumstances the child never has all the information needed to solve the problem in the working space at the same time in order to obtain a satisfactory solution.

Development level. Although the M-space construct appeared basic to our investigation, it was also obvious that this could not be the only measure we should make because it was clear from the correlational data in the research literature that other influences must also be at work. Moreover, it has been very clear to mathematics educators for a decade now that deductions drawn from pure psychological theory rarely apply directly to mathematics learning (see Bauersfeld, 1979). In addition to M-space, we felt that a child's developmental level in specific areas of relevance to the content area under consideration could also be an influence. Thus we incorporated Piagetian tests appropriate to mathematical learning and used both the M-space and developmental data to define a construct, cognitive capacity for mathematical material, which would be more useful in a study that was primarily concerned with mathematics instruction rather than cognitive theory.

In summary, to identify cognitive capacity, we gave two batteries of tests. The first battery of tests was designed to measure the short-term memory capacity (M-space) of the child for processing mathematical material. The second battery included tests constructed to measure the child's level of cognitive development on dimensions from the Piagetian model, such as conservation and transitivity, and presumably related to mathematical ability. We then used psychometric procedures, factor analysis, and cluster analyses to interpret the data from both batteries and to group children. From this approach, we assumed that well-defined groups of children with specific cognitive characteristics could be identified.

Classroom Instruction

Throughout this project the children in the study were being taught to add and subtract in school. To identify some aspects of classroom instruction, we observed in five classrooms to gather data on a sample of students at grades 1, 2, and 3. It is at these grades that addition and subtraction skills are taught. The sample of students we observed was selected to reflect differences in cognitive capacity.

Data on the performance of the students were collected using an achievement monitoring battery developed by Buchanan and Romberg (1983). This battery provides information on a variety of aspects of adding and subtracting, and in several administrations profiles of growth can be obtained. The profiles then can be used as indicators of the effectiveness of instruction.

Third, we decided to observe teacher actions, pupil actions, and teacher-pupil interactions for children at each grade level who differed in cognitive capacity. The proposition that "teachers make a difference" had been central to much of the previous work done on mathematics education at the Wisconsin Center for Education Research. For example, the steps in the Individually Guided Education (IGE) instructional programming model (Klausmeier, 1977) are all descriptions of actions teachers are to take. In addition, as DMP was being developed, behaviors that teachers were to use in teaching the program were specified. Despite these efforts, little evidence is available to substantiate the importance of teacher actions.

Berliner (1975) pointed to the probable reasons for this lack of data and identified a long list of problems facing researchers who attempt to examine the relationships between teacher behaviors and pupil performance. He saw methodology as a major impediment to progress in this area, particularly the inadequate framework for the conceptualization of teacher tasks and the assumed direct relationship between teacher tasks and pupil performance. It is possible that the logical analyses of the problem subsequent to Berliner's rather pessimistic overview are as far from classroom realities as the analyses carried out on mathematics curriculum programs in the 1950s and 1960s or the logical application of general psychologists' theories of development and learning to mathematics programs of a decade ago. Perhaps what was needed was a fresh look at the problem.

In this study we decided to concentrate attention on teachers' actions as they related to children of known cognitive characteristics and, moreover, on the same children's reactions related to the teachers' initiating actions. The approach should make some progress toward conceptualizing teachers' instructional tasks and testing the notion that teachers have some discernible effect on pupils' performance. The approach used in this investigation was a "time-on-task" approach based on that used in the Beginning Teacher Evaluation Study (BTES), which in turn was influenced by Carroll (1963, 1973), Bloom (1974), and Harnischfeger and Wiley (1975).

Another major criticism made by Berliner was the lack, at that time, of instruments that gave researchers a clear understanding of the meaning of data gathered by objective tests or surveys. Moreover, even when observational techniques were employed, it was not usual to code pupil actions. We decided to take advantage of recent advances in this area by using the observational instrument developed for the study of instructional time with DMP (Romberg, Small, Carnahan, & Cookson, 1979). This instrument takes into account the behavior of both teachers and children.

The instrument is used with a limited sample of pupils who are identified as target students. Then a trained observer fills out a time based observational form for each day of instruction. At the end of each minute the observer codes pupil activities, teacher activities, content categories, and classroom characteristics. Data from target students are then aggregated to estimate mean class time on the variables. This methodology provides reliable and generalizable information about how time is spent in classrooms.

Conclusion

The five studies reported in this monograph represent an attempt to draw together data gathered from four different perspectives. Each perspective is viable in its own right. However, our intent was to see whether in combination the perspectives could better portray how students develop addition and subtraction skills. The first approach, from the classical individual differences perspective, was to use psychometric techniques in two studies to identify students with different cognitive capacities. The second approach, from the cognitive-processing perspective, was to gather interview data about the strategies children use to solve verbal addition and subtraction problems. The third approach, from the quasi-experimental perspective, was to assess changes in student achievement using test monitoring procedures. The final approach, the direct instruction teaching perspective, was to use a time-on-task observation procedure to determine how features of classroom instruction relate to student engagement.

The studies were designed not only to gather and analyze data on the four perspectives described above, but also to examine the interactions between the four factors. Obviously, a number of interactions would be of considerable interest, but in view of our interests and to examine some new hypotheses we decided to concentrate on the interaction between children's cognitive processing capability and the other variables.

We first identified a sample of children aged 4-8 years with specific cognitive characteristics. Sample selection required measuring M-space (study 1) and measuring cognitive development (study 2) of a population of 4- to 8-year-olds. Next, we studied the mathematics performance, strategies used, and instruction provided the sample over a 3-month period. In clinical interviews the children's performance and strategies were determined with verbal addition and subtraction problems (study 3). Achievement was measured with standard written addition and subtraction tasks (study 4). The nature of the instruction provided and children's actions and engagement were determined in classroom observations (study 5).

We assumed that from these five studies we would be able to relate performance at a given time (in terms of performance level achieved and strategy adopted) to a child's cognitive capability and to specific instructional activities the child's teacher had used. In this way, we could consider various questions about change in performance and strategy and their possible causes.

The various research techniques used, the data gathered, and their analysis are described in the next four chapters. Chapter 2 is concerned with the means we used to characterize the cognitive processing capabilities examined in studies 1 and 2. Chapters 3 and 4 relate the cognitive level of each group to their performance on addition and subtraction problems. In chapter 3 the individual clinical interview data coded for both performance and strategies used by children are presented (study 3). In chapter 4 achievement on paper-and-pencil tests of addition and subtraction is presented. In chapter 5 we attempt to relate cognitive level to teacher-pupil interactions. Chapter 6 provides a summary of the findings and some conclusions that draw together the understandings obtained through the studies and suggests some direction for further research.

Chapter 2

IDENTIFICATION OF GROUPS OF CHILDREN WHO DIFFER IN COGNITIVE-PROCESSING CAPABILITIES

In this chapter the classification of children into groups according to their cognitive-processing capabilities with mathematical materials is presented. Cognitive-processing capability is a derived categorization label based on a combination of measures of working memory capacity (M-space) and measures of the level of cognitive development as determined by the Piagetian model. The M-space measures were the basis of the classification of children into categories, and the developmental tests gave an indication of developmental criteria that are applicable within each category.

Study 1--M-space

Information-processing theories are based on the idea that mental functions can be characterized in terms of the way information is stored, accessed, and operated on. Mental structures are discussed in terms of an intake register through which information from the environment enters the system, a working or short-term memory (M-space) in which the actual information processing occurs and a long-term memory in which knowledge is stored.

The working memory's growing capacity to process information appears as a fundamental characteristic of cognitive development in a number of theories (Bruner, 1966; Case, 1978a; Flavell, 1971). Young children are quite limited in their ability to deal with all the information demands of complex tasks. Their limited capacity seems to be a critical developmental factor that constrains learning in instructional situations (Case, 1975, 1978a, 1978b).

Pascual-Leone (1970, 1976) proposed a theory that operationalizes the development of information-processing capacity or M-space. According to this theory, learning is a change in behavior resulting from factors extrinsic to the psychological system. Learning produces a change in the repertoire of schemes (internally represented behavioral units or patterns) available to the learner. Since M-space is limited, the number of information chunks that can be coordinated to produce a new scheme is limited. Therefore, the complexity of schemes learned is also limited; the processes of learning are constrained by the developing psychological system. Pascual-Leone's theory is concerned with the functional aspects of development and the mental processing of information. Learning through instruction depends on the child's capacity to process all of the essential incoming information.

To generate hypotheses about children's performance on specific tasks, both the information-processing capacity (M-space) of the child and the information-processing demands of the task must be known. This study addresses the problem of assessing information-processing capacity.

The rationale for administering different tests to measure this construct is based on the results of two recent studies, one by Hiebert (1979), in which a measure of M-space (backward digit span) did not predict learning of mathematical skills and another by Case and Kurland (1978) in which three different measures of M-space (counting span, Mr. Cucui, and digit placement) were given. Although in Case and Kurland's study positive correlations (.50 to .60) were found between the three tests, the consistency between the measures was not high. Recent work by Case and associates (Case, Kurland, Daneman, & Emmanuel, 1979) suggests that it may be very difficult to construct one general measure of M-space that will predict performance on a wide range of tasks. Their data indicate that task variables may be more important than previously supposed in determining M-space demands. Thus, we decided to use the three tests from Case and Kurland's study along with the backward digit span test from Hiebert's study to see whether together they would yield a reliable estimate of a child's M-space. The tests chosen also seemed appropriate in terms of the task variables involved in learning to add and subtract.

Method

Sample

All of the 139 children in grades K-2 at the Sandy Bay Infant School in Hobart, Tasmania, were tested for this study. The school is located on the Derwent River in Sandy Bay, a suburb of Hobart near the University of Tasmania. The community is middle to upper-middle class. Table 2 gives details about the age, grade and gender of the sample and the number of children involved.

Tests

Counting span. This test was developed by Case and Kurland (1978). Conceptually, it is straightforward. The operation required is counting. The items that must be stored are the products of a series of counting operations. Children are presented with a sequence of arrays

Table 2

Characteristics of Sample

Characteristic		Class and Grade 1 K-AM	2 K-PM	3 Prep	4 Gr. 1	5 Gr. 1/2*	6 Gr. 2	Total
Gender	Boys	16	11	8	8	15	15	73
	Girls	9	9	13	14	9	12	66
	Total	25	20	21	22	24	27	139
Age	Youngest	4.9**	5.0	5.4	6.2	6.5	7.3	
	Oldest	5.1	5.7	6.1	7.3	7.10	8.2	
	Average	4.11	5.4	5.10	6.7	7.3	7.8	

*Gr. 1/2 was a mixed class with both Grade 1 and Grade 2 students.

**4.9 means 4 years 9 months as of October 1, 1979.

of geometric shapes to count and are asked to recall the number of objects in the arrays preceding the current trial as soon as they have finished counting the shapes on the current stimulus card. The number of arrays in the set is incremented from trial to trial and children's M-space is assumed to be equal to the maximum number of arrays that they can count while maintaining perfect recall.

The test includes 33 items. However, at most, only five items were scored at any one of five M-space levels. To reduce the total number of trials a modified "ceiling basal" method was used (Bachelder & Denny, 1977). Children were presented with sets from different M-space levels until it was determined at what level they passed and at what level they failed. They were then presented with a larger number of trials until the level of complete success and the level of complete failure had been determined.

Mr. Cucui. This measure was designed in Pascual-Leone's laboratory by DeAvila, for use with children with an imperfect command of English (DeAvila & Havassy, 1974). It can be administered quickly and is suitable for use with four-year-olds as well as older children.

On each trial, children are presented with the outline of Mr. Cucui. After viewing it for five seconds, they are told to remember what parts of his body are colored. They are then presented with a blank outline drawing of Mr. Cucui and told to point to the parts that are colored. There are 25 items, five different items at each of five levels; a level is defined as the number of body parts that are colored.

This test is the only one that does not require students to count or use numbers. Instead, recall of spatial location is required to respond correctly. The ceiling-basal method was followed for the Counting Span Test.

Digit placement. This is a measure of M-space developed and standardized by Case. It is known to yield the same norms as other tests of M-space (cf. Case, 1972) and to correlate highly with the general factor defined by more lengthy M-tests (Case & Globerson, 1974). The basic procedure is to present subjects with a set of numbers. The first *n* - 1 of these are in ascending order of magnitude and the *n*th is out of order (e.g., 2, 5, 9, 12, 7). After the numbers have disappeared from view, the children are asked to indicate where the final number belongs in the original series. M-space corresponds to the maximum set size for which the task can be executed successfully. There are 15 items on this test, five for each of three levels; levels 1 and 5 as measured in the two tests above are not tested. All items were given to each subject.

Backward digit span. The form used in this study was developed by Hiebert (1979). On each trial, the experimenter reads a series of digits. The subject is to repeat them in reverse order. M-space corresponds to the maximum series size correctly repeated. In this test, there are 40 items (10 at each of four levels; level 1 as measured in the first two tests is not tested and all items are given to each student).

Test Administration

A research assistant and two experienced teachers were hired to administer the tests. All were trained before the testing proceeded. One interviewer administered the counting span test; a second the Mr.

Cucui test; and the third the digit placement and the backward digit span tests. Children were randomly selected by their teacher to come to the interview room and randomly assigned to an interviewer. Most children took two tests on one day and the other two a day or two later. All testing was completed within 10 days.

Scoring the Tests

Although each item could obviously be scored correct or incorrect and the total correct counted to estimate each child's M-space level, there were at least two sound reasons why this procedure was not followed. First, because sets of items in each test were designed to measure different levels of M-space, item scores would need to be weighted to reflect those levels--especially as two of the tests did not aim to measure all five levels. Second, since the "ceiling basal" procedure was used with two of the tests, some items were not actually administered to each child; items not administered but at a level lower than where the child responded correctly were scored correct and all items at a level higher than where the child responded correctly were scored incorrect. Four scoring rules were devised for each test. The full details regarding those are available in Romberg and Collis (1980a) and will not be reported here.

Results and Discussion

Table 3 shows the frequency of scores (M-space level) for children in each class and for the total population for each test. In addition, class means and standard deviations are presented. The distributions of scores for the four memory tests provide two interesting results.

Table 3

Frequency of Scores on the M-Space Tests

Class	Score							
	0	1	2	3	4	5	M	SD
				Counting Span Test				
(1) K-AM		22	3				1.12	.33
(2) K-PM	1	9	9	1			1.50	.69
(3) Prep		11	9	1			1.52	.60
(4) Gr. 1		4	15	3			1.96	.57
(5) Gr. 1/2		1	15	8			2.29	.55
(6) Gr. 2		2	12	12	1		2.44	.70
Totals	1	49	63	25	1	0	1.83	.75
				Mr. Cucui Test				
(1) K-AM		12	12	1			1.56	.58
(2) K-PM		5	11	2			1.75	.64
(3) Prep		4	12	5			2.05	.67
(4) Gr. 1		1	9	8	4		2.68	.84
(5) Gr. 1/2			6	9	7	2	3.21	.93
(6) Gr. 2		1	6	7	11	2	3.26	1.02
Totals	0	25	56	32	22	4	2.45	1.05
				Digit Placement Test				
(1) K-AM		24		1			1.08	.40
(2) K-PM		20					1.00	.00
(3) Prep		19	2				1.10	.30
(4) Gr. 1		18	3	1			1.23	.53
(5) Gr. 1/2		12	6		6		2.00	1.25
(6) Gr. 2		5	1	2	19		3.30	1.20
Totals	0	98	12	4	25	0	1.68	1.17
				Backward Digit Span Test				
(1) K-AM		13	12				1.48	.51
(2) K-PM		2	18				1.90	.31
(3) Prep		1	20				1.95	.22
(4) Gr. 1			18	4			2.18	.40
(5) Gr. 1/2			16	8			2.33	.48
(6) Gr. 2			8	15	4		2.85	.66
Totals	0	16	92	27	4	0	2.14	.64

First, although older children generally have higher scores, the overlap of scores among children at different grade levels is quite striking. Scores are clearly age-related but do not appear to be specifically determined by age. Second, the variation of M-space level for individual children across tests (variation in within-class frequencies across tests) could imply that the context of the text may give students a cue that helps them answer questions. In addition, if partial level scores are allowed for children answering items on a test at a higher level, it is a reasonable deduction, on the evidence from the protocols, that the move from one level of M-space to another is gradual.

Relationship of Scores on the Tests

Each of the tests, it was hoped, would reflect the amount of M-space available to the children for processing early math-related material. However, the tasks were different, the student population covered a wide age/grade range, and the children demonstrated considerable variation in performance. Thus, it was important to investigate with some care whether or not the different tests yielded similar classifications of children. Three statistical procedures were performed on the data: (1) a correlation matrix was set up to show the correlations between the scores from the four tests for the total population; (2) the data for all pairs of tests were cross tabulated to see how many classifications were the same; and (3) a factor analysis was performed on the correlation matrix to determine the dimensionality of the scores.

Correlations of test scores. Although all the correlations (see Table 4) are positive and statistically significant, they are not

Table 4

Correlations of Scores for the Four Memory Tests

Test	CS	MC	DP	BDS
Counting Span (CS)	1.00			
Mr. Cucui (MC)	.49	1.00		
Digit Placement (DP)	.61	.50	1.00	
Backward Digit Span (BDS)	.52	.40	.64	1.00

particularly high. The highest is only .64. It seems clear that different tests do not necessarily classify children into the same M-space levels.

Cross-tabulation of scores for the four tests. To examine the similiarity between classification schemes based on the four tests, we cross tabulated the data for each test with each other test. The proportion of students who were classified in the same categories and in different categories in each comparison is shown in Table 4. The percentage of individuals who were differently classified in the comparisons ranges from 68% to 46%.

This cross tabulation demonstrates that the tests classify children in different ways. If these various classifications are along a single dimension, there is not a serious problem; this would mean that each test identifies different cutoff points on this one dimension. However, if these tests are found to measure more than one dimension, then each test is measuring something different.

Factor analysis. The results of the cross-tabulation made examining the dimensionality question more critical. A factor analysis was performed on the correlation matrix presented in Table 5 for the four tests across the total population. The model used was a multifactor solution model. All extractions were principle factor extractions with iterative estimates of commonalities, and the varimax rotation procedure was used. The data from this factor analysis appear in Table 6. A single factor was extracted. However, it should be noted that the Mr. Cucui test did not load heavily on this factor, and a considerable amount of the variance is still unaccounted for. The Mr. Cucui test is the only one of the four that does not ask children to

Table 5

Number and Percentage of Classifications That are the Same, Higher for the First Test, and Lower for the First Test for all Test Comparisons

	Test Comparisons (A/B)					
Classi-fication	CS/DP N(%)	CS/MC N(%)	CS/BDS N(%)	DP/MC N(%)	DP/BDS N(%)	MC/BDS N(%)
Same (A=B)	58(42)	47(34)	75(54)	49(35)	44(32)	57(41)
Higher (A>B)	36(26)	16(12)	13(9)	19(14)	31(22)	55(40)
Lower (A<B)	45(32)	76(55)	51(37)	71(51)	64(46)	27(19)

Note: CS = Counting Span
DP = Digit Placement
MC = Mr. Cucui
BDS= Backward Digit Span

Table 6

Factor Analysis for the Four Memory Tests

	Factor
	1
Eigenvalue	2.59
% variance	64.8
Raw (rotated) factor matrix	
Counting Span	.44(.56)
Digit Placement	.54(.72)
Mr. Cucui	.30(.37)
Backward Digit Span	.44(.51)

count, and it is also less English dependent. This suggests that the factor is a quantitative M-space factor involving memory of number or counting sequences. The Mr. Cucui test, on the other hand, requires memory of spatial orientation.

In summary, the four tests measure one primary factor, quantitative M-space. Thus, to classify children into M-space levels, it would seem best to administer a combination of tests as was done in this study and then to classify children with regard to that underlying structure. No single test, it appears, can reliably classify children into an M-space level. The next section indicates that a classification made on the basis of the results of three tests should be fairly reliable for most children.

Cluster Analysis

Since the factor analysis showed that one dimension accounted for nearly two-thirds of the variance, it seemed desirable for the next stage of the project to classify the children in the population along this single dimension. A cluster analysis procedure, which uses Euclidian distances between points,[1] was used for the classification.

This analysis indicated that there were six groups. Table 7 gives the estimated group vectors for the six groups identified. In the analysis, the last four groups (3, 4, 5, and 6) were closer together

---

[1]The usual Euclidian distance between points in four dimensions was used, i.e.,

$$d = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2 + (x_4 - y_4)^2}$$

Table 7

Estimated Vectors for the Six Groups Derived from a Cluster Analysis Where the Distance Between Score Vectors is Less than 1.50

Group	Amalgamated distance	Number of children	Test				Overall M-space classification
			CS	DPT	BDS	MC	
1	1.05	59	1.32	1.07	1.73	1.61	1
2	1.44	38	1.90	1.66	2.13	2.76	2
3	1.43	16	2.25	2.10	2.25	3.69	2S+
4	1.03	11	2.91	4.00	2.91	2.46	3S-
5	1.06	4	2.17	3.83	2.67	4.50	3S+
6	1.23	6	2.50	4.00	3.75	3.75	4S-

than Groups 1 and 2. This suggests that Groups 1 and 2 are distinct and that Groups 3, 4, 5, and 6, although different from each other, are less distinct.

Group 1 is largest, with 59 members. For the tests separately, the levels for this group are CS, Level 1; DPT, Level 1; BDS, Level 1; and MC, Level 1. This group is clearly at M-space Level 1, the lowest M-space level in the domain being measured. Only for BDS could some children be placed at Level 2.

Group 2 has 38 members. The levels for this group are CS, Level 2; DPT, Level 1; BDS, Level 2; and MC, Level 2. These children exhibit a basic M-space Level 2. They are below that level on the DPT and nearly reach Level 3 on the Mr. Cucui test. These differences seemed, from the protocols, to be due to contextual factors: the children found the instructions for DPT more complex than instructions for other tests, and either spatial perception or ability to check information contributed to scores on the Mr. Cucui test.

Group 3, with 16 members, scored slightly above Level 2 on three tests and nearly reached Level 4 on the Mr. Cucui test. Either their spatial perception is quite high or they are able to chunk information on that test, but they still exhibit a basic M-space level of 2. We have labeled this group Level 2S+ to highlight the fact that these children are above that level on the spatial test.

Group 4 has 11 members. On two tests, CS and DBS, the children are at Level 3; on DPT they are at Level 4, but on the Mr. Cucui test, they are only at Level 2. Their basic M-space level is probably 3. Their spatial perception involved in Mr. Cucui appears not to be as highly developed as their quantitative abilities. Therefore, we have classified them 3S-.

Group 5 has only four members who have a similar pattern of levels to children in Group 4 except that Group 5 scores very well on the Mr. Cucui test. Their basic pattern seems to place them at M-space Level 3, and therefore we have classified them 3S+.

Group 6 has six members. They are basically at M-space Level 4 on three tests but score below Level 3 on CS. It is not clear what the discrepancy on this test implies, and the protocols did not assist in this case. It could, of course, be simply a sampling or testing variation especially since the numbers in the category are so small. This variation needs closer examination than we were able to perform in this study. However, this group is lower than Group 5 on the Mr. Cucui test, but overall their quantitative skills are at Level 4. Therefore, we have labeled them 4S-.

Overall, these results suggest an underlying cognitive mechanism. The contextual setting has a significant effect on the child's ability to respond on any given occasion. This suggests possible significant differences in children's use of problem-solving strategies or their reception to instruction even though they have the same basic cognitive-processing potential. One could hypothesize that spatial development (qualitative) and number development (quantitative) strategies appear to be interwoven and develop close together in time, but some children achieve number skill prior to spatial skill and others vice versa.

Study 2--Cognitive Development

The reason for wanting a battery of tests that measure cognitive development is based on the theory of Jean Piaget (1974). According to Piaget, the development of cognition is inseparable from the growth of biological and psychological faculties. Development is a broad-based process, generalizing to a wide variety of situations.

Piaget's position is summarized in the following statement:

> I think that development explains learning, and this option is contrary to the widely held opinion that development is a sum of discrete learning experiences. (1974, p. 176)

The phrase "development explains learning" implies that the outcome of a learning experience is in part accounted for by developmental capabilities. That is, learning potential is defined (or explained) to a large extent by developmental level.

For this project a battery of 10 tests was devised, all measuring the early development of the child's ability to work with elementary quantitative and logical concepts concerned with premathematical skills. We tested the entire population in order to relate developmental characteristics to characteristics already derived from the M-space tests.

Cognitive Development Tests

As stated earlier, the choice of specific tests was based on our intent to examine the relationship of cognitive capability to children's performance on addition and subtraction tasks. Of the 10 tests, seven were selected from a large battery of tests constructed by Fullerton (1968); two from tests devised by Romberg, Carpenter, and Moser (1978);

and one was constructed by the authors for this study. Details of each test can be found in Romberg and Collis (1980b).

Extension (E). This group test was developed by Fullerton (1968). Children are to decide which of three choice boxes has the same number of dots as a sample box. The term extension refers to the fact that the number sets extend beyond the usual level of subitemization to a higher portion of the number scale. The test contains 12 items. The number of correct responses is scored. A correct answer is interpreted to mean that the child is able to set up a one-to-one correspondence between sets.

Ordinal Correspondence (OC). In this group test, also developed by Fullerton (1968), the format for the items is similar to that in the Extension Test. This test also contains 12 items. The number of correct responses is scored. A correct answer is interpreted as meaning that the child is able to establish an ordinal correspondence between sets.

Conservation of Number (Wohwill, CN-W). This group test, also developed by Fullerton (1968), is based on an earlier test developed by Wohlwill (1960). Six items are given. The number of correct responses is scored. A correct response is interpreted to mean that the child is able to preserve one-to-one correspondence between sets after one set has been rearranged (i.e., is able to overcome perceptual distractions).

Addition-Subtraction (Wohlwill, AS-W). The items for this group test, also developed by Fullerton (1968) and based on Wohlwill's earlier work (1960), are interspersed with those of the previous test (CN-W) because of the similarity between the two tests. This test differs only in that a single object is either added to or subtracted from the collection of objects in front of the children. In this case a correct response is interpreted to mean that the child recognizes that an increase or decrease in one of two sets in one-to-one correspondence means these sets are no longer in such correspondence. Six items are given and the number correct scored.

Transitivity (T). The authors developed this six-item group test because the Coordination of Relations Equivalence Test (CRE, described next) requires a child to attend to both transitivity and a linear rearrangement of sets. The present test was designed to assess only transitivity. A correct response is interpreted as the child being able to preserve both equivalence and order relationships. A total correct score is recorded for each child.

Coordination of Relations of Equivalence Test (CRE). This six-item group test was developed by Fullerton (1968). The items are similar to those in the T test except that the fixed set is also transformed (lengthened, shortened, or heaped together). A correct response here is interpreted as the child being able to preserve equivalence relationships even after rearrangement. The scoring procedure is identical to that for T.

Class Inclusion (CI). This individually administered test of two items was developed by Romberg, Carpenter, and Moser (1978). A correct response is interpreted as a child being able to subdivide logically a set into distinct subsets.

Additive Composition of Number (ACN). This individually administered test, developed by Fullerton (1968), includes three items that ask children to respond to three quite different composition tasks. A correct response implies the child can establish an equivalence relationship by the common practice of sharing and preserve such a correspondence when distracting information is presented.

Counting On (CO). This individually administered test was developed by Romberg, Carpenter, and Moser (1978). The test includes three items for each of the three levels of counting on: small number onto a number less than 10, small number onto a number between 10 and 20, and a large number onto a number between 10 and 20. The typical question asked was, Could you start counting at 13 to find the number that is four more than 13? Children are marked as passing a level if they answer two of three items correctly. A total score is then recorded of the number of levels passed (0, 1, 2, or 3).

Counting Back (CB). This test is like formal CO; however, in this case the typical question asked was, Could you count back starting at 15 to find the number that is four less than 15? The scoring procedure used is the same as in CO.

Test Administration

Because the order in which these tests were administered was important, and because they would be administered to children of varying ages, two decisions were made to gather the data more efficiently. First, the tests were separated into four sets to be administered at separate times. Second, not all of the tests were given to all children. The organization of the tests and the rules for selecting who was to take which test are given in Table 8. The interview tests and set 2 were given to all children. A child passing the two tests in set 2 (CN-W and AS-W) was assumed to have passed set 1 and was given set 3.

Table 8

Tests Included in Each Set, Sequence of Administration, and Rules for Selecting Subjects

Order	Set (tests)	Rule
1	Interview (ACN, CI, CO, CB)	All children
2	Set 2 (CN-W, AS-W)	All children
3	Set 1 (E, OC)	Children failing either test in set 2
4	Set 3 (T, CRE)	Children passing both tests in set 2

However, if a child failed either of the tests in set 2, set 1 was administered, and the child was assumed to have failed set 3.

On the interview tests one assistant administered the CO and CB Tests, and the other administered the CI and the ACN Tests. Again children were randomly selected by their teachers to come to the interview room (the teachers' lounge). Each interviewer was in a corner of the room. Children were randomly assigned to an interviewer. Children took two tests on one day and the other two a day to two after. Shortly after the interviews were completed, the group batteries were given. Set 2 was given first to groups of six to eight children at a time from each class. The research assistant presented the stimulus information for each test following a script and using a large magnet board. The other assistants observed the children to make sure they were working on the correct page, responding in the right place, and not copying from others. Set 1 was given next, followed by set 3. All testing was completed within four weeks.

Results

Intercorrelations Among Cognitive Development Tests

Full summary tables for the raw score data for each of the tests are given by Romberg and Collis (1980b). To examine the relation between the tests and the structure of the battery itself, a two-step procedure was followed.

Fullerton (1968) used scalogram analysis to organize the battery of tests he developed. He found tests that grouped together, and he established an order for the tests based on test difficulty. Unfortunately, that method fails to establish the underlying dimensionality of the data matrix or the possible structure of the assumed hierarchy. A more satisfactory method is to determine first the dimensionality of the intercorrelations of the tests. If the matrix is unidimensional, then a hierarchy can be established.

The intercorrelations across the whole population for the 10 cognitive processing tests appear in Table 9. The correlations are all positive but fairly low, ranging from .24 to .79; 17 of the 28 correlations fall between .40 and .58. We decided to exclude the E and OC tests from the correlation matrix for further analysis on the grounds that they were baseline tests on which most children scored at the ceiling.

Factor Analysis of Cognitive Development Tests

To determine the dimensionality of the intercorrelations, a factor analysis was performed on the matrix shown in Table 8 with tests E and OC excluded. A multifactor solution model was used. All extractions were principal factor extractions with iteration estimates of commonalities; the varimax rotation procedure was employed. The results of this analysis are shown in Table 10.

A two factor solution was derived, although the eigenvalue for the first factor is considerably larger than that for the second factor. An examination of this rotated factor matrix shows that the counting tests (CB, CO) load heaviest on the rotated first factor, followed by the tests in Battery 4, T and CRE. This factor may reflect a mature level of counting skill. The other four tests also load on this factor, but not to the same degree. At best we can say that it is probably a quantitative factor influenced by the ability to count. The second

Table 9

Intercorrelations of the Ten Cognitive Development Tests

	E	OC	ACN	CN-W	AS-W	CI	CO	CB	T	CRE
E	1.00									
OC	.45	1.00								
ACN	.22	.25	1.00							
CN-W	.30	.32	.35	1.00						
AS-W	.35	.37	.48	.51	1.00					
CI	.13	.13	.32	.24	.28	1.00				
CO	.22	.28	.55	.43	.42	.44	1.00			
CB	.15	.21	.49	.40	.39	.45	.79	1.00		
T	.13	.16	.43	.42	.36	.47	.52	.61	1.00	
CRE	.17	.21	.51	.55	.48	.39	.58	.62	.68	1.00
Maximum	12	12	3	6	6	2	3	3	6	6
Mean	10.88	10.63	1.97	4.86	5.03	.51	1.35	1.06	3.93	1.75
Std. deviation	1.90	2.28	.94	1.56	1.34	.81	1.29	1.21	4.68	1.49

Table 10

Factor Analysis for Eight Cognitive Development Tests

	Factors	
	1	2
Eigenvalue	4.34	.92
% variance	54.30	11.50
Raw (rotated) factor matrix		
ACN	.64(.46)	.07(.45)
CN-W	.61(.25)	.35(.65)
AS-W	.60(.24)	.37(.66)
CI	.52(.49)	-.13(.23)
CO	.80(.76)	-.22(.33)
CB	.83(.85)	-.33(.26)
T	.73(.60)	-.06(.41)
CRE	.81(.56)	.11(.59)

factor seems more qualitative, involving the ability to make comparisons and see transformations without having to count. In particular, the Wohlwill tests (AS-W and CN-W) load heaviest on this rotated factor load. One test, Class Inclusion, does not load heavily on either factor. Since Class Inclusion involves logical reasoning and is the only nonquantitative test, this finding gives credence to our interpretations of the first two factors.

The factor analysis of the data seemed to show that there were two interpretable dimensions underlying performance on the tests. However, since the first factor accounted for such a large proportion of the variance (54.30%), we examined the possible hierarchical ordering of the tests using Guttman's (1954) simplex procedures. It is clear as one examines the correlation matrix as a whole that the tests are not in simplex order. Even when we take a subset of the matrix, the five tests (ACN, CN-W, CI, T, CRE) that might be considered to test aspects of logical functioning at this level do not satisfy the criteria. It seems from all the evidence, then, that there is no basis for a hierarchical ordering of these tests. In summary, the cognitive development tests do not seem to measure a single dimension. Rather, about two-thirds of the variance on the tests can be explained in terms of two dimensions, a quantitative factor influenced by the ability to count, which accounts for over half of the variance, and a qualitative factor that involves an ability to make comparisons and see transformations without counting.

Relationships Between Cognitive Development and M-Space Tests

In this section of the analysis, we attempted to combine the information from the M-space tests and the cognitive development tests with a view to grouping the children/pupils into categories that have distinct describable cognitive characteristics.

To begin with, a correlation matrix (Table 11) was drawn up for the four M-space tests and the eight cognitive development tests (tests E and OC being omitted for reasons given earlier). The correlations range from .29 to .79, with 20 of the 32 falling between .40 and .59. The higher correlations with the M-space tests occur with both the counting tests (CO, CB). This is not surprising. The counting tests undoubtedly rquire a larger memory capacity than some of the other tests. However, there is no apparent variation in correlations of the different memory tests with the cognitive processing tests. This suggests that the positive correlation is along a single dimension.

Factor Analysis of Cognitive Development and M-Space Tests

To check this suggested unidimensional relationship, a factor analysis was carried out in which the four M-space tests were added to the eight cognitive development tests. The data for that factor analysis appear in Table 12. Again, as was the case with the factor analysis of the cognitive development tests (see Table 9), two factors appeared. The two factors have the same structure as the two factors that appeared in the earlier analysis. The memory tests load heavily on the first factor but not the second.

At this point, we decided that we had enough information to look for a pattern in the achievement on all cognitive tests for each of the six groups formed by the cluster analysis of the M-space tests. The proportion correct in each cognitive test for each M-space category is set out in Table 13; a graphical representation of the same information

Table 11

Correlations of the Eight Cognitive Development Tests and the Four M-Space Tests

M-Space Tests	Cognitive Processing Tests							
	ACN	CN-W	AS-W	CI	CO	CB	T	CRE
Counting Span	.54	.32	.39	.43	.63	.61	.47	.53
Digit Placement	.54	.44	.41	.45	.77	.79	.69	.63
Mr. Cucui	.46	.32	.37	.48	.53	.55	.46	.47
Backward Digit Span	.48	.48	.50	.38	.61	.58	.55	.54

Table 12

Factor Analysis for Eight Cognitive Development Tests and the Four M-Space Tests

	Factors	
	1	2
Eigenvalue	6.52	1.02
% variance	54.40	8.50
Raw (rotated) factor analysis		
ACN	.65(.56)	.08(.36)
CN-W	.58(.40)	.41(.50)
AS-W	.59(.37)	.41(.60)
CI	.55(.56)	-.12(.12)
CO	.83(.78)	-.18(.28)
CB	.84(.85)	-.25(.15)
T	.73(.73)	-.01(.16)
CRE	.78(.70)	.16(.31)
CS	.71(.68)	-.13(.25)
DP	.86(.74)	-.19(.10)
MC	.63(.68)	-.09(.25)
BDS	.73(.62)	.13(.43)

is shown in Figure 1. It can be seen that there are clear differences between Groups 1 and 2 and the other four groups. Within the latter groups, Groups 3 and 4 differ little from each other but do differ from Groups 5 and 6, which are also very similar.

Group 1 children with M-space level 1 were below the other groups in all four areas and were in general incapable of handling quantitative tasks. They could make qualitative comparisons and transformations only at a moderate level.

Group 2 children with M-space level 2 were also without specific quantitative skills, although they performed considerably better than Group 1 on all the tests. They could handle qualitative correspondence at an acceptable level although they scored somewhat lower than the other groups on the conservation of number test.

Group 3 children with M-space level 2S+ were high on qualitative correspondence, had developed the specific counting skills of counting on and counting back, but were inadequate in their use of those skills on the transitive reasoning test. Their logical reasoning was also deficient, although they performed considerably better than Groups 1 or 2 on that test.

Group 4 children with M-space level 3S- were high on qualitative correspondence and all the quantitative tests, but inadequate on the logical reasoning test. In fact, they differed significantly from Group 3 only on the additive composition test and the transitivity test.

Groups 5 and 6 with M-space levels 3S+ and 4S- presented similar profiles on these tests. They reached the ceiling on the qualitative correspondence tests, scoring a little higher than Groups 2, 3, and 4.

Table 13

Percent Correct for the Six M-Space Groups on the Ten Cognitive Development Tests

Group (M-space level)	Test									
	Baseline		Factor 1 (Quantitative)							Logical
			Factor 2 (Qualitative)							
	E	OC	AS-W	CN-W	ACN	CO	CB	CRE	T	CI
1(1)	90	84	46	48	45	9	3	6	2	6
2(2)	100	100	91	71	80	52	35	43	14	20
3(2S+)	100	100	87	93	71	82	78	73	53	50
4(3S-)	100	100	100	90	87	93	87	80	90	50
5(3S+)	100	100	100	100	100	91	75	100	100	88
6(4S-)	100	100	100	100	93	100	87	100	80	90

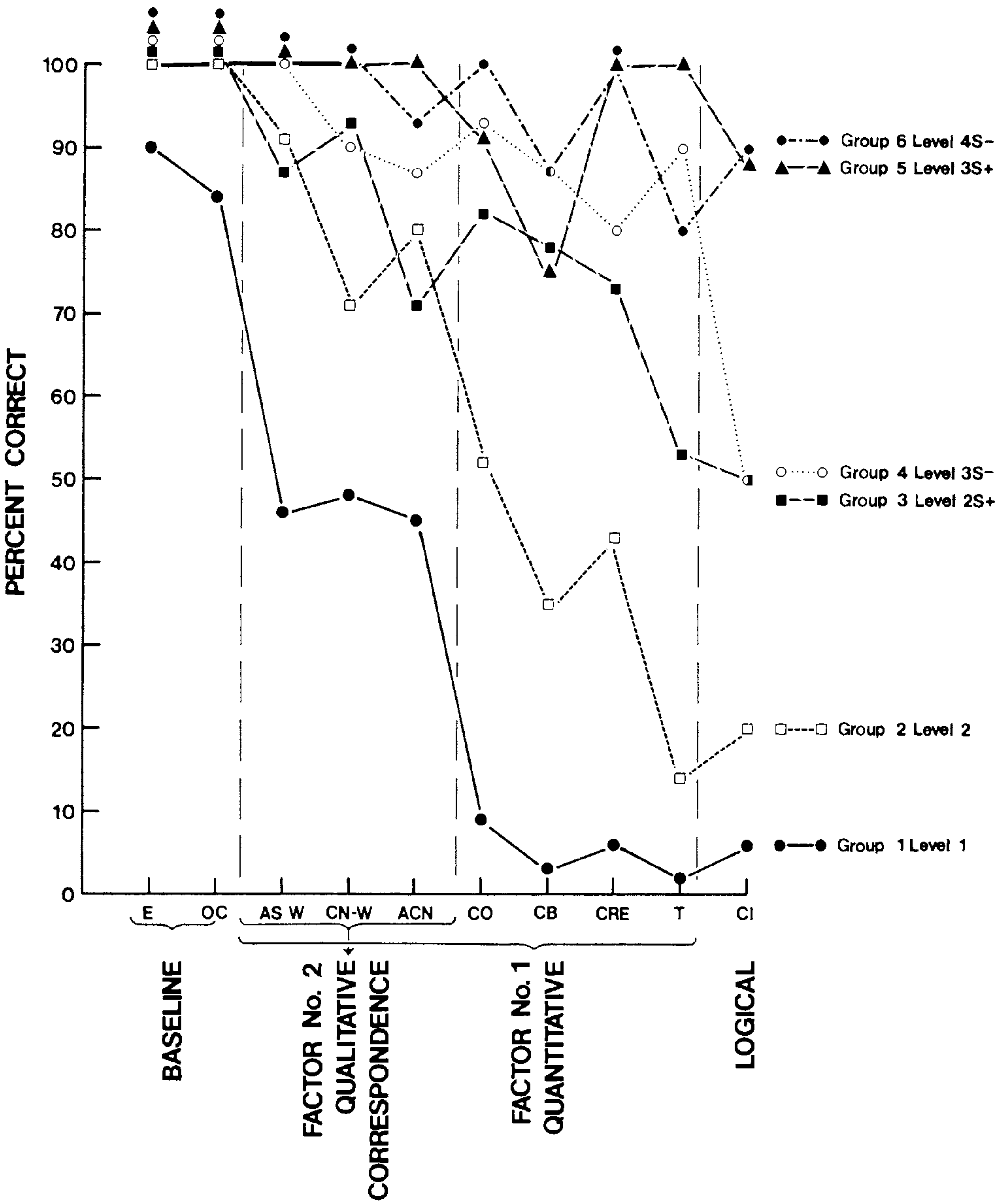


Figure 1. Pattern of scores (percent correct) for the six M-space groups on ten cognitive process tests grouped by factors.

Like Group 4 children, they had high scores on all the quantitative tests. Children in these groups were high on the class inclusion test.

From these cluster groups, a sample of students was drawn for Studies 3, 4, and 5 in this series in the following school year.

Summary and Conclusion

Based on data from four memory tests and eight cognitive development tests, we were able to identify groups of children who had well-defined but different cognitive-processing capabilities. This identification was accomplished in the following steps. First, using cluster analysis on the memory test scores, we identified six groups of students with similar patterns of responses. Second, from the results of a factor analysis, we found that the cognitive development tests loaded on two factors: a quantitative factor that involves mature counting strategies and a qualitative correspondence factor. Third, by examining how the six groups defined by the M-space analysis performed on the cognitive tests, we demonstrated that the cognitive-processing scores of five of the six groups differed systematically.

This last step was the basis for the remainder of the project. We formed five distinct groups of students (cluster groups 5 and 6 were combined) with known cognitive capabilities related to the learning of mathematical materials. In the following chapters, we describe how this information was used to study several aspects of the children's interaction with mathematics instruction in early elementary school.

In conclusion, the data gathered and analyzed in this chapter suggest that the following propositions deserve close attention by both researchers and practitioners:

1. A global qualitative/quantitative distinction is apparent in children's mathematical thinking in the early school year;
2. M-space level seems to be related to the development of other cognitive skills;
3. The suggested developmental sequence in the preschool to early elementary years in mathematically related reasoning appears to be: comparison -- qualitative correspondence -- quantitative -- logical operations;
4. An M-space level of 1 is enough for handling simple comparison tasks;
5. An M-space level of 2 is enough for understanding qualitative correspondence and is a prerequisite for the development of number skills;
6. An M-space level of 3 seems necessary for success on sophisticated counting tasks.

In all, these data suggest that simple correspondence (both equivalence and order) appears to be the first ability to develop. This is followed by a qualitative correspondence capacity that involves understanding how correspondence between two sets is preserved or changed under varying circumstances. Next, the quantitative skills of counting on and counting back develop, followed by their use in transitivity tasks. Finally, the capacity for logical reasoning develops.

Chapter 3

COGNITIVE-PROCESSING CAPACITY AND CHILDREN'S PERFORMANCE ON VERBAL ADDITION AND SUBTRACTION PROBLEMS

In this chapter, the third study in this set is reported. Its purpose was to study the relationships among children's cognitive capacity, and their performance and use of strategies on verbal addition and subtraction problems. The importance of knowing how children learn the concepts and procedures of addition and subtraction should be self-evident. It is frequently assumed that children must first master computational skills and then begin to solve verbal addition and subtraction problems. However, it has been clearly demonstrated that children develop a variety of strategies for solving such mathematical problems, independent of instruction (cf. Carpenter & Moser, 1983; Ginsburg, 1977; Resnick, 1978). In fact, many of the strategies they use are more sophisticated and demonstrate more insight than the procedures that are taught.

A sample of the children tested in the previous studies (Chapter 2) and selected to reflect different cognitive capabilities was interviewed on three occasions over a 3-month period in 1980 (27-29 February, 9-1 April, and 26-28 May). In each interview, a set of verbal addition and subtraction problems was given to each child. The interviewer coded each student's performance and strategies.

Method

Sample

The children from the earlier studies had advanced a grade in school since previous testing. Furthermore, the grade 2 students who were in Sandy Bay Infant School in October now were in third grade and in different primary schools. Most, however, were enrolled at Waimea Heights Primary School.

Our intent was to have a sample of two to four students from each cognitive level in each grade. We began with rosters of students from each grade and their cognitive level. Then an initial selection of students was made. However, after school began, some third graders originally in one class were switched to another. This created some imbalance across classes but should not have affected the results. The 44 students in this study are shown by cognitive group, class, and grade in Table 14.

Procedure

Types of problems. An interview consisted of six problem types (tasks) given under four of six conditions. The six types included two problems solvable by addition of the two given numbers and four problems solvable by subtraction of the two given numbers. The types differed in terms of semantic structure. The semantic characterization for these six problem types is detailed in Moser (1979) and in Carpenter and Moser (1979, 1983).

Table 15 presents representative problems in the order in which the problems were administered to the children. The actual wording for each problem type differed but the semantic structure remained constant.

Table 14

Children in Each Cluster Group in Each Class

Cognitive Group	Sandy Bay Infant School		Waimea Heights Primary School			Total
	Class		Class			
	1	2	3	4	5	
	Grade 1	Grade 2	Grade 3	Grade 3	Grade 3	
1	3	2	0	0	0	5
2	3	6	0	4	0	13
3	1	2	2	3	3	11
4	0	0	2	3	3	8
5,6	0	0	3	1	3	7
Totals	7	10	7	11	9	44

Table 15

Problem Types

Task	Sample Problem
1. Change/Join (Addition)	Pam had 3 shells. Her brother gave her 6 more shells. How many shells did Pam have altogether?
2. Change/Separate (Subtraction)	Jenny had 7 erasers. She gave 5 erasers to Ben. How many erasers did Jenny have left?
3. Combine/Part Unknown (Subtraction)	There are 5 fish in a bowl. 3 are striped and the rest are spotted. How many spotted fish are in the bowl?
4. Combine/Whole Unknown (Addition)	Matt has 2 baseball cards. He also has 4 football cards. How many cards does Matt have altogether?
5. Compare (Subtraction)	Angie has 4 lady bugs. Her brother Todd has 7 lady bugs. How many more lady bugs does Todd have than Angie?
6. Change/Join, Change set Unknown (Subtraction)	Gene has 5 marshmallows. How many more marshmallows does he have to put with them so he has 8 marshmallows altogther?

Within each problem, two of three numbers from a number triple (x, y, z) defined by $x + y = z$, $x \leq y \leq z$, were given. In the two addition problems, x and y were presented, with the smaller number x always given first. In the four subtraction problems, z and the larger addend y were presented. The order of presentation of y and z varied among problem types.

The six semantic problem types used were presented under six conditions, although not all children responded to all conditions. Four conditions resulted from crossing smaller zs and larger zs with presence and absence of manipulative materials. In the smaller number problems (called SN problems), the addition guideline of $5 \leq z \leq 9$ was imposed. In the larger number problems (called LN problems), the restriction on the sum was $11 \leq z \leq 15$. Problem sets SNp and LNp were given with manipulatives present; the same sets given with manipulative absent were called SNa and LNa.

For the interviews with third-grade children, the domain of two-digit numbers was included. In the two-digit domain, two subdomains were identified. In the NR probems, no regrouping (borrowing or carrying) was required to determine a difference or sum when a computational algorithm was used. In the second subdomain R problems, regrouping was required. For the two-digit problems, the sum z was restricted to numbers in the 20s and 30s. All the third-grade children took the LN, NR, and R problems. Complete details of the procedures used are reported in Romberg, Collis, and Buchanan (1981).

Interview method. Three trained interviewers administered the interviews (see Cookson & Moser, 1980, for details of interviewer-training procedures and reliability). One interviewer worked at Sandy Bay Infant School and the other two at Waimea Heights Primary School. Each interviewer was able to conduct from 8 to 12 interviews a day, depending on the schools' schedules and on the task level. (The LN tasks took longer than the SN tasks.) At the schools, the interviewers were assigned interview areas, which were quiet rooms separate from distracting activities. The verbal tasks were read and reread to the child as often as necessary so that remembering the given numbers or relationships caused no difficulty. An individual interview required two sessions, one for the SN tasks and the other for LN tasks (or one for the LN and the other for NR and R). The sessions lasted 15-25 minutes each, with each child receiving the same sequence of problems. No child was interviewed twice in one day.

Coding student responses. All of the possible codings of student responses are presented in detail in Cookson and Moser (1980). Three or four elements were coded for each child: model used, correctness, strategy, and, if incorrect, error. A record of each subject's responses to the tasks was compiled from the coding sheets. These profiles are the basis for all other statistical information appearing in this chapter and are reported in Romberg, Collis, and Buchanan (1981).

Data Aggregation and Analysis

The interview data are summarized in terms of percent correct and general strategy. The data for percent of items answered correctly by children are summarized by examining the differences for children with differing cognitive processing capabilities. It was anticipated that children in Group 5-6 would answer more items correctly than those in

Group 4, who in turn would answer more items correctly than the Group 3 children, and so forth.

Pupil strategy was categorized according to type of model used (if any), strategy or process used, and errors (if any). Five general categories for the SN and LN problems are the following:

1. Direct modeling--use of the manipulatives provided, or fingers, to stand for the problem entities. Actions performed on the objects generally correspond to the action or relationship described in the problem.
2. Use of counting sequences--use of the string of counting words, either forward or backward, where the entry point in the sequence is a number other than 1. Counting may proceed in either direction a given number of counts, or until a desired number (usually one of the numbers given in the problem) is reached. This requires a second counting or some sort of a tracking mechanism, often aided by the use of fingers.
3. Routine mental operations--use of memorized number facts by direct recall.
4. Nonroutine mental operations--derivation of a nonmemorized fact through manipulation of some other recalled fact. As an example, the fact for 6 + 8 can be derived by determining it to be two more than the easily remembered doubles fact of 6 + 6.
5. Inappropriate Behaviors--guessing, using one of the given numbers in the problem, adding instead of subtracting, or giving no answer at all.

For the NR and R data, the five categories used with the SN and LN tasks were used if students did not write a sentence. If students did write a sentence, three other categories were used.

6. Correct sentence/algorithmic. This category of behavior includes the standard algorithms taught in school as well as any "invented" (Carpenter & Moser, 1982) ones that involve considerations of place value. Algorithmic behavior must be exhibited by use of paper and pencil.
7. Correct sentence/nonalgorithmic. After writing a sentence, the work is done mentally as was frequently seen in problems in which no regrouping (NR tasks) was required (Moser, 1980).
8. Inappropriate sentence. This behavior involves writing and working with the wrong sentence (e.g., addition instead of subtraction).

Details of what specific model, strategy, and error data were used to form these categories are presented in Romberg, Collis, and Buchanan (1981).

The plan for analyses of the aggregated data was based on the two primary dimensions in this study--differences in the level of problem administered and differences in children's cognitive capacity. The problem dimension involves a completely crossed repeated assessment (three interviews) of six problem sets (SNp, SNa, LNp, LNa, NR, and R), with six tasks in each set (combine/join, combine/separate, and so on). The student dimension includes children nested in cognitive levels within classes and in turn, within grades.

The data matrix is incomplete since not all cognitive levels are represented in each grade level, grade 1 and grade 2 children did not take the NR and R problems, and the grade 3 children did not take the SN problems. The small number of subjects, the unequal cell sizes, and the extensive incompleteness of the matrix limited us to describing the frequencies and testing a few of the differences with chi-square statistics.[1]

For purposes of this report, frequency and percent correct and frequency of use of strategy are presented for children with different cognitive processing capabilities. The data are presented for three problem sets (SN, LN, and NR and R combined) and for each semantic task within each set. Other analyses performed for each interview and by grade level are not reported here. Those analyses can be found in Romberg, Collis, and Buchanan (1981).

Results

Performance by Cognitive Groups

All SN and LN tasks. To examine whether or not differences in cognitive capacity are reflected in different percentages of correct responses, separate tables are presented for each problem set. In Table 16, the data for the SN problems that were given only to grade 1 and grade 2 children clearly show that there is a significant increase in percent correct (56% to 75% to 88%) for children in cognitive Groups 1, 2, and 3, respectively ($\chi^2 = 47.19$, $p < .01$).

[1] Because of the large number of trials and lack of a systematic plan to test differences, an alpha level of .01 was arbitrarily chosen to test significance. In addition, tests that yielded probability values between an alpha of .05 and .01 ($.05 > p > .01$) were considered marginally significant. All χ^2 values were calculated via 2 x 2 contingency tables where frequency of correct answers or strategy was dichotomized.

Table 16

Frequency and Percent Correct by Cognitive Group for
All SN Tasks

Cognitive Group	N	Total Responses[a]	Correct Responses	
			Frequency	Percent
1	5	180	100	56
2	9	312[a]	235	75
3	3	108	95	88
4	-	-	-	-
5,6	-	-	-	-
Total	17	600	430	72

[a]When all children were present for all three interviews, number of trials equals N times 12 problems (6 SNp and 6 SNa) times three occasions.

For the LN problems given to all children, the percent correct for children in different cognitive groups is shown in Table 17. The differences are striking. The Group 1 children only got 22% correct, while children in Group 5-6 got 96% correct. There is a significant increase from Group 1 to Group 2 (22% to 65%, $\chi^2 = 94.38$, $p < .01$), from Group 2 to Group 3 (65% to 81%, $\chi^2 = 26.74$, $p < .01$), and again from Group 4 to Group 5-6 (83% to 96%, $p < .01$). The lack of difference in percent correct between Cognitive Group 3 and Group 4 children is not surprising, since these groups differed very little on the cognitive tests.

<u>All NR and R tasks</u>. For the D and E problems given only to grade 3 children, the pattern of correct responses were very similar. Thus, for summary purposes, the data on these problems are combined in Table 18. For these students, the difference between percentage correct for children in Cognitive Groups 2 and 3 (49% and 67%) is significant ($\chi^2 = 11.76$, $p < .01$), as are the differences between Cognitive Groups 4 and 5-6 children (62% and 83%, $\chi^2 = 30.05$, $p < .01$). Again, the differences in performance between Cognitive Groups 3 and 4 on both sets of problems are not significant.

Overall, our predictions about percentage of the items answered correctly were found to be accurate, except that children in Cognitive Groups 3 and 4 differed very little in terms of their general performance.

<u>Performance by Task</u>

Within each problem set, one item representing each of six tasks (change/join; change/separate; combine/part unknown; combine/whole

Table 17

Frequency and Percent Correct by Cognitive Group for
All LN Tasks

Cognitive Group	N	Total Responses[a]	Correct Response	
			Frequency	Percent
1	5	180	40	22
2	13	456	206	65
3	11	396	320	81
4	8	264	220	83
5,6	7	252	241	96
Total	44	1548	1117	72

[a]When all children were present for all three interviews, number of trials equals N times 12 problems (6 LNp and 6 LNa) times three occasions.

Table 18

Frequency and Percent Correct by Cognitive Group for
All NR, R Tasks

Cognitive Group	N	Total Responses[a]	Correct Response	
			Frequency	Percent
1	-	-	-	-
2	4	144	71	49
3	8	264	176	67
4	8	252	155	62
5,6	7	252	210	83
Total	27	912	612	67

[a]When all children were present for all three interviews, number of trials equals N times 12 problems (6 NR and 6 R) times three occasions.

unknown; compare and change/join; change/set unknown) was given. Because the different semantics of each problem type elicit different cognitive demands, we anticipated that performance would vary with the tasks. Following Greeno's (1980) categorization of the six tasks given (see Table 15) we expected Tasks 1 and 2 (change/join and change/separate) to be the easiest, for they demand only a change/cause schema; Task 4 (combine/whole unknown) to be next in difficulty, for it involves a harder combination schema; Tasks 6 and 3 (the missing addends problem) to follow in difficulty because of the location of the missing information; and Task 5 (comparison/subtraction) to be hardest because it involves a comparison schema which requires more units of memory.

Each SN task. The percent correct data for each cognitive group for each task in the SN set of problems are presented in Table 19. The pattern of differences between cognitive groups is consistent with Group 3 children performing better than Group 2 who, in turn, perform better than the Group 1 children. As expected for the SN level Tasks 1 and 2 were easy for all children. Tasks 4 and 6, however, were just as easy. Task 3 was more difficult, and Task 5 was hard for all children.

Each LN task. The percent correct data for each cognitive group on each task for the LN set of problems are presented in Table 20. If two thirds of the items were correct, then this was used as a rough criterion for success for those data. Again, a consistent pattern of the children in the higher cognitive group getting as many or more items correct is apparent. The one exception to this pattern was on Task 3 (combine/part unknown), the Group 4 children did not do as well as the Group 3 children on those tasks. Group 1 children were generally unable to work any of the LN problems successfully. The majority of Group 2

Table 19

Frequency and Percent Correct by Cognitive Group for Each SN Task

Cognitive Group	N	Total Responses	Correct Response Frequency	Correct Response Percent
Task 1 Change/Join (+)				
1	15	30	23	77
2	26	52	44	85
3	9	18	18	100
4	-	-	-	-
5,6	-	-	-	-
Total	50	100	85	85
Task 2 Change/Separate (-)				
1	15	30	21	70
2	26	52	42	81
3	9	18	16	89
4	-	-	-	-
5,6	-	-	-	-
Total	50	100	79	79
Task 3 Combine/Part Unknown (-)				
1	15	30	12	40
2	26	52	40	77
3	9	18	16	89
4	-	-	-	-
5,6	-	-	-	-
Total	50	100	68	68

Cognitive Group	N	Total Responses	Correct Response Frequency	Correct Response Percent
Task 4 Combine/Whole Unknown (+)				
1	15	30	21	70
2	26	52	43	83
3	9	18	16	89
4	-	-	-	-
5,6	-	-	-	-
Total	50	100	80	80
Task 5 Compare (-)				
1	15	30	7	23
2	26	52	22	42
3	9	18	12	67
4	-	-	-	-
5,6	-	-	-	-
Total	50	100	41	41
Task 6 Change/Join, Change set unknown (-)				
1	15	30	16	53
2	26	52	44	85
3	9	18	17	94
4	-	-	-	-
5,6	-	-	-	-
Total	50	100	77	77

Table 20

Frequency and Percent Correct by Cognitive Group for Each LN Task

Cognitive Group	N	Total Responses	Correct Response Frequency	Correct Response Percent
Task 1 Joining (+)				
1	15	30	9	30
2	38	76	55	72
3	33	66	57	86
4	22	44	40	91
5,6	21	42	40	95
Total	129	258	201	78
Task 2 Separating (-)				
1	15	30	7	23
2	38	76	52	68
3	33	66	51	77
4	22	44	34	77
5,6	21	42	40	95
Total	129	258	184	71
Task 3 PPW, missing addend (-)				
1	15	30	6	20
2	38	76	54	71
3	33	66	56	85
4	22	44	33	75
5,6	21	42	41	98
Total	129	258	190	74

Cognitive Group	N	Total Responses	Correct Response Frequency	Correct Response Percent
Task 4 PPW (+)				
1	15	30	11	37
2	38	76	52	68
3	33	66	53	80
4	22	44	38	86
5,6	21	42	40	95
Total	129	258	194	75
Task 5 Comparison (-)				
1	15	30	1	3
2	38	76	30	39
3	33	66	49	74
4	22	44	38	86
5,6	21	42	39	93
Total	129	258	157	58
Task 6 Joining, missing addend (-)				
1	15	30	6	20
2	38	76	53	70
3	33	66	54	82
4	22	44	37	84
5,6	21	42	41	98
Total	129	258	191	74

children worked all problems except Task 5. The children in the higher groups were able to work all problems. However, except for the difficult comparison problems (Task 5), the tasks were of equal difficulty.

Each NR and R task. The same data for the NR and R sets of problems are shown in Table 21. Again, the same pattern is evident except for the Cognitive Group 4 children whose performance was marginally lower than Group 3 children on Tasks 1 and 2 and was about the same as Group 2 children on Task 5. Overall, Group 2 children were only successful on Task 1. Group 3 and 4 children were successful on Tasks 1, 4, and 6. And, Group 5-6 children were successful on all tasks. However, unexpectedly, Task 2 was as hard as Tasks 3 and 5 for all of the children. These data suggest that when problems have large enough numbers, children should use algorithms, because the implied computational procedures become more important than the semantics. Thus, addition problems are easier than subtraction problems. Although Task 6 is a subtraction problem, it was often solved using additive notions, making it easier than Task 2.

In summary, although there were important variations in performance due to problem set (size of number) to specific task, and to grade, it is clear that children who had been identified as having different cognitive-processing capabilities performed differently on these addition and subtraction tasks regardless of the other important factors.

Table 21

Frequency and Percent Correct by Cognitive Group for Each NR, R Task

Cognitive Group	N	Total Responses	Correct Response Frequency	Correct Response Percent
Task 1 Change/Join (+)				
1	-	-	-	-
2	12	24	16	67
3	22	44	39	89
4	21	42	33	78
5,6	21	42	37	88
Total	76	152	125	82
Task 2 Change/Separate (-)				
1	-	-	-	-
2	12	24	10	42
3	22	44	27	61
4	21	42	21	50
5,6	21	42	31	74
Total	76	152	89	58
Task 3 Combine/Part Unknown (-)				
1	-	-	-	-
2	12	24	9	38
3	22	44	22	50
4	21	42	22	52
5,6	21	42	35	83
Total	76	152	88	58

Cognitive Group	N	Total Responses	Correct Response Frequency	Correct Response Percent
Task 4 Combine/Whole Unknown (+)				
1	-	-	-	-
2	12	24	14	58
3	22	44	31	70
4	21	42	31	74
5,6	21	42	39	93
Total	76	152	115	76
Task 5 Compare (-)				
1	-	-	-	-
2	12	24	12	50
3	22	44	28	64
4	21	42	20	48
5,6	21	42	30	71
Total	76	152	90	59
Task 6 Change/Join, Change set unknown (-)				
1	-	-	-	-
2	12	24	10	42
3	22	44	29	66
4	21	42	28	67
5,6	21	42	38	90
Total	76	152	105	68

Strategies Children Used

As outlined in the first part of this chapter, the data on strategies children used are summarized in terms of five categories for the SN and LN problem sets (direct modeling, counting sequences, routine mental operations, nonroutine mental operations, and inappropriate strategy) and eight categories for the NR and R problem sets (the same five no-sentence categories as for SN and LN tasks plus correct sentence-algorithmic, correct sentence-nonalgorithmic, and incorrect sentence).

We expected that children with low cognitive capacity would either use inappropriate strategies or directly model problems. Children with higher capacities would use counting sequences and routine mental operations. Algorithms would be used in increasing frequency by children at higher levels of competency.

All SN tasks. To show whether children with different cognitive capacities used different strategies, separate tables are presented for each problem set. For the SN problems given only to grades 1 and 2 children (Table 22), as expected, there was a significant increase in use of routine mental operations (8% to 27% to 35%) by children with higher cognitive capacity ($\chi^2 = 36.97$, $p < .01$) and a corresponding significant decrease in use of an inappropriate strategy (39% to 18% to 7%; $\chi^2 = 34.80$, $p < .01$). However, the frequency of use of the other categories unexpectedly remained constant over cognitive levels.

All LN tasks. For the LN problems given to all children, the strategy data for children in different cognitive groups are shown in Table 23. The picture here is more dramatic. As anticipated, children in Cognitive Group 1 either directly modeled the problems (28% of the

Table 22

Frequency of Use of Strategies by Cognitive Group and Category for All SN Tasks

Cognitive Group	Responses	Direct Modeling		Counting Sequences		Routine Mental Operation		Nonroutine Mental Operation		Inappropriate	
		Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent
1	180	69	38	20	11	15	8	5	3	71	39
2	312	120	38	43	14	85	27	8	2	56	18
3	108	39	36	17	16	39	36	5	5	8	7
Total	600	228	38	80	13	139	23	18	3	135	22

Table 23

Frequency of Use of Strategies by Cognitive Group and Category for All LN Tasks

Cognitive Group	Responses	Direct Modeling		Counting Sequences		Routine Mental Operation		Nonroutine Mental Operation		Inappropriate	
		Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent
1	180	50	28	1	0+	2	1	0	0	127	70
2	456	166	36	82	18	59	13	27	6	122	27
3	396	71	18	130	33	104	26	38	10	53	13
4	264	30	11	79	30	92	35	38	14	25	9
5,6	252	32	13	101	40	105	42	14	6	0	0
Total	1548	349	22	393	25	362	23	117	8	327	21

trials) or used an inappropriate strategy (70% of the trials). Use of an inappropriate strategy goes down consistently with increase in cognitive capacity (70% for Group 1 children to 0% for Group 5-6 children). Direct modeling is the strategy most often used by Cognitive Group 2 children; counting sequences by Group 3 children; and routine mental operations by Groups 4 and 5-6 children who also used counting sequences frequently.

All NR tasks. Data for the NR problems, which were given only to the third-grade children, are summarized in Table 24. As expected, between Cognitive Groups 2 and 5-6 there is a significant increase in use of counting strategies from 12% to 33% ($\chi^2 = 10.40$, $p < .01$) and a corresponding decrease in use of inappropriate strategies from 29% to 2% ($\chi^2 = 30.86$, $p < .01$). Unexpectedly, children at all cognitive levels used other strategies at about the same frequency.

All R tasks. The data for the R problems, also given only to third graders, are summarized in Table 25. As for the NR problems, from Group 2 to Group 5-6, the use of counting strategies increased significantly from 4% to 32% ($\chi^2 = 20.50$, $p < .01$) and the use of inappropriate strategies decreased from 44% to 5% ($\chi^2 = 46.52$, $p < .01$). For both NR tasks and R tasks, there was no appreciable increase in the use of algorithms by children in higher cognitive groups (NR, 21% to 25%; R, 26% to 25%).

Strategies by Task Type

Within each problem set, one item representing each of six tasks (change/join; change/separate; combine/part unknown; combine/whole unknown; compare, and change join/change/set unknown) was given. From

Table 24

Frequency of Use of Strategies by Cognitive Group and Category for All NR Tasks

Cognitive Group	Responses	Direct Modeling		Counting Sequences		Routine Mental Operation		Nonroutine Mental Operation		Inappropriate		Algorithm		Non-Algorithm		Incorrect Sentence	
		Frequency	%	Frequency	%	Frequency	%	Frequency	%	Frequency	%	Frequency	%	Frequency	%	Frequency	%
2	72	14	19	9	12	8	11	3	4	21	29	15	21	1	1	1	1
3	132	29	22	27	20	24	18	5	4	21	16	25	19	0	0	1	1
4	127	25	20	28	22	23	18	9	7	15	12	24	19	1	1	2	2
5,6	126	20	16	42	33	27	21	1	1	3	2	31	25	2	2	0	0

Table 25

Frequency of Use of Strategies by Cognitive Group and Category for All R Tasks

Cognitive Group	Responses	Direct Modeling		Counting Sequences		Routine Mental Operation		Nonroutine Mental Operation		Inappropriate		Algorithm		Non-Algorithm		Incorrect Sentence	
		Frequency	%	Frequency	%	Frequency	%	Frequency	%	Frequency	%	Frequency	%	Frequency	%	Frequency	%
2	72	5	7	3	4	7	10	2	3	32	44	19	26	2	3	2	1
3	132	28	21	30	23	17	13	0	0	30	23	24	18	1	1	2	1
4	127	29	23	17	13	17	13	1	1	32	25	27	21	0	0	2	2
5,6	126	28	22	40	32	16	13	1	1	6	5	31	25	1	1	3	2

past research (e.g., Carpenter & Moser, 1982), we anticipated that different strategies would be used on tasks with differing semantic structures (particularly on the missing addend problems, Tasks 3 and 6, and on the compare problem, Task 5).

Each SN task. The strategy data for each cognitive group for each task for the SN set of problems are presented in Table 26. A consistent inverse relationship between use of inappropriate strategies and cognitive level is apparent. Although the percentages of various strategies used with each of the tasks differ, the patterns seem to be consistent across cognitive groups. For example, few students used direct modeling for the compare and change/join missing addend tasks (Tasks 5 and 6), regardless of cognitive group. In particular, students used counting sequences most frequently with Task 6.

Each LN task. The strategy data for each cognitive group on each task for the LN set of problems are presented in Table 27. Again, the use of direct modeling decreases as cognitive capacity increases as does the use of inappropriate strategies, while the use of counting sequences and routine mental operations generally increases with capacity. Cognitive Group 1 children directly modeled or used inappropriate strategies for all types of tasks. The use of other strategies varies by task. Again, direct modeling is not used often with Tasks 5 and 6.

Each NR and R task. The data for the NR and R sets of problems are shown in Table 28. Again, the same pattern is evident. Students used direct modeling strategies only for subtraction tasks. Routine mental operations or algorithms were used on Tasks 1 and 4 (the addition tasks) and algorithms on Task 2 (the simplest subtraction task). Students often used direct modeling on all subtraction tasks but rarely for

Table 26

Frequency of Use of Strategies by

Cognitive Group	N	Direct Modeling Frequency	Direct Modeling Percent	Counting Sequences Frequency	Counting Sequences Percent
Task 1					
1	15	16	53	3	10
2	26	23	44	5	10
3	9	6	33	3	17
Total	50	45	45	11	11
Task 2					
1	15	17	57	1	3
2	26	27	52	3	6
3	9	10	56	1	6
Total	50	54	54	5	5
Task 3					
1	15	12	40	1	3
2	26	26	50	1	2
3	9	9	50	2	11
Total	50	47	47	4	4
Task 4					
1	15	17	57	3	10
2	26	25	48	8	15
3	9	10	56	2	11
Total	50	52	52	13	13
Task 5					
1	15	2	7	1	3
2	26	10	19	7	13
3	9	3	17	1	6
Total	50	15	15	9	9
Task 6					
1	15	5	17	11	37
2	26	9	17	19	36
3	9	1	6	8	44
Total	50	15	15	38	38

Cognitive Group and Category for Each SN Task

Routine Mental Operation		Nonroutine Mental Operation		Inappropriate		
Frequency	Percent	Frequency	Percent	Frequency	Percent	Trials
Change/Join (+)						
3	10	4	13	4	13	30
17	33	3	6	4	8	52
7	39	2	11	0	0	18
27	27	9	9	8	8	100
Change/Separate (-)						
2	7	0	0	10	33	30
16	31	2	4	4	8	52
6	33	0	0	1	6	18
24	24	2	2	15	15	100
Combine/Part Unknown (-)						
2	7	0	0	15	50	30
14	27	1	2	10	19	52
5	28	2	11	0	0	18
21	21	3	3	25	25	100
Combine/Whole Unknown (+)						
3	10	1	3	6	20	30
14	27	0	0	5	10	52
6	33	0	0	0	0	18
23	23	1	1	11	11	100
Compare (-)						
2	7	0	0	25	83	30
4	8	1	2	30	58	52
8	44	0	0	6	33	18
14	14	1	1	61	61	100
Change/Join, Change Set Unknown (-)						
3	10	0	0	11	37	30
20	38	1	2	3	6	52
7	39	1	6	1	6	18
30	30	2	2	15	15	100

Table 27

Frequency of Use of Strategies by

Cognitive Group	N	Direct Modeling		Counting Sequences	
		Frequency	Percent	Frequency	Percent
Task 1					
1	15	11	37	0	0
2	38	32	42	13	17
3	33	12	18	23	35
4	22	1	2	15	34
5,6	21	1	2	14	33
Total	129	57	22	65	25
Task 2					
1	15	12	40	0	0
2	38	33	43	5	6
3	33	15	23	17	26
4	22	5	11	17	39
5,6	21	8	19	20	48
Total	129	73	28	59	23
Task 3					
1	15	11	37	0	0
2	38	32	42	14	18
3	33	19	27	17	26
4	22	4	9	14	32
5,6	21	7	17	13	31
Total	129	73	28	58	22
Task 4					
1	15	11	37	0	0
2	38	34	45	18	24
3	33	9	14	28	42
4	22	7	16	11	25
5,6	21	5	12	16	38
Total	129	66	26	73	28
Task 5					
1	15	1	3	0	0
2	38	14	18	12	16
3	33	9	14	24	36
4	22	9	20	12	27
5,6	21	7	17	20	48
Total	129	40	16	68	26
Task 6					
1	15	4	13	1	3
2	38	21	28	20	26
3	33	7	11	21	32
4	22	4	9	10	23
5,6	21	4	10	18	43
Total	129	40	16	70	27

Cognitive Group and Category for Each LN Task

Routine Mental Operation		Nonroutine Mental Operation		Inappropriate		
Frequency	Percent	Frequency	Percent	Frequency	Percent	Trials
Change/Join (+)						
0	0	0	0	19	63	30
15	20	6	8	10	13	76
19	29	7	11	5	8	66
18	41	7	16	3	7	44
21	50	6	14	0	0	42
73	28	26	10	37	14	258
Change/Separate (-)						
0	0	0	0	18	60	30
12	16	5	6	21	28	76
12	18	12	18	10	15	66
12	27	5	11	5	11	44
10	24	4	9	0	0	42
46	18	26	10	54	21	258
Combine/Part Unknown (-)						
1	3	0	0	18	60	30
8	10	5	6	17	22	76
17	26	8	12	5	8	66
11	25	8	18	7	16	44
20	48	2	5	0	0	42
57	22	23	9	47	18	258
Combine/Whole Unknown (+)						
0	0	0	0	19	63	30
10	13	0	0	14	18	76
17	26	4	6	8	12	66
17	39	7	16	2	4	44
20	48	1	2	0	0	42
64	25	12	5	43	17	258
Compare (-)						
0	0	0	0	29	97	30
4	5	6	8	40	53	76
16	24	2	3	15	23	66
14	32	5	11	4	9	44
15	36	0	0	0	0	42
49	19	13	5	88	34	258
Change/Join, Change Set Unknown (-)						
1	3	0	0	24	80	30
10	13	5	6	20	26	76
23	35	5	8	10	15	66
20	45	6	14	4	9	44
19	45	1	2	0	0	42
73	28	17	6	58	22	258

Table 28

Frequency of Use of Strategies by

		No Sentence					
		Direct Modeling		Counting Sequences		Routine Mental Operation	
Cognitive Group	N	Frequency	Percent	Frequency	Percent	Frequency	Percent
							Task 1
2	12	1	4	2	8	2	8
3	22	7	16	6	14	15	34
4	21	2	5	2	5	16	38
5,6	21	4	10	3	7	16	38
Total	76	14	9	13	9	49	32
							Task 2
2	12	4	17	0	0	2	8
3	22	15	34	7	16	2	4
4	21	13	31	6	14	6	14
5,6	21	13	31	5	12	2	5
Total	76	45	30	18	12	12	8
							Task 3
2	12	4	17	2	8	3	12
3	22	13	29	10	23	4	9
4	21	13	31	7	17	3	7
5,6	21	14	33	17	40	5	12
Total	76	44	29	36	24	15	10
							Task 4
2	12	1	4	1	4	4	17
3	22	7	16	2	4	12	27
4	21	5	12	4	9	8	19
5,6	21	4	9	6	14	11	26
Total	76	17	11	13	9	35	23
							Task 5
2	12	4	17	4	17	2	8
3	22	8	18	16	36	3	7
4	21	9	21	13	31	3	7
5,6	21	6	14	25	60	3	7
Total	76	27	18	58	38	11	7
							Task 6
2	12	5	21	3	12	2	8
3	22	7	16	16	36	5	11
4	21	12	29	13	31	4	9
5,6	21	7	17	26	62	6	14
Total	76	31	20	58	38	17	11

Cognitive Group and Category for Each NR, R Task

Nonroutine Mental Operation		Inappropriate		Correct Sentence: Algorithm		Correct Sentence: Non-Algorithm		Incorrect Sentence: All Strategies		
Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent	Trials
Change/Join (+)										
0	0	6	25	12	50	1	4	0	0	24
1	2	2	4	12	27	1	2	0	0	44
1	2	5	12	16	38	0	0	0	0	42
0	0	0	0	19	45	0	0	0	0	42
2	1	13	9	59	39	2	1	0	0	152
Change/Separate (-)										
0	0	8	33	8	33	1	4	1	4	24
1	4	6	13	13	29	0	0	0	0	44
0	0	4	10	11	26	1	2	1	2	42
0	0	2	5	16	38	2	5	2	5	42
1	0	20	13	48	31	4	3	4	3	152
Combine/Part Unknown (-)										
1	4	10	41	4	17	0	0	0	0	24
0	0	11	25	5	11	0	0	1	2	44
1	2	13	31	4	9	0	0	1	2	42
0	0	2	5	3	7	1	2	0	0	42
2	1	36	23	16	10	1	1	2	1	152
Combine/Whole Unknown (+)										
0	0	10	42	8	33	0	0	0	0	24
0	0	10	23	13	29	0	0	0	0	44
0	0	6	14	19	45	0	0	0	0	42
0	0	0	0	21	50	0	0	0	0	42
0	0	26	17	61	40	0	0	0	0	152
Compare (-)										
0	0	12	50	2	8	0	0	0	0	24
1	2	11	25	3	7	0	0	2	4	44
5	12	10	24	0	0	0	0	2	5	42
1	2	4	9	2	5	0	0	1	2	42
7	5	37	24	7	5	0	0	5	3	152
Change/Join, Change Set Unknown (-)										
4	17	7	29	0	0	1	4	2	8	24
2	4	11	25	3	7	0	0	0	0	44
3	7	9	21	1	2	0	0	0	0	42
1	2	1	2	1	2	0	0	0	0	42
10	7	28	18	5	3	1	1	2	1	152

addition, and often used counting sequences with the missing addend problems (Tasks 3 and 6) and the compare problem (Task 5). A considerable increase in use of counting sequences is apparent from Cognitive Group 2 to Cognitive Group 5-6 on these tasks. There remains a significant relationship between use of inappropriate strategies and cognitive group.

Summary and Conclusion

In summary, there are important variations in strategies associated with problem set (size of number) and specific tasks. There also appear to be important interactions between the strategies used by children who have been identified as having different cognitive-processing capabilities and problem set and task. Children with different capacities use different strategies on these addition and subtraction tasks, regardless of the other important factors. It should be understood that other factors, such as grade, have been confounded with cognitive-processing capacity.

Chapter 4

COGNITIVE-PROCESSING CAPACITY AND CHILDREN'S PERFORMANCE ON STANDARD ADDITION AND SUBTRACTION PROBLEMS

In this chapter, the fourth study in this series is reported. Its purpose was to relate the children's cognitive capacity and grade level to their performance on a standard set of items related to addition and subtraction. The procedure used in this study was achievement monitoring (Romberg & Braswell, 1973), which involves repeatedly measuring groups of students in a quasi-experimental design (Campbell & Stanley, 1963). The measures were objective-referenced sets of items on various aspects of addition and subtraction. The quasi-experimental design involved combining longitudinal and cross-sectional designs. Figure 2 shows the design for describing both the longitudinal and cross-sectional data. The within-grade longitudinal growth is represented by the relative heights of the unshaded planes for the groups of students in each grade. The shaded plane across grades perpendicular to the time-of-testing axis represents cross-sectional growth.

The data gathered in this study are summarized first in terms of percent correct on the scales for each grade to portray longitudinal growth. Second, cross-sectional growth profiles are presented on the common scales across grades. Third, summarizations of performance are made for students belonging to the same cognitive groups by grade and across grades. Finally, we relate these data for third-grade children to the strategies they used to solve the verbal problems (NR and R problems) discussed in Chapter 3.

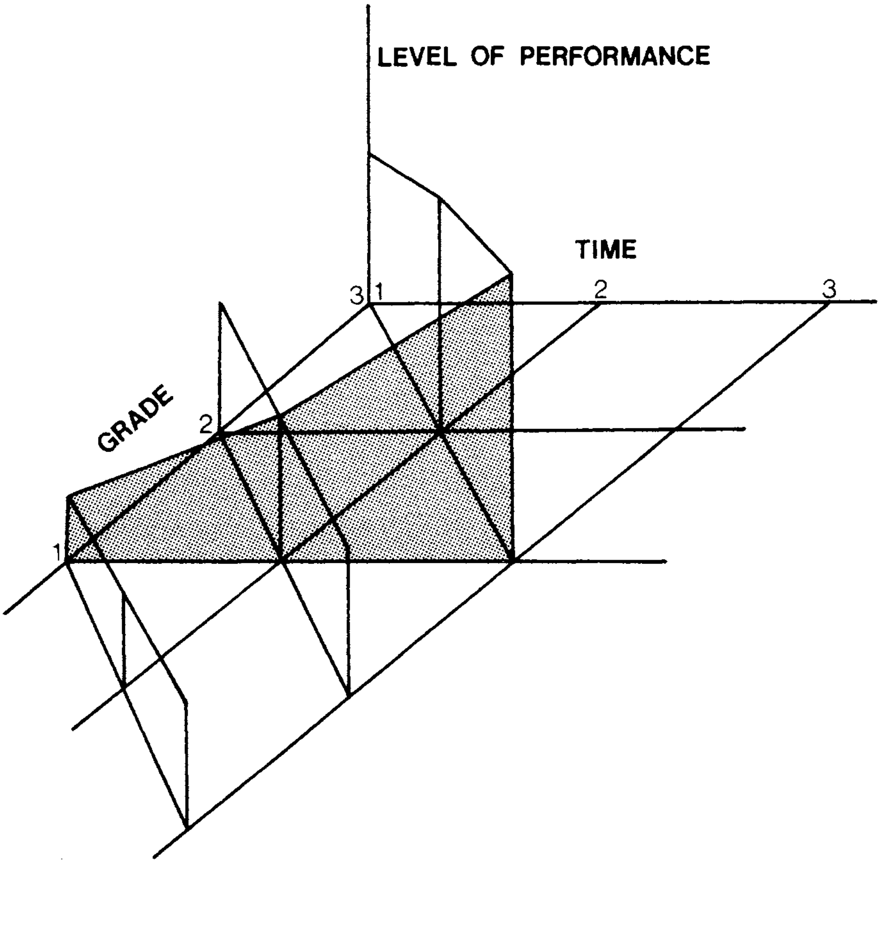


Figure 2. Longitudinal mean growth (unshaded planes) and cross-sectional growth (shaded plane) for students in grades 1, 2, and 3.

Method

Sample

A sample of 44 children in grades 1-3 from the population examined in the previous studies in this series (see Chapters 2 and 3) were administered a set of items on three occasions over a 3- to 4-month period in 1980 (29 February, 11 April, and 28 May or 6 July). The number of children at each grade level is shown in Table 29. In each administration, a set of test items was given to each student. Each child's performance on all items was scored. This report presents the data from those test administrations.

Description of the Tests

A battery of paper-and-pencil objective-referenced tests had previously been developed to assess student achievement on addition and subtraction skills at grades 1, 2, and 3 (Buchanan & Romberg, 1983). The battery contained three test forms for each grade. The items were written to assess the instructional objectives of 10 experimental topics designed to teach addition and subtraction as well as to measure performance on certain prerequisite objectives and noninstructional objectives (Romberg, Carpenter, & Moser, 1978). A summary of all objectives included in the battery is provided in Table 30. Not all objectives were assessed at all grade levels, however. Because of the small sample of students to be tested, one of the three forms was administered at each grade (Form K at grade 1, Form S at grade 2, Form V at grade 3).

Form K was a 30-minute test containing three subtests: a 15-item multiple-choice subtest and two separate 9-item subtests assessing

Table 29

Children at each Cognitive Group in each Grade

	Sandy Bay Infant School		Waimea Heights Primary School	
Cognitive Group	Grade 1	Grade 2	Grade 3	Total
1	3	2	0	5
2	3	6	4	13
3	1	2	8	11
4	0	0	8	8
5,6	0	0	7	7
Total	7	10	27	44

Table 30

Objectives Assessed in Addition and Subtraction

Achievement Monitoring Battery

Prerequisite Instructional Objectives	Non-instructional Objectives
Numerousness	Problem-Solving 0-20
0-10	add-simple joining
11-20	subt-simple separating
0-99, writes	subt-part part whole-addend
0-99, represents	add-part part whole
	subt-comparison
Ordering, Place Value	subt-join-addend
sets, one-to-one correspondence	
numbers 0-20	Problem-Solving 0-99
numbers 0-99, orders	add-simple joining
numbers 0-99, notation	subt-simple separating
	subt-part part whole-addend
Instructional Objectives for the S and A Topic Series	add-part part whole
	subt-comparison
	subt-join-addend
Open Sentences	
add 0-20	Counting 9-31
subt 0-20	on
	back
Sentence-Writing 0-20	
add-simple joining	Basic Facts--Speeded Test
subt-simple separating	add 0-20
subt-part part whole-addend	subt 0-20
add- part part whole	
subt-comparison	Algorithms--Timed Test
subt-join-addend	add 0-99
	add 0-99
Sentence-Writing 0-99	
add-simple joining	
subt-simple separating	
subt-part part whole-addend	
add-part part whole	
subt-comparison	
subt-join-addend	
Algorithms	
add 0-99	
subt 0-99	

recall of addition and subtraction facts under time limits. Form S was a 35-minute test containing four subtests; three of the subtests were similar to the Form K subtests with some items dropped and some added to form a 19-item multiple-choice subtest and two 12-item recall tests. The fourth subtest was a 4-item free-response sentence-writing measure. Form V for third grade was a 40-minute test containing six subtests. In this case, the two recall subtests and the sentence-writing subtest were identical for the Form S subtests. Five items were dropped from the Form S multiple-choice subtest, leaving 14 items. The two new subtests were 24-item timed measures of performance on addition and subtraction algorithms.

Multiple-choice subtests. Individual objectives in the areas of numerousness, ordering, place value, open sentences, and algorithms were represented by one multiple-choice item in each test form on which they were assessed. For the two objectives for counting, counting on and counting back for numbers to 18, there was one item per form; however, an additional counting item for numbers to 31 was included in each test because information on these numbers was of potential interest relative to interview problem situations using larger numbers (see Chapter 3).

Four individual objectives for sentence-writing were represented by a multiple-choice item in each form. For grade 1, these items contained numbers 5-9 or 11-15; for grades 2 and 3 the number domains were 11-15 and 0-99. Since there was no way in a multiple-choice format to have students actually write a sentence, the items required listening to a verbal problem read aloud and then choosing the sentence that correctly represented the verbal situation. The problem situation itself was not printed on the test page. This prevented reading difficulties and also was in keeping with the procedures for the interviews, in which the problems were presented orally.

For Form K, two objectives for the problem-solving area were assessed while for Forms S and V, four objectives were included. The number domains were the same as for the sentence-writing objectives, and again, the problem situations were not printed in the student booklets.

All of the questions in the multiple-choice section of the tests were read to the children and then the key phrases were repeated; in the case of the verbal problems for the sentence-writing and problem-solving objectives, the entire story situation was read twice. The children then marked an X on one of the four response choices: the solution, two distractors, and the "puzzled face," an option that indicated "I have not learned this yet." The response choices, symbols, and pictures were not read or explained to the children (with the exception of the puzzled face).

The puzzled face option was provided to avoid unnecessary frustration and to reduce random guessing. Although we expected students to use this choice throughout the achievement testing because there would always be objectives not yet introduced and/or mastered, this option was particularly useful at the baseline period. Marking the puzzled face allowed children to give a positive response indicating that they had not yet learned to find the answer to a question.

Speeded subtests. There were 9 addition and 9 subtraction facts on Form K and 12 on each of Forms S and V. The first six problems in each case covered the facts from 4 to 9; the last three (or six) involved 10 to 18. The test administered introduced the addition and subtraction recall subtests; then specific directions on a tape recording preceded the items presented, with intervals of 4 seconds working time for Form K and 2 seconds for Forms S and V. The children wrote their answers in designated spaces, leaving spaces for unknown facts empty. There was a short break between the two subtests.

Sentence-writing free-response subtests. Four of the 12 individual sentence-writing objectives (verbal problem types) for the numbers 0-20 and 0-99 were assessed in Forms S and V. A free-response format was employed in which a verbal problem was read twice to the students, who were directed to write a sentence for the situation and not to solve the sentence. There were two 0-20 and two 0-99 items per test.

Addition and subtraction algorithms timed subtests. These subtests, in Form V only, each contained 24 items. The items were either two-digit or three-digit; 18 items required regrouping, 6 did not. The items were arranged in order of difficulty. For example, three-digit problems not requiring regrouping preceded three-digit problems that required regrouping, and three-digit regrouping problems in which only the ones were regrouped preceded problems in which both ones and tens were regrouped. The students were instructed to try each problem in order (the problems were alphabetized) and to go on to the next problem if unable to do a particular example. Six minutes were allowed for each subtest.

Test Administration

The three assistants who gathered data in Study 3 (Chapter 3) also carried out that task in this study. Guidelines for administering the achievement tests were provided to each assistant. The guidelines indicated which tests were to be given, dates for administration, and so forth.

The first administration was supervised by Professor Romberg and went smoothly. The second and third administrations were carried out after Professor Romberg had returned to the United States. These test admininstrations at grade 1 went smoothly as scheduled. At grade 2, one item on Form S did not copy well, so students could not read that question. At grade 3 there were two administrative mixups. First, Form S rather than Form V was given in April to all three classes and in May to two of the classes. This was not a serious problem, since many items are the same, except that the timed algorithms tests were not given. Second, in the third class, Form V was given in July rather than May. The May administration was scheduled near the end of the term, but the assistant failed to administer the tests at that time. After a short break, children returned to school to start the next term. The assistant asked whether she should still gather the data and was advised to administer Form V in July. The results of this administration would not reflect much additional instruction, since there had been a break between terms. All data were then shipped to Madison and scored by Center staff. Each subject's responses were recorded and are the basis for all summary information appearing in this paper.

Results and Discussion

Longitudinal Growth Within Grades

Grade 1. The percent correct for students at grade 1 on the individual objectives and composite objectives for each of the three administrations is shown in Table 31. Overall, the data show that, at

Administration Time for Grade 1, Form K

Description of Objectives	Results for Objectives: Number of Items	Feb. N=7	April N=7	May N=7	Results for Composite Objectives: Number of Items	Feb. N=7	April N=7	May N=7
Prerequisite instructional objectives								
Numerousness								
0-10	1	100	100	100	2	86	71	93
11-20	1	71	43	86				
Ordering								
Sets, one-to-one correspondence	1	86	71	86	2	93	86	86
Numbers 0-20	1	100	100	86				
Instructional objectives for S topics								
Open sentences								
Add 0-20	1	43	57	86	2	29	36	50
Subt 0-20	1	14	14	14				
Sentence-writing 0-20								
Subt-simple separating (11-15)	1	14	0	0	4	21	11	21
Subt-comparison (5-9)	1	29	14	0				
Add-simple joining (11-15)	1	29	14	57				
Subt-part part whole-addend (11-15)	1	14	14	29				
Noninstructional objectives								
Problem solving 0-20								
Add-part part whole (5-9)	1	100	100	100	2	64	57	86
Subt-comparison (11-15)	1	29	14	71				
Counting on 9-31	2	29	43	57	3	19	33	43
Counting back 9-31	1	0	14	14				
Recall of basic facts--speeded test								
Add 0-20					9	33	49	76
Subt 0-20					9	29	44	56

the start of the school year (February), this sample of students had acquired the prerequisite objectives and could solve the verbal addition problems (but probably not by addition), and some (43%) could find the answer to an open addition problem. They could not solve subtraction problems, write sentences, count in or count back, nor could they recall basic facts.

By the end of the autumn term (May), these students' addition skills had improved dramatically. The percent correct increased for solving an open sentence, 43% to 86%; writing a correct addition sentence, 29% to 57%; counting on, 29% to 57%; and addition facts, 33% to 76%. However, the same cannot be said for subtraction. Only for solving a verbal comparison problem (29% to 71%) and for subtraction facts (29% to 56%) was there marked improvement. Obviously, instruction in grade 1 had some effect.

Grade 2. The picture is somewhat different for grade 2 students (see Table 32). At the beginning of the school year, this sample of nine students generally had a low percent correct. In fact, on only three items did more than half of the students get the correct answer. Part of the difficulty was that Form S used large numbers (0-99) in several of the questions. By May, improvement on several composite objectives was apparent. The students were comfortable with numerousness of larger sets (56% to 75%), had improved on basic facts (29% to 51% and 23% to 53%, but not yet to any level of mastery), could solve simple open sentences (17% to 88%), and had improved in counting (30% to 63%) and writing sentences for verbal problems (28% to 59%). Again, instruction had an effect, but increases in performance were not

Table 32

Percent Correct for Objectives and Composite Objectives by

Administration Time for Grade 2, Form S

Description of Objectives	Results for Objectives				Results for Composite Objectives			
	Number of Items	Feb. N=9	April N=9	May N=8	Number of Items	Feb. N=9	April N=9	May N=8
Prerequisite instructional objectives								
Numerousness								
Writes 0-99	1	--	--	--[a]	1	56	67	75
Represents 0-99	1	56	67	75				
Ordering, place value								
Ordering 0-99	1	11	0	25	2	6	0	19
Place value 0-99	1	0	0	13				
Instructional objectives for S and A topics								
Open sentences								
Add 0-20	1	22	78	100	2	17	39	88
Subt 0-20	1	11	0	75				
Sentence-writing 0-20, 0-99 (multiple choice)								
Subt-simple separating (11-15)	1	33	33	25	4	17	14	16
Subt-comparison (0-99)	1	0	0	0				
Add-simple joining (0-99)	1	11	11	25				
Subt-part part whole-addend (11-15)	1	22	11	13				
Sentence-writing 0-20, 0-99 (free response)								
Subt-simple separating (0-99)	1	56	44	75	4	28	53	59
Subt-part part whole-addend (0-99)	1	0	0	0				
Add-part part whole (11-15)	1	56	89	100				
Subt-join-addend (11-15)	1	0	78	63				
Algorithms								
Addition algorithm	1	11	33	13	2	11	17	25
Subtraction algorithm	1	11	0	38				

continued

Table 32 (continued)

Description of Objectives	Results for Objectives: Number of Items	Feb. N=9	April N=9	May N=8	Results for Composite Objectives: Number of Items	Feb. N=9	April N=9	May N=8
Noninstructional objectives								
Problem-solving 0-20, 0-99					4	22	39	25
Add-part part whole (0-99)	1	0	22	25				
Subt-comparison (11-15)	1	22	56	50				
Subt-part part whole-addend (11-15)	1	44	67	13				
Subt-join-addend (0-99)	1	22	11	13				
Counting on 9-31	2	33	28	81	3	30	33	63
Counting back 9-31	1	22	44	25				
Recall of basic facts--speeded test								
Add 0-20					12	29	35	51
Subt 0-20					12	23	30	53

[a]Students were unable to complete item because tests duplicated poorly.

apparent for ordering large numbers, problem solving, selecting written sentences for verbal problems, and algorithms.

Grade 3. The picture was more encouraging for grade 3 students (see Table 33). In February, their performance was not high (above 80%) except on two items, but by the end of May (or early July) performance on all composite objectives except one approached 80%. The exception was the item on place value for numbers 0-99. Sentence writing-selecting skills had improved, but scores were not yet high on some subtraction situations (comparison and part-part-whole addend).

Grade 3 students' performance on the timed algorithms test is shown in Table 34. When all 22 children were tested in February, they performed well on the six addition-without-regrouping problems and acceptably on the three items testing two-digit subtraction without regrouping. On all others, they did poorly. Part of the difficulty was that because of the timed conditions most students did not attempt the last items in the test. Those children who did reach the items did fairly well on the addition regrouping items but had considerable difficulty with the subtraction items requiring regrouping.

Unfortunately, no children were given this test again in April or May, and only 12 in July. By then, those students' performance was considerably better. They still had some difficulty with the three-addend addition problems and the subtraction regrouping problems, but the improvement in every case is striking.

In summary, for this small sample of children assessed at each grade level, growth within each grade on some aspects associated with addition and subtraction is clear. Growth, however, is not uniform across objectives. In addition, overall level of performance on many

Table 33

Percent Correct for Objectives and Composite Objectives by

Administration Time for Grade 3, Forms S, V

Description of Objectives	Number of Items	Feb. N=22	April N=22	May/July[a] N=11/12	Results for Composite Objectives: Number of Items	Feb. N=22	April N=22	May/July N=11/12
Prerequisite instructional objectives								
Numerousness					2	68	61	82/92
Writes 0-99	1	45	32	64/92				
Represents 0-99	1	91	91	100/91				
Ordering, place value					2	30	70	32/58
Ordering 0-99	1	36	91	64/75				
Place value 0-99	1	23	50	0/42				
Instructional objectives for S and A topics								
Sentence-writing 0-20, 0-99 (multiple choice)					4	41	61	41/83
Subt-simple separating (11-15)	1	60	91	73/100				
Subt-comparison (0-99)	1	18	14	18/58				
Add-simple joining (0-99)	1	77	91	64/100				
Subt-part part whole-addend (11-15)	1	10	50	9/75				
Sentence-writing 0-20, 0-99 (free response)					4	39	64	64/81
Subt-simple separating (0-99)	1	36	77	82/92				
Subt-part part whole-addend (0-99)	1	5	23	18/67				
Add-part part whole (11-15)	1	68	95	100/92				
Subt-join-addend (11-15)	1	45	60	55/75				

continued

Table 33 (continued)

Description of Objectives	Results for Objectives				Results for Composite Objectives			
	Number of Items	Feb. N=22	April N=22	May/July[a] N=11/12	Number of Items	Feb. N=22	April N=22	May/July N=11/12
Noninstructional objectives								
Problem-solving 0-20, 0-99					4	67	78	80/87
Add-part part whole (0-99)	1	55	68	64/92				
Subt-comparison (11-15)	1	91	77	100/100				
Subt-part part whole-addend (11-15)	1	77	95	91/83				
Subt-join-addend (0-99)	1	45	73	64/75				
Recall of basic facts--speeded test								
Add 0-20					12	44	66	66/94
Subt 0-20					12	40	69	52/84
Algorithms--timed test								
Addition algorithm					24	41	--[b]	--/81[b]
Subtraction algorithm					24	15	--[b]	--/65[b]

[a]Form S was used in April and May; Form V was used in February and July.

[b]Form S did not assess this objective.

Table 34

Percent Correct for Addition and Subtraction Algorithms

Timed Tests by Problem Type for Grade 3, Form V

Item Type	Number of Items	Percent Correct Feb. N=22	Percent Correct July N=12
Addition			
2-digit (without regrouping)	3	86	100
3-digit (without regrouping)	3	93	94
2-digit (with regrouping)[a]	6	49	89
3-digit (with regrouping)	9	16	78
3-digit addends	3	0	44
Subtraction			
2-digit (without regrouping)	3	68	94
3-digit (without regrouping)	3	33	89
2-digit (with regrouping)[a]	6	8	75
3-digit (with regrouping)	12	0	47

[a]Three items are 2-digit $\pm$ 1-digit.

objectives is not high. By mid-third grade, students had yet to master many aspects of either addition or subtraction.

Cross-sectional Growth Across Grades

To portray cross-sectional growth (see Figure 2), five objectives were assessed in all three grades: sentence writing: subtraction-simple separating (11-15); sentence-writing: subtraction-part-part-whole missing addend (11-15); problem solving subtraction-comparison (11-15); recall of basic facts-addition; and recall of basic facts-subtraction. Two composite scales were also administered to grades 1 and 2, and the composite scale ordering, place value was administered at grades 2 and 3.

The cross-sectional data for these scales are presented in Table 35. On each objective, considerable growth is evident. But, as with the longitudinal data, the growth is not uniform or smooth.

Performance of Children in Cognitive Groups Within Grades

Because of the small sample of students we summarize data for children in different cognitive groups by aggregating each group's scores into a total score across all three administrations of the tests. Small sample size also dictates that conclusions drawn from these data must be regarded as tentative and should be the subject of further study.

Grade 1. The relative performance on the test items for children in grade 1 in different cognitive groups is shown in Table 36. There were three children in both Cognitive Groups 1 and 2, but only one child in Cognitive Group 3. The differences in performance for the eight

Table 35

Percent Correct for Common Objectives and Composite

Objectives for Cross-sectional Growth Across Grades 1, 2, and 3

Description of Objective	Percent Correct		
	Feb. Grade 1 N=7	April Grade 2 N=9	May/July[a] Grade 3 N=23
Sentence writing			
Subt-simple separating (11-15)	14	33	87
Subt-part part whole-addend (11-15)	14	11	39
Problem solving			
Subt-comparison	29	56	100
Recall of basic facts--speeded test			
Add 0-20	33	35	78
Subt 0-20	29	30	69
	Feb. Grade 1 N=7		May Grade 2 N=8
Open sentences	29		88
Counting on and back	19		63
	Feb. Grade 2 N=9		May/July[a] Grade 3 N=23
Ordering, place value	6		46

[a]Data gathered on these dates have been combined.

Table 36

Frequency and Percent Correct for Composite Objectives by

Description of Objectives	Number of Items	Cognitive Group 1 (N=3) Frequency	Percent	Trials
Prerequisite instructional objectives				
Numerousness 0-20	2	14	78	18
Ordering 0-20	2	16	89	18
Instructional objectives for the S topics				
Open sentences	2	7	39	18
Sentence-writing 0-20	4	4	11	36
Noninstructional objectives				
Problem solving 0-20	2	12	67	18
Counting	3	2	7	27
Addition facts recall--speeded test	9	24	30	81
Subtraction facts recall--speeded test	9	24	30	81

Table 37

Frequency and Percent Correct for Composite Objectives by

Description of Objectives	Number of Items	Cognitive Group 1 (N=2) Frequency	Percent	Trials
Prerequisite instructional objectives				
Numerousness 0-99	1[a]	4	67	6
Ordering, place value 0-99	2	3	25	12
Instructional objectives for the S and A topics				
Open sentences	2	4	33	12
Sentence-writing 0-20, 0-99 (multiple choice)	4	1	4	24
Sentence-writing 0-20, 0-99 (free response)	4	9	38	24
Algorithms	2	1	8	12
Noninstructional objectives				
Problem solving 0-20, 0-99	4	11	46	24
Counting	3	6	33	18
Addition facts recall--speeded test	12	17	24	72
Subtraction facts recall--speeded test	12	16	22	72

[a]Two items were administered for the numerousness objective; students had so data for this item were discarded.

Cognitive Group for All Administration Times for Grade 1, Form K

Cognitive Group 2 (N=3)			Cognitive Group 3 (N=1)			Total		
Frequency	Percent	Trials	Frequency	Percent	Trials	Frequency	Percent	Trials
17	94	18	4	67	6	35	83	42
15	83	18	6	100	6	37	88	42
7	39	18	2	33	6	16	38	42
9	25	36	2	17	12	15	18	84
13	72	18	4	67	6	29	69	42
16	59	27	2	22	9	20	32	63
65	80	81	11	41	27	100	53	189
49	60	81	8	30	27	81	43	189

Cognitive Group for All Administration Times for Grade 2, Form S

Cognitive Group 2 (N=5)			Cognitive Group 3 (N=2)			Total		
Frequency	Percent	Trials	Frequency	Percent	Trials	Frequency	Percent	Trials
8	58	14	5	84	6	17	67	26
0	0	28	1	8	12	4	8	52
13	46	28	7	58	12	24	46	52
11	20	56	4	17	24	16	15	104
25	45	56	14	58	24	48	46	104
4	14	28	4	33	12	9	17	52
11	20	56	8	33	24	30	29	104
14	33	42	12	67	18	32	41	78
58	35	168	43	60	72	118	38	312
50	30	168	43	60	72	109	35	312

difficulty reading one of the items due to poor quality of the test duplication composite objectives favor the Group 2 children over those in Group 1 on six composites, with some of the differences being quite large. In addition, the Group 2 students increased in performance from February to May over all the objectives, but the Group 1 students improved only in recall of facts. The single Group 3 child fails to fit any pattern.

Grade 2. The relative performance of grade 2 children in different cognitive groups is shown in Table 37. There were two children each in Cognitive Groups 1 and 3 and five in Group 2. In general, Group 3 children performed better than Group 2 children, who in turn performed better than Group 1 children. Some of the differences are striking; for example, open sentences (58%-46%-33%) and addition facts (60%-35%-24%). However, there is one anomaly. For the four problem-solving items, the Group 1 children did better than either of the other groups (46% to 20% to 33%). However, since these children were low on facts, algorithms, and counting skills, the results suggest that they found answers to the verbal problems using other strategies. The children with better arithmetic skills (but not close to mastery) may have attempted to use those skills to solve the problems, but made errors. This explanation is further substantiated by the decrease in performance of Group 1 children on those items as the year progressed and as their arithmetic skills improved.

Grade 3. Results for the grade 3 students in different cognitive groups are striking but somewhat ambiguous (see Table 38). The Group 5-6 children performed better on all objectives than any other group, and Group 2 children were lower than other groups on all the objectives. However, Groups 3 and 4 did not differ consistently. Obviously, the differing characteristics of these two groups are not related to

Table 38

Frequency and Percent Correct for Composite Objectives by Cognitive Group

for All Administration Times for Grade 3

Description of Objectives	Number of Items	Cognitive Group 2 (N=4)			Cognitive Group 3 (N=8)			Cognitive Group 4 (N=8)			Cognitive Group 5,6 (N=7)			Total		
		Frequency	Percent	Trials	Frequency	Percent	Trials	Frequency	Percent	Trials	Frequency	Percent	Trials	Frequency	Percent	Trials
Prerequisite instructional objectives																
Numerousness 0-99	2	15	63	24	29	73	40	24	67	36	29	85	34	97	72	134
Ordering, place value 0-99	2	8	33	24	20	50	40	14	39	36	23	68	34	65	49	134
Instructional objectives for the S and A topics																
Sentence-writing 0-20, 0-99 (multiple choice)	4	21	44	48	42	53	80	39	54	72	44	65	68	146	54	268
Sentence-writing 0-20, 0-99 (free response)	4	25	52	48	49	61	80	37	51	72	44	65	68	155	58	268
Noninstructional objectives																
Problem solving 0-20, 0-99	4	34	71	48	58	73	80	55	76	72	58	85	68	205	76	268
Addition algorithms--timed test[a]	24	30	31	96	134	51	264	111	51	216	172	72	240	447	55	816
Subtraction algorithms--timed test[a]	12	12	13	96	78	30	264	55	25	216	122	51	240	267	33	816
Addition facts recall--speeded test	12	65	45	144	152	63	240	133	62	216	162	79	204	512	64	804
Subtraction facts recall--speeded test	12	61	42	144	135	56	240	126	58	216	154	75	204	476	59	804

[a]This objective was assessed in February for 22 students representing all cognitive groups and in May for 12 students in all groups except 2. It was not assessed in April.

differences in performance. Most of the differences between the Group 5-6 and the Group 2 children are large (selecting sentences 65% to 44%, ordering 68% to 33%, subtraction algorithms 51% to 13%, and so forth).

Performance of Children in Cognitive Groups Across Grades

Performance data from Tables 35, 36, and 37 for children in cognitive capacity Groups 1, 2, and 3 are compared in this section. (Children in Groups 4 and 5-6 are only in grade 3.)

Group 1. The performance of children in this group at grades 1 and 2 is shown in Table 39. In general, the performance of these children at both grades is consistent with their capacity. Only for ordering small numbers is their performance adequate. More striking, there is little difference in performance between grades. There is a marked difference (7% to 33%) only on the counting items, although performance is very low.

Group 2. Children in this capacity group are at all three grade levels. The comparative data for these children are presented in Table 40. Performance gains by grade are apparent, but in most cases very modest. For example, performance on solving open sentences goes from 39% to 46% from grade 1 to grade 2 and performance on writing sentences (free response) from 45% to 52% from grade 2 to grade 3. Only for problem solving (0-99) was there a large gain (20% to 71%). Also, there is a large decrease in performance from grade 1 to grade 2 on recall of both addition and subtraction facts. The decrease is undoubtedly due to the increased number of facts and decreased time for response over forms. This clearly suggests that the high performance at grade 1 was not due to having committed the facts to memory. Also, it should be

Table 39

Frequency and Percent Correct for Composite Objectives

for Cognitive Group 1 for All Administration Times Across Grades

Description of Objectives	Grade 1 (N=3)			Grade 2 (N=2)		
	Frequency	Percent	Trials	Frequency	Percent	Trials
Prerequisite instructional objectives						
Numerousness 0-20	14	78	18	--	--	--
Numerousness 0-99	--	--	--	4	67	6
Ordering 0-20	16	89	18	--	--	--
Ordering, place value 0-99	--	--	--	3	25	12
Instructional objectives for the S and A topics						
Open sentences	7	39	18	4	33	12
Sentence-writing 0-20	4	11	36	--	--	--
Sentence-writing 0-20, 0-99 (multiple choice)	--	--	--	1	4	24
Sentence-writing 0-20, 0-99 (free response)	--	--	--	9	39	24
Algorithms	--	--	--	1	8	12
Noninstructional objectives						
Problem solving 0-20	12	67	18	--	--	--
Problem solving 0-20, 0-99	--	--	--	11	46	24
Counting	2	7	27	6	33	18
Addition facts recall--speeded test	24	30	81	17	24	72
Subtraction facts recall--speeded test	24	30	81	16	22	72

Table 40

Frequency and Percent Correct for Composite Objectives

for Cognitive Group 2 for All Administration Times Across Grades

Description of Objectives	Grade 1 (N=3)			Grade 2 (N=6)			Grade 3 (N=4)		
	Frequency	Percent	Trials	Frequency	Percent	Trials	Frequency	Percent	Trials
Prerequisite instructional objectives									
Numerousness 0-20	17	94	18	--	--	--	--	--	--
Numerousness 0-99	--	--	--	8	58	14	15	63	24
Ordering 0-20	15	83	18	--	--	--	--	--	--
Ordering, place value 0-99	--	--	--	0	0	28	8	33	24
Instructional objectives for the S and A topics									
Open sentences	7	39	18	13	46	28	--	--	--
Sentence-writing 0-20	9	25	36	--	--	--	--	--	--
Sentence-writing 0-20, 0-99 (multiple choice)	--	--	--	11	20	56	21	44	48
Sentence-writing 0-20, 0-99 (free response)	--	--	--	25	45	56	25	52	48
Algorithms	--	--	--	4	14	28	--	--	--
Noninstructional objectives									
Problem solving 0-20	13	72	18	--	--	--	--	--	--
Problem solving 0-20, 0-99	--	--	--	11	20	56	34	71	48
Counting	16	59	27	14	33	42	--	--	--
Addition facts recall--speeded test	65	80	81	58	35	168	65	45	144
Subtraction facts recall--speeded test	49	60	81	50	30	168	61	42	144
Addition algorithms	--	--	--	--	--	--	30	31	96
Subtraction algorithms	--	--	--	--	--	--	12	13	96

noted that, at grade 3, these children had not learned to use the addition and subtraction algorithms with any facility. Again, the overall performance of these students reflects capacity more than grade level.

Group 3. The data for children in this cognitive capacity group at grades 2 and 3 are compared in Table 41. There are some important differences in performance at this grade. For example, performance increases from 8% to 50% on ordering, place value; from 17% to 53% on sentence-writing (multiple choice); and from 33% to 73% on problem solving. However, for all other scales, performance is similar across grades. Overall performance is fair and students show some facility with the addition algorithm, but not with the subtraction algorithm.

Relationship of Performance on Algorithms to Strategies Used to Solve Problems

One general goal of instruction on addition and subtraction is to have students solve verbal problems (such as those presented in Chapter 3) by using an addition or subtraction algorithm. We now examine the relationship of the performance of the third-grade children on the timed algorithm problems to the strategies they used to solve verbal problems that could be done using those algorithms. The strategy data were collected in the interview study discussed in Chapter 3. We were particularly interested in examining whether or not students who had learned to use the addition and subtraction algorithms chose to use them when solving such verbal problems.

For addition problems requiring no regrouping, at Time 1, students attempted 62 items and got 57 correct (92%); in July, students answered

Table 41

Frequency and Percent Correct for Composite Objectives

for Cognitive Group 3 for All Administration Times Across Grades

Description of Objectives	Grade 2 (N=2)			Grade 3 (N=8)		
	Frequency	Percent	Trials	Frequency	Percent	Trials
Prerequisite instructional objectives						
Numerousness	5	84	6	29	73	40
Ordering, place value 0-99	7	8	12	20	50	40
Instructional objectives for the A and S topics						
Open sentences	7	58	12	--	--	--
Sentence-writing 0-20, 0-99 (multiple choice)	4	17	24	42	53	80
Sentence-writing 0-20, 0-99 (free response)	14	58	24	49	61	80
Algorithms	4	33	12	--	--	--
Noninstructional objectives						
Problem solving 0-20	8	33	24	58	73	80
Counting	12	67	18	--	--	--
Addition facts recall--speeded test	43	60	72	152	63	240
Subtraction facts recall--speeded test	43	60	72	135	56	240
Addition algorithms	--	--	--	134	51	264
Subtraction algorithms	--	--	--	78	30	264

NOTE: The one group 3 child at grade 1 was not included in this comparison.

all 36 items they attempted correctly. With one exception (student 517), these students knew how to add two-digit numbers without regrouping. However, on the interviews at Time 1, students used algorithms only 59% of the time (54% correctly). At Times 2 and 3, the percent of use increased but only to 79% and 72%, respectively.

The data for addition with regrouping at Time 1 were similar; students attempted 95 items and got 66 correct (69%). By Time 2, they attempted 74 items and got 66 correct (89%). Thus, although students had some difficulty with regrouping at the start of the year, by July, all the students could add with regrouping (with the exception of one student who made six errors in six problems).

The interview data show that, in spite of this high level of performance, many students did not use the algorithms to solve verbal addition problems. At Time 1, about half (54%) of the children tried using an algorithm (46% correctly). At Time 2, this had changed to 60% using an algorithm (48% correctly), and by Time 3, 78% used an algorithm with no errors.

For subtraction without regrouping, results of a comparison between performance on three achievement items and strategies used on the four verbal subtraction problems were similar. At Time 1, students attempted 55 items and got 45 correct (82%). By Time 2, 34 of 36 attempts were correct (94%). In fact, only one student made any errors in July. One can conclude that these students were able to subtract without regrouping. However, on the four verbal subtraction problems only 14% of the strategies used were algorithmic (only 9% correct) at the start of the year. At Time 2, this had increased to 25% and finally to 34% by

Time 3. Furthermore, over half of the total attempts (59%) were on Task 2 (simple separate), the most obvious subtraction problem.

The same pattern, but more pronounced, occurred for subtraction with regrouping. At Time 1, students attempted 38 items and got only 12 correct (32%). Many children managed to complete only the first six no-regrouping items in this timed test, so there was no real measure of their capability. It is hard to imagine why they were so slow. One must assume that they would have been unable to do the regrouping problems had they attempted them. By the second administration (July), students attempted 66 items and got 55 correct (82%). At this time, only two students made more than one error on the six problems. Thus, although there was evidence that students had considerable difficulty in subtracting with regrouping in February, by the end of the autumn term most were capable of using a subtraction algorithm.

Again, however, in spite of knowing the algorithmic procedures for subtraction, most children did not attempt to use them to solve verbal problems. On the first interview, students used algorithms on only 13% of the items (5% correct). On the second interview, this had increased to 23% (11% correct), and by the third interview, it was 35% (26% correct). As with subtraction no-regrouping, most of the attempts were on the simple separating tasks (44%).

On this last set of verbal problems, the cognitive Group 2 students made the most total attempts to use algorithms (35% of the time), even though they got no items correct on the achievement test and made the most errors (only 10% correct) on verbal problems. In contrast, the Group 5-6 students attempted to use algorithms only 22% of the time.

The relationship between proficiency in performing addition and subtraction algorithms and using the algorithms to solve verbal problems is interesting. Most of the third-grade students used other strategies (counting, fingers, and so forth) until they became confident in using algorithms. However, Group 2 children who had not acquired other strategies to solve these problems tended to use the taught algorithms even though they were not proficient in their use. This suggests that most third graders recognized that these problems could be solved using algorithms but chose to use other familiar strategies. The problem structures (verbal semantics) clearly influenced how the students worked the problems. In fact, the semantics seemed to be more important than the realization that the problems could be done algorithmically.

Summary and Conclusions

The overall picture these data presents is of children struggling to learn the complex arithmetic skills associated with addition and subtraction and to use those skills to solve verbal problems. The children had difficulty with place value even though they correctly solved three-digit problems. Work on algorithms improved even though basic facts were weak. With little arithmetic competence, students correctly solved some simple verbal problems.

Children who were identified as being in a particular cognitive group, with one important exception, performed differently than children in other groups within each grade. The exception was the lack of consistent differences between Groups 3 and 4 at grade 3. Again, it should be noted that Group 3 at grade 3 also did not differ on the interview tasks (see Chapter 3) and only differed on transitivity on the cognitive tasks (see Chapter 2). Overall, however, it is apparent that children who differed in cognitive-processing capacity (Group 1, Group 2, Groups 3 and 4, and Group 5-6) performed differently regardless of specific objectives, instruction over time, or grade.

Although it cannot be denied that teaching or experience accounts for some differences in the children's performance on standard addition and subtraction tasks, it is striking that the actual level of performance appears to be consistent with capacity. Differences in performance between groups and within groups across grades are differences one would expect, based on the nature of the groups (e.g., quantitative skills, memory capacity).

Chapter 5

COGNITIVE-PROCESSING CAPACITY AND CLASSROOM INSTRUCTION

The fifth and last study in this series is reported in this chapter. Its purpose was to examine the question, Do children who differ in cognitive capacity receive different instruction?

Method

Sample

A sample of 35 children from the population used in studies 3 and 4 in this series (see chapter 2) was observed during instruction over a three-month period in 1980 (February 27 through May 28). The number of children observed in each cognitive group, class, and grade is shown in Table 42.

Our attempt was to determine the way in which aspects of content influence certain teacher behaviors during instruction and in turn how these actions affect pupil outcomes. In particular, the extent to which children are engaged in learning mathematics was examined. To do this we developed a model of classroom instruction in which "content segmentation and sequencing" and "content structuring" were hypothesized to influence teacher planning, which in turn influences classroom organization, the allocation of instructional time, verbal interactions within classroom, and, eventually, pupil engaged time (see Romberg, Small, & Carnahan, 1979, for a complete explication of the model). To test this model, data were gathered on various components of the model in actual classroom settings for several periods of time (see Romberg,

Table 42

Children in Each Cognitive Group, Class and Grade Used in the Observation Study

Cognitive Group	Sandy Bay Infant School		Waimea Heights Primary School		
	Class		Class		
	1	2	3	4	5
	Grade 1	Grade 2	Grade 3	Grade 3	Grade 3
1	2	2			
2	3	4		3	
3	1	2	2	2	2
4			2	2	2
5-6			3	1	2
Totals	6	8	7	8	6

Small, Carnahan, & Cookson, 1979, for a description of coding procedures used as well as detailed explanations of coding categories). With such data the relationship of the model to classroom instruction can be examined.

Summary of the Coding Procedure and Aggregation of Data

Data were collected on content covered and on certain teacher and pupil behaviors involved in the teaching and learning of mathematics using two procedures (complete details appear in Romberg, Collis, Buchanan, & Romberg, 1982).

Content. First, to estimate time spent on various mathematics objectives, the teachers were asked to log the number of minutes of instruction in nine content areas spent for each target child. Seven of the nine areas dealt with aspects of learning to add and subtract. The "other arithmetic" area included time spent on both multiplication and division activities, and "other mathematics" encompassed all other activities such as measurement, fractions, or geometry.

Classroom observation. Three trained observers gathered the data. These were the same persons who gathered data in Studies 3 and 4. One observer worked at Sandy Bay Infant School and observed both the grade 1 and grade 2 classes. The other two worked at Waimea Heights Primary School, where one observed two classes. Each observer was able to observe instruction in a class approximately 24 days during the observation period. At the schools, the observers sat in a class and over time became fixtures who did not distract either teacher or children.

Data on pupil and teacher behavior were gathered using an observation coding form. The exact nature of the data collected and the method used to gather it are described fully in the manual produced by the project staff to train observers (Romberg, Small, Carnahan, & Cookson, 1979).

In brief, student and teacher verbal behaviors were observed in each class on a sample of days. A time-sampling procedure was used in which each of six to eight target students was observed in a particular sequence at different moments throughout the observation period. The sequence in which the students were observed was fixed prior to the beginning of the observation period and was invariant while observations were taking place. The teacher's behavior was coded for instances of relevant verbal behavior each time a target student was observed. The observation of all six to eight students (along with the teacher six to eight times) represented a coding cycle. It was estimated that one minute was needed: (a) to observe the target student's behavior, (b) to observe the teacher, (c) to observe organizational aspects of the classroom, and (d) to code the appropriate categories on the observation form. The behavior to be coded consisted <u>only</u> of those activities the teacher and pupil were involved in precisely at the beginning of the one-minute time interval. Through this process, observer bias in sampling moments is minimized. The coding categories were used to record a description of what was occurring at that one instant for both the target student and the teacher. In this way, a series of codings was obtained that gave a running account of what took place in the classroom for a particular observation period.

The observation for a class session began when the mathematics instruction began and ended when the mathematics instruction for that class session ended. The beginning and ending of the observation period did not always coincide with the beginning and ending of mathematics instruction as scheduled. As a result, two measures of time involved in mathematics class were obtained. Available time represented the scheduled time period in which mathematics instruction was to take place. Actual time, on the other hand, represented the amount of time mathematics instruction occurred. In most cases, the amount of time observing coincided closely with the measure of available time.

Data aggregation and analysis. The basic observational data were aggregated in the form of frequency counts for each behavior category coded. For purposes of interpretation, the proportional occurrence of each behavior (based on total observed instances) is used. Data were aggregated separately for each class for the total period. The data give an overall picture of the teaching of mathematics in each class and yield estimates of how instructional factors affect engagement rates.

The observational data were summarized in terms of three categories: pupil actions, teacher behaviors, and teacher behavior-pupil engagement interactions. Pupil actions were summarized in terms of engaged time; if engaged, whether it was on content or directions; grouping; interactions; and if interacting, with whom. Teacher behaviors were summarized in terms of interactions, speaking to group, speaking on content or directions, questions, feedback and type of explanations. Interactions of teacher behaviors and pupil engagement were summarized in terms of whether or not pupils were engaged when the teacher was speaking, speaking to groups, listening, no teacher interactions, questioning, and providing information.

The plan for the analysis of the observational data was based on the fact there were three primary dimensions in the study: grade, class, and cognitive group of the pupils. The raw data are observed minutes. Thus, the number of minutes and percent of time are aggregated in this analysis in five ways. First, we have aggregated data for all pupils with respect to grade. Second, since three different classrooms were observed in grade 3, we have examined the data by class. Third, we examined the data for all students with respect to cognitive group. Fourth, we have examined the data by cognitive level within grade. Finally, we present the data in terms of cognitive level within class.

Results and Discussion

Content Covered

Table 43 presents the percentage of total time teachers spent on various content areas. These data reflect the curricular emphasis common in these grades. Almost half of the time is spent on addition and subtraction. The emphasis obviously varies across grades. In grade 1, the highest percentage is on addition facts, numerousness and counting. In grade 2, basic facts for both addition and subtraction are still emphasized as are counting skills. In grade 3, most of the emphasis is on computational algorithms. The only disappointing percentages are the little time spent on writing sentences and finding solutions to verbal problems. However, this differential emphasis is program-related, not child-related. For example, the reduction in percent of time spent on counting at grade 3 was not matched by the

Table 43

Percentage of Time Spent on Mathematical Content Area

by Grade--Teacher Log Data

Content Area	Grade 1 (24 days, 50-60 min/day)	Grade 2 (25 days, 50-55 min/day)	Grade 3 (111 days, 30 min/day) (3 classes combined)
Numerousness	14.3	6.4	4.5
Ordering	5.2	5.6	2.1
Basic facts	15.5	13.3	4.0
(Add)	(14.7)	(6.8)	(3.1)
(Subtract)	(.8)	(6.5)	(.9)
Problem solving	2.6	1.4	4.2
Sentence writing	.8	.8	3.1
Algorithms	0	3.1	24.0
(Add)	(0)	(3.1)	(13.4)
(Subtract)	(0)	(0)	(10.6)
Counting	9.3	12.4	1.4
TOTAL Addition and subtraction	47.7	50.2	44.3
Other arithmetic	13.2	16.8	15.6
Other math	39.1	33.0	41.1

children's failure to use this technique to solve problems. In fact, it can be argued that the structure of the program at grade 3 (emphasis on algorithms) presumes that children have mastered most of the prerequisites (like counting and basic facts) and have acquired a high level of reasoning about numbers (are in Cognitive Group 5-6). This is, of course, at odds with the data on these students presented in the last two chapters.

Also, this description is fair in terms of the content included in the mathematics curricula in these schools, but it fails to capture important structural features of those programs. The program in Sandy Bay Infant School was filled with manipulative materials, many opportunities to explore independently or in small groups, learning stations, etc., and no basal text was used. However, in the third grades at Waimea Heights, a single text was followed and most activities involved paper-and-pencil seatwork.

Pupil Actions

Grade. The data on pupil actions by grade are presented in Table 44. Significant engagement rate and grouping differences are apparent across grades. Both are likely due in part to the differences in the structure of the curriculum in the schools. The high amount of time spent on small-group and individual activities in grades 1 and 2 (85% and 68%, respectively) is consistent with the manipulative-based, learning station approach at the Sandy Bay School. Similarly, 70% of the time spent in large-group instruction at grade 3 is consistent with the text-based instruction used at the Waimea Heights School. However, it is interesting to note that the largest difference in engagement is

Table 44

Observed Minutes and Percent of Time

of Pupil Actions by Grade

Pupil Action	Grade 1		Grade 2		Grade 3	
	Minutes	%	Minutes	%	Minutes	%
Engagement						
Engaged time	559	55	771	71	1369	77
Off-task time	449	45	317	29	403	23
Types of engagement						
Content	488	89	656	86	1149	88
Directions	62	11	107	14	164	12
Grouping						
Individual	302	30	165	15	11	1
Small group	553	55	583	53	524	29
Large group	156	15	343	31	1259	70
Interactions						
Target speaking	62	6	51	5	105	6
Target listening	91	9	163	15	279	15
None	858	85	880	80	1427	79
Interaction other party						
Teacher	99	65	161	76	296	78
Pupil	48	31	36	17	77	20
Other adult	6	4	16	8	6	2

between grade 1 and grade 2 students, who are following the same curriculum (in the same school). In fact, it was observed that in grade 1, many children spent considerable time waiting for instructions about what to do next when they had completed an activity. By grade 2, this behavior was observed less frequently. Many students now proceeded to the next task with little hesitation. Part of this change is probably due to increased student maturity or familiarity with behavior expectations of the system and part is probably due to the particular teacher a student had.

Class. Grade 3 data were further subdivided into pupil actions by class as shown in Table 45. For comparative purposes, the data for grade 1 (class 1) and grade 2 (class 2) are shown again. Classes 3, 4, and 5 are all in grade 3. Class 4 is clearly different from the other two classes. Pupils in that class were off-task more of the time. Furthermore, if they were engaged, they were more likely to be engaged in receiving directions, and if interacting they were more likely to be interacting with other pupils. Differences in grouping are not likely a function of curriculum, since all third-grade classes are similar on that dimension. The large differences in engagement and interactions are probably a function of the teacher.

Cognitive group. The number of minutes and percent of time coded for the five pupil action categories for all students in the cognitive groups are presented in Table 46. Overall, the percent of engaged time steadily increases across cognitive groups. Also, differences in grouping are striking, with percent of time in large-group instruction varying from 21% for Group 1 to 68% for Group 6 children. All other differences in percentage of time coded for the pupil action categories

Table 45

Observed Minutes and Percent of Time of Pupil Actions by Class

Pupil Action	Grade 1-Class 1		Grade 2-Class 2		Grade 3					
					Class 3		Class 4		Class 5	
	Minutes	%	Minutes	%	Minutes	%	Minutes	%	Minutes	%
Engagement										
Engaged time	559	55	771	71	402	98	650	64	317	90
Off-task time	449	45	317	29	8	2	358	36	37	10
Types of encouragement										
Content	488	89	656	86	364	95	496	79	289	97
Directions	62	11	107	14	21	5	135	21	8	3
Grouping										
Individual	302	30	165	15	6	1	0	0	5	1
Small group	553	55	583	53	101	24	247	25	176	47
Large group	156	15	343	31	317	75	750	75	192	51
Interactions										
Target speaking	62	6	51	5	24	6	52	5	29	8
Target listening	91	9	163	15	112	26	127	13	40	11
None	858	85	880	80	289	68	835	82	303	81
Interaction: other party										
Teacher	99	65	161	76	122	92	119	67	55	81
Pupil	48	31	36	17	10	8	57	32	10	15
Other adult	6	4	16	8	1	1	2	1	3	4

Table 46

Observed Minutes and Percent of Time of Pupil Actions by Cognitive Group

Pupil Actions	Cognitive Group 1		Cognitive Group 2		Cognitive Group 3		Cognitive Group 4		Cognitive Group 5	
	Minutes	%	Minutes	%	Minutes	%	Minutes	%	Minutes	%
Engagement										
Engaged time	420	64	850	65	721	70	331	76	377	87
Off-task time	237	36	460	35	310	30	106	24	56	13
Types of engagement										
Content	361	86	690	83	634	90	282	88	326	91
Directions	57	14	140	17	68	10	37	12	31	9
Grouping										
Individual	167	25	201	15	104	10	0	0	6	1
Small group	356	54	593	45	444	43	129	29	138	31
Large group	135	21	510	39	496	48	317	71	300	68
Interactions										
Target speaking	37	6	61	5	63	6	19	4	38	9
Target listening	76	12	164	12	162	15	62	14	69	16
None	545	83	1090	83	825	79	367	82	338	76
Interaction: other party										
Teacher	80	71	167	74	162	73	67	83	80	78
Pupil	24	21	46	20	55	25	14	17	22	21
Other adult	9	8	12	5	6	3	0	0	1	1

are not striking or of practical interest. However, these differences in engagement and grouping are clearly confounded by the grade, class, and teacher effects described earlier. This is due to the fact that at grade 1, five of six children observed were in cognitive Groups 1 and 2; at grade 2, six of eight children were in Groups 1 and 2; and at grade 3, 12 of 21 children were in Groups 4, 5, and 6.

Cognitive level within class. To answer the question of whether or not children with different cognitive capacities received different instruction, the data for children within each class are presented. The data for children of different cognitive levels within Class 1 (grade 1) are presented in Table 47. Only the difference in time pupils interact with other pupils is of interest and then only between Group 1 and Group 3 children (24% to 45%).

The data for Class 2 (grade 2) children in different cognitive groups are presented in Table 48. As with grade 1, the only observable difference is in pupil interactions with other pupils (17% for Group 1 children and 32% for Group 3 children).

Tables 49, 50, and 51 contain the within-class data for children in different cognitive groups for the three third-grade classes. Class 3 and Class 5 show high engagement on content with virtually no differences between students. Class 4, on the other hand, exhibits much lower engagement with more time on directions for all students. Again, only pupil interactions with other pupils vary by cognitive level (31% for Group 2 children to 46% for Group 5-6 children).

Summary. Overall, these data suggest that differences in grouping of students are due to grade (structure of the curriculum) or teacher. Grade 1 and grade 2 children often worked in small groups and

Table 47

Observed Minutes and Percent of Time of Pupil Actions

by Cognitive Group Within Class 1, Grade 1

Pupil Action	Cognitive Group 1		Cognitive Group 2		Cognitive Group 3	
	Minutes	%	Minutes	%	Minutes	%
Engagement						
Engaged time	260	60	189	51	110	54
Off-task time	174	40	181	49	94	46
Types of engagement						
Content	230	89	159	87	99	90
Directions	28	11	23	13	11	10
Grouping						
Individual	129	30	119	32	54	26
Small group	235	54	197	53	121	59
Large group	70	16	56	15	30	15
Interactions						
Target speaking	26	6	24	6	13	6
Target listening	41	9	30	8	20	10
None	368	85	318	85	172	84
Interaction: other party						
Teacher	46	70	36	67	17	52
Pupil	16	24	17	31	15	45
Other adult	4	6	1	2	1	3

Table 48

Observed Minutes and Percent of Time of Pupil Actions

by Cognitive Group Within Class 2, Grade 2

Pupil Action	Cognitive Group 1		Cognitive Group 2		Cognitive Group 3	
	Minutes	%	Minutes	%	Minutes	%
Engagement						
Engaged time	160	72	399	72	212	69
Off-task time	63	28	158	28	96	31
Types of engagement						
Content	131	82	336	85	189	90
Directions	29	18	57	15	21	10
Grouping						
Individual	38	17	82	15	45	14
Small group	121	54	294	53	168	54
Large group	65	29	179	32	99	32
Interactions						
Target speaking	12	5	20	4	19	6
Target listening	35	16	84	15	44	14
None	177	79	454	81	249	80
Interaction: other party						
Teacher	34	72	87	84	40	65
Pupil	8	17	8	8	20	32
Other adult	5	11	9	9	2	3

Table 49

Observed Minutes and Percent of Time of Pupil Actions

by Cognitive Group Within Class 3, Grade 3

Pupil Action	Cognitive Group 3		Cognitive Group 4		Cognitive Group 5,6	
	Minutes	%	Minutes	%	Minutes	%
Engagement						
Engaged time	144	98	80	96	178	99
Off-task time	3	2	3	4	2	1
Types of engagement						
Content	127	93	76	96	161	95
Directions	10	7	3	4	8	5
Grouping						
Individual	5	3	0	0	1	0
Small group	33	22	20	23	48	26
Large group	114	75	67	77	136	74
Interactions						
Target speaking	8	5	2	2	14	8
Target listening	47	31	16	18	49	29
None	98	67	69	79	122	66
Interaction: other party						
Teacher	52	95	16	94	54	89
Pupil	3	5	1	6	6	10
Other adult	0	0	0	0	1	1

Table 50

Observed Minutes and Percent of Time of Pupil Actions

by Cognitive Group Within Class 4, Grade 3

Pupil Action	Cognitive Group 2		Cognitive Group 3		Cognitive Group 4		Cognitive Group 5,6	
	Minutes	%	Minutes	%	Minutes	%	Minutes	%
Engagement								
Engaged time	262	68	151	69	148	62	89	67
Off-task time	121	32	101	40	92	38	44	33
Types of engagement								
Content	195	76	119	83	112	77	70	80
Directions	60	24	24	17	33	23	18	20
Grouping								
Individual	0	0	0	0	0	0	0	0
Small group	102	27	62	25	51	21	32	24
Large group	275	73	187	75	187	79	101	76
Interactions								
Target speaking	17	4	14	6	8	3	13	10
Target listening	50	13	36	14	30	12	11	8
None	318	83	204	80	203	84	110	82
Interaction: other party								
Teacher	44	66	33	67	29	76	13	54
Pupil	21	31	16	33	9	24	11	46
Other adult	2	3	0	0	0	0	0	0

Table 51

Observed Minutes and Percent of Time of Pupil Action

by Cognitive Group Within Class 5, Grade 3

Pupil Action	Cognitive Group 3		Cognitive Group 4		Cognitive Group 5,6	
	Minutes	%	Minutes	%	Minutes	%
Engagement						
Engaged time	104	87	103	90	110	92
Off-task time	16	13	11	10	10	8
Types of engagement						
Content	100	98	94	99	95	95
Directions	2	2	1	1	5	5
Grouping						
Individual	0	0	0	0	5	4
Small group	60	48	58	48	58	46
Large group	66	52	63	52	63	50
Interactions						
Target speaking	9	7	9	7	11	9
Target listening	15	12	16	13	9	7
None	102	81	95	79	106	84
Interaction: other party						
Teacher	20	83	22	85	13	72
Pupil	1	4	4	15	5	28
Other adult	3	13	0	0	0	0

individually for mathematics instruction while large group work was common in grade 3. Differences in engaged time are likely due to teachers or to students' familiarity with the instructional pattern. Only pupil interactions with other pupils are plausibly due to childrens' cognitive level (with children in higher groups more likely to interact with others), but this behavior only occurs where such interactions are allowed and even then is infrequent.

Teacher Behaviors

The data for number of minutes and percent of time teachers at various grade levels engaged in the behaviors coded are discussed first. Then, the teacher behaviors are presented by class, by cognitive group, and by cognitive group/class interactions.

Grade. The data on teacher behaviors by grade are presented in Table 52. The differential time teachers spent explaining or giving directions vs. content varies with grade level and is consistent with program expectations discussed earlier. Time spent on directions is inversely related to grade level.

Class. The data on teacher behaviors by class within grade 3 are shown in Table 53. The differences in speaking about content appear to be teacher or class specific. The differences between the first-grade teacher and two of the third-grade teachers on content are large. For example, the teacher of class 1 (grade 1) spent 51% of the observed time speaking on content while the teacher in class 3 (grade 3) spent 82% on content. Another grade 3 teacher (class 4) spent 57% of the time on content. However, the percent of time teachers explain directions appears to be a grade effect or curriculum effect, since all three grade

Table 52

Observed Minutes and Percent of Time

of Teacher Behaviors by Grade

Teacher Behavior	Grade 1		Grade 2		Grade 3	
	Minutes	%	Minutes	%	Minutes	%
Interaction						
Listening	187	17	206	18	216	11
Speaking	640	58	677	59	1238	64
None	276	25	254	22	485	25
Speaking/large group	91	14	209	31	313	25
Speaking/small group	82	13	65	10	227	18
Speaking/individual	467	73	402	59	697	56
Speaking/content	367	57	404	60	823	66
Speaking/directions	268	42	256	38	347	28
Low-level questions	135	12	157	14	338	17
Direction-related questions	33	3	29	3	199	10
No feedback	1006	91	1035	91	1819	94
Feedback/individual	79	90	89	94	109	92
Low-information feedback	97	100	101	98	115	93
High-information feedback	0	0	2	2	9	7
Explaining content	130	12	117	10	323	17
Explaining directions	235	21	228	20	165	9

Table 53

Observed Minutes and Percent of Time of Teacher Behaviors by Class

Teacher Behavior	Grade 1-Class 1		Grade 2-Class 2		Grade 3					
					Class 3		Class 4		Class 5	
	Minutes	%	Minutes	%	Minutes	%	Minutes	%	Minutes	%
Interaction										
Listening	187	17	206	18	55	12	129	12	32	8
Speaking	640	58	677	59	290	63	681	62	267	71
None	276	25	254	22	116	25	294	27	75	20
Speaking/large group	91	14	209	31	128	44	134	20	51	19
Speaking/small group	82	13	65	10	41	14	107	16	79	29
Speaking/individual	467	73	402	59	121	42	439	64	137	51
Speaking/content	367	57	404	60	239	82	391	57	193	71
Speaking/directions	268	42	256	38	45	15	240	35	62	23
Low-level questions	135	12	157	14	94	20	172	16	72	19
Direction-related questions	33	3	29	3	22	5	125	11	52	14
No feedback	1006	91	1035	91	434	94	1025	93	360	95
Feedback/individual	79	90	89	94	24	92	71	93	14	83
Low-information feedback	97	100	101	98	23	82	77	99	15	83
High-information feedback	0	0	2	2	5	18	1	1	3	17
Explaining content	130	12	117	10	96	21	139	13	88	23
Explaining directions	235	21	228	20	26	6	126	11	13	3

3 teachers spend less time (6%, 11%, and 3%) than either the grade 1 (21%) or grade 2 (20%) teachers.

Cognitive group. The number of minutes and percent of time coded for six teacher behavior categories are presented in Table 54. Overall, three differences are striking across cognitive groups. First, the percent of time speaking to individual children decreases from 67% for Group 1 children to 53% for Group 5-6 children. Second, the percent of time teachers spent speaking about directions shifts from 39% for Group 1 children to 27% for Group 5-6 children. Similarly, the percent of time explaining directions decreases from 22% for Group 1 children to 6% for Group 5-6 children. However, all of these differences are undoubtedly confounded by grade level.

Cognitive group within class. Data on the percent of time teacher behaviors were observed in each class in relation to students in different cognitive groups are not presented here. For four of the classes (1, 2, 3, and 4), there were no striking differences in terms of time spent interacting with different children. Only one important difference was found. In class 5, the time that the teacher spoke on content decreased across Groups 1-5/6 from 82% to 66%.

In summary, although teachers varied considerably in their behavior, differences seem due more to grade, or individual teaching style, or grouping patterns within classes than to differential treatment of students with various cognitive capacities. Teachers may treat some students different from others, but these data suggest that cognitive capacity is not the basis for such differentiation.

Table 54

Observed Minutes and Percent of Time of Teacher Behaviors by Cognitive Group

Teacher Behavior	Cognitive Group 1		Cognitive Group 2		Cognitive Group 3		Cognitive Group 4		Cognitive Group 5,6	
	Minutes	%	Minutes	%	Minutes	%	Minutes	%	Minutes	%
Interaction										
Listening	127	19	231	16	139	13	50	10	62	14
Speaking	394	60	837	58	722	66	310	60	292	64
None	141	21	380	26	235	21	154	30	97	21
Speaking/large group	83	21	190	23	189	26	90	23	81	28
Speaking/small group	47	12	116	14	100	14	56	18	55	19
Speaking/individual	264	67	529	63	433	60	184	59	156	53
Speaking/content	233	59	479	57	475	66	204	66	203	69
Speaking/directions	155	39	327	39	228	31	81	26	80	27
Low-level questions	84	13	178	12	196	18	82	16	90	20
Direction-related questions	12	2	78	5	68	6	51	10	52	11
No feedback	592	83	1334	92	1018	93	481	93	427	94
Feedback/individual	60	91	98	93	70	95	27	87	22	88
Low-information feedback	69	99	114	100	75	94	31	91	24	92
High-information feedback	1	1	0	0	5	6	3	9	2	8
Explaining content	72	11	166	11	182	17	75	15	75	17
Explaining directions	143	22	254	18	164	15	38	7	29	6

Teacher Behavior/Pupil Engagement Interactions

The number of minutes and percent of coded time that teachers performed various actions and children were engaged are reported in this section. As with the previous sections, the data were first aggregated for children differing by grade, then class, cognitive group, and finally, cognitive group within class.

Grade. The data on pupil engagement for various teacher actions by grade are presented in Table 55. Pupil engagement when teachers are speaking increases from 59% in grade 1 to 78% in grade 3. Engagement when teachers are not speaking increases from 50% to 76%. Similarly, pupil engagement when there are no interactions increases from 42% to 78% across grades, as do all engagement rates related to teacher questioning and providing information.

Class. The information on pupil engagement when teachers performed certain actions is presented for all five classes in Table 56. As would be expected from previous analyses, Class 4 in grade 3 is different from Classes 3 and 5 in grade 3. Engagement rates in Class 4 are lower in all categories than rates in the other two classes. In fact, the grade level effect noted previously is in part a teacher effect, and certainly would be higher for grade 3 if Class 4 were omitted.

Cognitive group. The overall data on time pupils in differing cognitive groups were engaged when teachers were doing different things is reported in Table 57. Many of the differences are striking. First, as cognitive level increases, children increase in engagement when teachers are speaking from 65% of the time to 86%. Second, the pattern across groups is similar regardless of whom the teacher is speaking to, and even when the teacher is not speaking (62% engagement to 89%).

Table 55

Observed Minutes and Percent of Time of Teacher Behaviors and Pupil Engagement by Grade

Interaction	Grade 1		Grade 2		Grade 3	
	Minutes	%	Minutes	%	Minutes	%
Teacher speaking/						
Pupil engaged	356	59	463	70	919	78
Pupil off-task	245	41	197	30	259	22
Pupil engaged when teacher speaking to:						
Individual	253	57	265	68	502	74
Small group	51	67	45	69	160	80
Large group	52	63	153	75	256	86
Not speaking	203	50	308	72	449	76
Pupil engaged when teacher:						
Listening	104	61	151	76	152	72
Pupil engaged when:						
No interactions	99	42	157	69	295	78
Pupil engaged when teacher asks:						
Low-level questions	75	60	108	71	263	81
High-level questions	8	67	19	83	42	95
Questions about directions	15	52	20	69	133	70
Pupil engaged when teacher provides:						
Low-information feedback	44	48	67	68	90	80
Positive feedback	32	54	56	72	52	87
Information about content	83	68	89	77	255	82
Explains directions	131	58	149	67	114	72

Table 56

Observed Minutes and Percent of Time of Teacher Behaviors and Pupil Engagement by Class

Interaction	Grade 1-Class 1		Grade 2-Class 2		Grade 3					
					Class 3		Class 4		Class 5	
	Minutes	%	Minutes	%	Minutes	%	Minutes	%	Minutes	%
Teacher speaking/										
Pupil engaged	356	59	463	70	263	99	437	66	219	88
Pupil off-task	245	41	197	30	2	1	226	34	31	12
Pupil engaged when teacher speaking to:										
Individual	253	57	265	68	109	99	275	63	118	87
Small group	51	67	45	68	31	100	69	59	60	90
Large group	52	63	153	75	123	99	92	74	41	85
Not speaking	203	50	308	72	138	96	213	62	98	94
Pupil engaged when teacher:										
Listening	104	61	51	76	51	98	73	57	28	90
Pupil engaged when:										
No interactions	99	42	157	69	88	95	141	65	66	96
Pupil engaged when teacher asks:										
Low-level questions	75	60	108	71	86	99	116	76	61	92
High-level questions	8	67	19	83	23	100	1	100	18	90
Questions about directions	15	52	20	69	21	100	70	60	42	82
Pupil engaged when teacher provides:										
Low-information feedback	44	48	67	68	22	100	56	73	12	92
Positive feedback	32	54	56	72	20	100	23	82	9	75
Information about content	83	68	89	77	86	100	95	69	74	86
Explains directions	131	58	149	67	24	96	80	66	10	91

Table 57

Observed Minutes and Percent of Time of Pupil Engagement

for Various Interactions, by Cognitive Group

Interaction	Cognitive Group 1		Cognitive Group 2		Cognitive Group 3		Cognitive Group 4		Cognitive Group 5,6	
	Minutes	%	Minutes	%	Minutes	%	Minutes	%	Minutes	%
Teacher speaking/										
Pupil engaged	254	64	518	66	509	73	216	76	241	86
Pupil off-task	137	35	267	34	188	27	69	24	40	14
Pupil engaged when teacher speaking to:										
Individual	174	67	306	61	289	69	128	73	123	79
Small group	26	55	78	71	75	82	33	72	44	92
Large group	54	65	133	77	145	78	55	87	74	95
Not speaking	166	62	332	63	211	63	115	76	136	89
Pupil engaged when teacher:										
Listening	94	75	145	68	81	60	35	74	52	87
Pupil engaged when:										
No interactions	72	51	188	60	130	66	79	76	82	91
Pupil engaged when teacher asks:										
Low-level questions	51	62	114	68	141	75	60	77	80	92
High-level questions	7	88	15	75	17	85	15	100	15	94
Questions about directions	8	67	41	59	46	70	34	68	39	78
Pupil engaged when teacher provides:										
Low-information feedback	43	62	68	62	50	70	20	71	20	83
Positive feedback	32	65	46	66	37	77	10	83	15	83
Information about content	53	74	106	68	147	82	57	83	64	90
Explains directions	87	61	159	67	101	63	25	69	22	79

Third, in the same manner, pupil engagement increases from 51% for Group 1 children to 91% for Group 5-6 children, when there were no teacher interactions. Finally, the same pattern of increase in engagement is apparent when teachers question students or provide information. However, as in the previous analyses, these are the same differences found across grade levels.

Cognitive level within class. The engagement data for children of different cognitive levels within each class was also calculated. Although there is some variation in engagement in each class for children of differing cognitive levels, no pattern of differences in any class was apparent. Thus, tables summarizing these data are not presented.

In summary, the data relating pupil engagement to type of teacher behavior suggest that differences are due to grade level and teacher style and not to differences in cognitive capacity among the students within each class.

Summary and Conclusions

Data from the sample of students in the five classes observed in this study indicate that children who differed in cognitive capacity did not receive different instruction. There were some overall differences in how the five teachers dealt with Group 1 and Group 5-6 children, but these differences are slight and are confounded with both grade and teacher effects. Nevertheless, the study provides some interesting insights about mathematics instruction. First, teachers tended to organize and teach mathematics based on school traditions. Differences in content emphasis and patterns of grouping students were based on school curricula. In particular, the differences in pupil actions and teacher actions from grades 1 and 2 to grade 3 reflected a shift in emphasis and organization of activities. Sandy Bay Infant School (grades 1 and 2) has an open, activity oriented program. Waimea Heights, on the other hand, is a school in which instruction is more formal and direct. Hence, the overwhelming grade level effect on pupil actions, teacher actions, and pupil engagement was to be expected.

Second, the mathematics program within schools was not related either to how students work problems or their capacity to reason. Third, there were important differences between one teacher and two others who used the same curriculum. Classes 3 and 5 in grade 3 clearly reflected good teaching that was following a direct instruction approach. Children were on task in large or small groups. Class 4, on the other hand, while following the same program, was not a successful class.

Fourth, the only interesting pupil behavior we found that was related to cognitive capacity was the tendency for children in higher groups to interact with other pupils more often when there was an opportunity to interact. This effect, however, may also be a function of grade, school, and teacher variables.

Chapter 6

SUMMARY, CONCLUSIONS, AND IMPLICATIONS

The question investigated in these set of five studies was, Do children who differ in cognitive capacity learn to add and subtract differently? In asking this question, we assumed that children's performance on addition and subtraction problems was related both to their cognitive capacity and to classroom instruction. This series of studies was reported from four different intellectual perspectives so that each study would shed light on a different aspect of the question. Then, by putting the information from each together, we hoped to answer the basic question.

In retrospect, we believe that the picture that has evolved from these studies is both interesting and provocative, but not at all clear. This chapter summarizes what we learned and specifies the strengths and weaknesses of each of the studies. Finally, implications are suggested to other researchers, to curriculum developers, and to teachers. We have organized this discussion under five headings: cognitive capacity, solving verbal addition and subtraction problems, using the concepts and skills of addition and subtraction, the influence of instruction on addition and subtraction performance, and final reflections.

Cognitive Capacity

The original question assumed that young children differ in their cognitive capacity to deal with mathematical information and that available psychometric techniques would yield groups of students with similar test scores. First, a set of tests measuring short-term memory capacity (M-space) was administered. Second, a set of developmental psychological tests was given to the same children. Data from the M-space tests were used to empirically derive six groups of students. The developmental tests were then used to assist in describing differences between the groups.

Cognitive Group 1 children have limited memory capacity (M-space level 1), are incapable of handling most quantitative tests, can serially count but have no sophisticated counting strategies, and can only deal with qualitative comparisons and transformations at a moderate level.

Cognitive Group 2 children have larger memory capacities (M-space level 2), have no difficulty with qualitative comparisons (they can preserve correspondence after rearrangement of sets and overcome perceptual distractions), and can determine whether sets are larger or smaller if an object has been put with or taken from particular sets. However, the quantitative skills of these children are limited. They can count sets, but have no sophisticated counting strategies and are unable to solve transitivity and rearrangement problems.

These first two groups are distinct from each other and distinct from the remaining four groups. The final four groups, both psychometrically and logically, are more similar to each other than they are different from each other in that all members have a memory capacity of level 3 or 4 and have sophisticated counting strategies.

Cognitive Group 3 children differ from Cognitive Group 4 children only on measures of transitivity and transitivity under rearrangement. Groups 3 and 4 children differ from Cognitive Group 5 and Cognitive Group 6 children only on the class inclusion test. Groups 5 and 6 children differ only on one measure of memory capacity; in all analyses we combined Groups 5 and 6.

The data gathered and analyzed with respect to cognitive capacity suggested the following six propositions. First, a global qualitative versus quantitative distinction is apparent in children's mathematical thinking in the early school years. Second, M-space level seems to be related to the developmental sequences in the preschool to early elementary years in mathematically-related tests. Third, reasoning appears to develop in the following sequence: comparison--qualitative--correspondence--quantitative--logical operations. Fourth, an M-space level of 1 enables a child to solve simple comparison tasks. Fifth, an M-space level of 2 is enough for qualitative correspondence and is a prerequisite for the development of most number skills. And sixth, an M-space level of 3 seems to be necessary for success in sophisticated counting tasks and probably is necessary for the development of addition and subtraction.

Problems and recommendations. The data indicate that children differ significantly in their ability to perform simple mathematical tasks. However, the approach that we took is purely empirical. It is not based on any theory of how children process mathematical information. The next step in research would be to use a theoretic model of cognitive processing such as that proposed by Campione and Brown (1978), which distinguishes between the "architectural" features of cognition (memory capacity, automaticity, speed of processing, etc.) and "executive" aspects of cognitive processing (metacognition, schema in long-term memory, etc.). Using such a model would enhance our understanding of cognitive capacity more than the psychometric approach followed in this study.

We found three sets of tests of cognitive processing to be especially important. First, memory capacity was most useful in identifying groups with different cognitive capacities. Unfortunately, the instruments used to assess this underlying trait leave much to be desired. In particular, on the Mr. Cucui Test, children can organize information by "chunking" it (e.g., left side of the body, head, and so on). As a result, higher M-space levels are indicated when children use a smaller part of memory to store information. This phenomenon is well known in the literature, but it is difficult to separate chunking from actual M-space. We believe that the four tests (Counting Span, Mr. Cucui, Digit Placement, and Backward Digit Span) indicate M-space levels 1, 2, and 3 relatively accurately. However, memory capacity levels above level 3 in many cases may be due to chunking. Nevertheless, we are convinced that memory capacity is an important feature of cognitive processing capacity and strongly suggest that other researchers measure memory capacity of their subjects.

The second set of tests that distinguished groups were the counting forward and counting back tests. Sophisticated counting skills are important in solving verbal addition and subtraction problems, as demonstrated in Chapter 3. We recommend that such tests be used in other research. Also, other tests that measure different counting skills (simple counting, counting on, counting back, counting all, etc.) should also be developed and used.

Finally, the class inclusion test distinguished groups of students. The relationship of class inclusion skills to how children work certain problems (particularly the part-part-whole problems) is not at all clear. We recommend the development and study of other tests that assess the way in which individuals logically reason about phenomena.

Solving Verbal Addition and Subtraction Problems

One indication that students have learned to add and subtract is that they can solve simple verbal problems. For such problems, children can write an addition or subtraction sentence about the problem and use learned addition or subtraction concepts or skills to find the appropriate answer. In Chapter 3, we examined both the performance of students (the number of questions they were able to answer correctly) and the strategies they used to solve a variety of addition and subtraction problems. The data were gathered in interviews of each child on several occasions in which six problems were given to each student at two or three of four levels of difficulty, determined by the size of numbers in the problem.

The results described in Chapter 3 indicate that there was considerable variability in the children's ability to solve a variety of verbal problems and in the strategies they used to solve those problems. The overall performance of students with different cognitive processing capacities on the tasks was relatively high. Students answered 72% of the SN level problems correctly; of the LN level problems, students answered 72% correctly; and of the NR and R level problems, 67% were answered correctly. However, there was considerable variability in both performance and strategies; this variation was influenced by several factors: the semantics of the problem, the size of the numbers in the problem, the implied operation in the problem, the grade level of the child, and the cognitive capacity of the child. Table 58 summarizes the level of performance across all items for the six different semantic types of tasks. In general, the results support the conclusion of Greeno and Riley (1981), that change problems are easier than combine problems, which in turn are easier than compare problems.

The most striking findings on both performance and strategies were for children in different cognitive groups. The performance and strategies that children in each cognitive group used are summarized on the following pages. The percent correct is noted only if children answered at least two-thirds of the tasks within a semantic category correctly. Similarly, to highlight the strategies that students in each particular group used, percentages are indicated only if students used a strategy on at least 20% on the same semantic set of problems.

The summary information on the performance and use of strategies for the children in cognitive Group 1 is presented in Table 59. This group of children performed satisfactorily only on three of the 12 tasks--the three SN tasks that can easily be solved by direct modeling. The strategies that these students used, with one exception, were either inappropriate or direct modeling. The exception was on task 6 at the SN level, when students used "counting on" 37% of the time.

Overall, this behavior clearly reflects the cognitive capacity of these children. They had low memory capacity, lacked systematic counting skills, and were only able to directly model the problems. Also, the compare task, which requires more memory capacity, was impossible for the children; they used inappropriate strategies on the SN and LN level compare tasks 83% and 93% of the time, respectively.

Table 58

Frequency and Percent Correct for Each Task on Different Items for All Students

Task	Level SN		Level LN		Level NR,R		All Levels	
	Frequency	%	Frequency	%	Frequency	%	Frequency	%
1 Change/join (+)	85	85	201	78	125	82	411	81
2 Change/separate (-)	79	79	184	71	89	58	352	69
3 Combine/part unknown (-)	68	68	190	74	88	58	346	68
4 Combine/whole unknown (+)	80	80	194	75	115	76	389	76
5 Compare (-)	41	41	157	58	90	59	288	56
6 Change/join, change set unknown (-)	77	77	191	74	105	68	373	73

Table 59

Performance and Common Use of Strategies for Cognitive Group 1

Task	Percent Correct	Direct Modeling	Counting Sequences	Routine Mental Operation	Inappropriate
SN Level					
1 Change/join (+)	77	53			
2 Change/separate (-)	70	57			33
3 Combine/part unknown (-)		40			50
4 Combine/whole unknown (+)	70	57			20
5 Compare (-)					83
6 Change/join, change set unknown (-)			37		37
LN Level					
1 Change/join (+)		37			63
2 Change/separate (-)		40			60
3 Combine/part unknown (-)		37			60
4 Combine/whole unknown (+)		37			63
5 Compare (-)					93
6 Change/join, change set unknown (-)					80

The summary information on the performance and use of strategies for students in cognitive Group 2 is shown in Table 60. This group of children could find answers satisfactorily on both the SN and LN sets of problems, with the exception of compare tasks. Although their performance was slightly lower on the LN than the SN set, the pattern of the performance was similar. However, on the NR and R sets (large numbers), group 2 students could only solve task 1 at a satisfactory level of performance.

The strategy information Group 2 students used was consistent with their cognitive capacity. Direct modeling was the most frequently used strategy for both the SN and LN level problems, although the use of routine mental operations was becoming commonplace with the small number of problems in the SN set. The students could not work most of the compare problems. Inappropriate strategies were coded in over half of the trials over all problems. Group 2 students only used systematic counting strategies on task 6 at both SN and LN levels and task 4 on the LN problems. Finally, for the problems with larger numbers (NR and R), they most frequently used inappropriate strategies on all tasks except task 1. Only on this task did these children use an algorithm, and they often made errors in the use of an algorithm.

The summary information for the children in cognitive Group 3 appears in Table 61. Their overall level of performance is quite satisfactory on all tasks at the SN and LN levels, although they had some difficulty with task 5. For the NR and R set, only on tasks 1, 4, and 6 was performance satisfactory. Direct modeling is a reasonable strategy, particularly on the small-number SN problems. Counting strategies, however, and routine mental operations were also being used

Table 60

Performance and Common Use of Strategies for Cognitive Group 2

Task	Percent Correct	Direct Modeling	Counting Sequences	Routine Mental Operation	Inappropriate	Algorithm
SN Level						
1 Change/join (+)	85	44		33		
2 Change/separate (-)	81	52		31		
3 Combine/part unknown (-)	77	50		27		
4 Combine/whole unknown (+)	83	48		27		
5 Compare (-)					58	
6 Change/join, change set unknown (-)	85		36	38		
LN Level						
1 Change/join (+)	72	42				
2 Change/separate (-)	68	43				
3 Combine/part unknown (-)	71	42				
4 Combine/whole unknown (+)	68	45	24			
5 Compare (-)					53	
6 Change/join, change set unknown (-)	70	28	26		26	
NR,R Level						
1 Change/join (+)	67				25	50
2 Change/separate (-)					33	33
3 Combine/part unknown (-)					41	
4 Combine/whole unknown (+)					42	33
5 Compare (-)					50	
6 Change/join, change set unknown (-)		21			29	

Table 61

Performance and Common Use of Strategies for Cognitive Group 3

Task	Percent Correct	Direct Modeling	Counting Sequences	Routine Mental Operation	Inappropriate	Algorithm
SN Level						
1 Change/join (+)	100	33		39		
2 Change/separate (-)	89	56		33		
3 Combine/part unknown (-)	89	50		28		
4 Combine/whole unknown (+)	89	56		33		
5 Compare (-)	67			44	33	
6 Change/join, change set unknown (-)	94		44	39		
LN Level						
1 Change/join (+)	72		35	29		
2 Change/separate (-)	68	23	26			
3 Combine/part unknown (-)	71	27	26	26		
4 Combine/whole unknown (+)	80		42	26		
5 Compare (-)	74		36	24	23	
6 Change/join, change set unknown (-)	82		32	35		
NR,R Level						
1 Change/join (+)	89			34		27
2 Change/separate (-)		34				29
3 Combine/part unknown (-)		29	23		25	
4 Combine/whole unknown (+)	70			27	23	29
5 Compare (-)			36		25	
6 Change/join, change set unknown (-)	66		36		25	

with small number problems. Sophisticated counting strategies were used on all LN level tasks and on three of the NR and R level tasks. Also, a fairly high frequency of inappropriate strategies was apparent on the NR and R tasks. The Group 3 students used algorithms for the NR and R problems less frequently than did the Group 2 children.

The summary information for Cognitive Group 4 children appears in Table 62. Not surprisingly, the performance and choice of strategies of these children differ very little from those of the Group 3 students. Group 4 students used counting strategies and routine mental operations on the LN problems and direct modeling and counting strategies on the NR and R problems. There was some increase in the use of algorithms by Group 4 students on the two addition problems.

The summary information of children in Cognitive Group 5-6 is shown in Table 63. These children solved all problems satisfactorily. They used counting strategies and routine mental operations to find most solutions. However, on the NR and R simple subtraction problems, tasks 2 and 3, they frequently employed direct modeling. They used routine mental operations and algorithms only with the three easiest tasks.

More important, these data show that a child's decision to use a particular strategy depends on several factors, including the semantics of the problem, the size of the numbers, and the implied operation. Furthermore, the availability or use of a strategy appears to depend on memory capacity.

Five general observations from the data relate to our understanding of how children learn to solve such problems. First, the frequent and persistent use of inappropriate strategies implies either an

Table 62

Performance and Common Use of Strategies for Cognitive Group 4

Task	Percent Correct	Direct Modeling	Counting Sequences	Routine Mental Operation	Inappropriate	Algorithm
LN Level						
1 Change/join (+)	91		34	41		
2 Change/separate (-)	77		39	27		
3 Combine/part unknown (-)	75		32	25		
4 Combine/whole unknown (+)	86		25	39		
5 Compare (-)	86	20	27	32		
6 Change/join, change set unknown (-)	84		23	45		
NR,R Level						
1 Change/join (+)	78			38		38
2 Change/separate (-)		31				26
3 Combine/part unknown (-)		31			31	
4 Combine/whole unknown (+)	74					45
5 Compare (-)		21	31		24	
6 Change/join, change set unknown (-)	67	29	31		21	

Table 63

Performance and Common Use of Strategies for Cognitive Group 5,6

Task	Percent Correct	Direct Modeling	Counting Sequences	Routine Mental Operation	Inappropriate	Algorithm
LN Level						
1 Change/join (+)	95		33	50		
2 Change/separate (-)	95		48	24		
3 Combine/part unknown (-)	98		31	48		
4 Combine/whole unknown (+)	95		38	48		
5 Compare (-)	93		48	36		
6 Change/join, change set unknown (-)	98		43	45		
NR,R Level						
1 Change/join (+)	88			32		39
2 Change/separate (-)	74	31				38
3 Combine/part unknown (-)	83	33	40			
4 Combine/whole unknown (+)	93			26		50
5 Compare (-)	71		60			
6 Change/join, change set unknown (-)	90		62			

unwillingness of some students to engage in the task, or an inadequate memory capacity to use a particular strategy. We agree with DeCorte and Verschaffel (1981) that some students fail to understand that they are to find an answer to a particular problem. However, we believe that most students try to solve problems but lose track of information. For example, Group 1 and Group 2 children do not have systematic counting strategies available to them to solve many of the problems. Thus, when they try, they may get mixed up and be unable to complete the task. We recommend a more careful investigation of the use of inappropriate strategies across tasks in order to obtain a better understanding of the difficulties some children have.

Second, direct modeling (the use of chips or fingers to represent sets) is the first and easiest strategy that students use. It is particularly appropriate for SN tasks 1, 2, and 4, where the change or combination can be physically represented. Also, direct modeling preserves all the important data; prior data need not be remembered. The strategy is appropriate for tasks 3 and 6, but additional memory storage is required to remember the whole and the original part. Finally, direct modeling could be used with comparison problems, but it requires even more memory storage. Even with large-number problems where physical modeling becomes more cumbersome, modeling is still an appropriate strategy. Many students appear to follow a "when in doubt one can always model" strategy for solving many problems. Even third-graders in Group 5-6 physically modeled many of the large-number problems to find answers. This suggests the importance of being familiar with efficient procedures; although children in Group 5-6 exhibited sophisticated counting strategies, knew basic facts, and could perform addition and subtraction algorithms efficiently, they still directly modeled some problems.

Third, the data also indicate that many children replaced direct modeling with either systematic counting strategies or routine mental operations. Counting strategies may be learned before routine mental operations; the choice of strategy depends on the size of the numbers involved in the problem. At all levels, for all cognitive groups, children solved the SN problems by using routine mental operations rather than counting strategies. Only for task 6 at the SN level was counting the dominant strategy. The LN, NR, and R problems were more likely to be solved by using sophisticated counting strategies. Furthermore, only on task 1 (combine/join) did children use routine mental operations with large-number problems.

Fourth, the use of addition and subtraction algorithms for many children was perceived as a cumbersome procedure for finding answers. Only the Group 2 children, who were limited in their knowledge of counting strategies or routine mental operations, used algorithms frequently, and they made many errors. Students at higher cognitive levels may see that algorithms are appropriate but know of and are comfortable in using other strategies. The children's teachers expected students to write the mathematical expression and use the algorithms to solve problems on the NR and R tasks. Most instruction had been on addition and subtraction algorithms, and the children's performance was reasonably good.

Fifth, it is apparent that the way in which students solved the problems was not directly related to classroom instruction. In grade 2, most instruction was on addition and subtraction facts (use of routine mental operations), but most students used direct modeling and counting skills to solve the problems. In grade 3, most of the instructional time was spent on algorithms, but students did not use them to solve the verbal problems.

Problems and recommendations. First, the sample of items--six tasks at each of the four levels--does not include all types of addition and subtraction problems. Nesher, Greeno, and Riley (1982) list 14 types. Use of a more comprehensive set of problems would give us a better picture of the overall development of strategies across tasks. Second, the small number of students and the method of selection for this study are limiting. Studies with a larger number of subjects are in order. Third, although some longitudinal data were gathered, there was relatively little change in performance over the three-month period. Although cross-sectional data indicate changes, studies of longer duration should be carried out. Fourth, there is an obvious confounding between age (grade level) and cognitive capacity.

Finally, these data need to be re-examined in light of the theory of the development of semantic categories for addition and subtraction proposed by Nesher, Greeno, and Riley (1982). Our data suggest that the decision sequence children use to select a strategy is more complex than this theory suggests.

Using the Concepts and Skills of Addition and Subtraction

Since most mathematics textbooks do not emphasize the solution of verbal problems, we also examined students' performance on the concepts and skills of addition and subtraction. This study is reported in chapter 4. A set of achievement monitoring tests that measured a variety of mathematics objectives was administered at each grade level. Instruction at each grade level had an effect on some objectives over the autumn term. In grade 1, at the start of the school year, students were unable to solve most problems; by the end of the term, their addition skills had improved dramatically, although the same could not be said for subtraction. In grade 2, although instruction had an effect, increases in performance were not as apparent for many objectives. Grade 3 students clearly improved on many of the objectives. In particular, performance on the addition and subtraction algorithms improved dramatically. Thus, growth within each grade on some aspects of addition and subtraction was very clear. However, improvement was not uniform across different concepts and skills, and the overall level of performance on many objectives was not high. Instruction did not seem to be related very systematically to the level of performance. Thus, in spite of the fact that overall performance on place value, knowledge of addition and subtraction facts, and writing number sentences was not high, time was not allocated for instruction on those topics. For example, in third grade, most students were still having trouble with writing open sentences and knowledge of basic addition and subtraction facts. Yet almost no time was allocated for instruction in these areas.

Performance differed by cognitive group within grade, although not all groups were represented at all grades. Group 1 children in grade 1 struggled with many of the objectives, while the Group 2 students improved in performance over all of the objectives. The children in the higher cognitive groups performed better than children in lower cognitive groups. Overall, children who differ in cognitive processing capacity performed differently regardless of specific objectives, instructional emphasis, or grade.

Thus, although teaching and pupil experience accounted for some of the differences between children, performance appears to be consistent with cognitive processing capacity.

Influence of Instruction on Addition and Subtraction

The final study in the series reported in this monograph attempted to determine whether or not children with different cognitive capacities received different instruction.

Observational data were collected on allocated time, pupil engagement, and teacher actions in relationship to pupil behavior. The findings of this study are not dramatic. However, what is portrayed is perhaps too typical of how instruction is carried out in many schools. First, about 50% of the total mathematics time in each grade is spent on addition and subtraction. In grade 1, primary emphasis is on addition facts, numerousness, and counting. In grade 2, basic facts for both addition and subtraction are taught. And in grade 3, computational algorithms are stressed. What pupils did in these classrooms seemed to be related to grade level and curriculum structure. In grades 1 and 2, children were working in small groups and individually for mathematics instruction while large group work was common in grade 3. Differences in pupil engaged time are likely due to teachers or student familiarity with the instructional pattern. Only the number of pupil interactions with other pupils is possibly due to the cognitive groups to which children belong. Teacher behaviors reflect grade level and individual teaching style. Certainly, cognitive capacity is not the basis of differentiation between students in these classrooms.

The data in the last two studies clearly indicate that children improve due to instruction on basic facts and algorithmic performance. What teachers do in classrooms varies, but within classrooms, they teach basically the same way to all children. What children learn appears to be consistent with their level of cognitive processing and with the content covered in each grade. The emphasis within classrooms seems to be on certain routine procedures (basic facts and algorithms) but not on others such as sentence writing, counting, or direct modeling of problems. The emphasis is on finding answers regardless of the procedure. Nothing is done to relate the semantics of various verbal problems to instruction in arithmetic.

Finally, there is no evidence that instruction attempts to build on or change the strategies that students use to solve verbal problems. In fact, instruction seems to proceed without consideration of the level of performance of individual children.

Final Reflections

In concluding this monograph, seven thoughts come to mind.

1. The information-processing approach to the study of how children solve a variety of addition and subtraction problems appears to provide a basis for a better understanding of the process of acquiring related concepts and skills and using them to solve problems. Our clustering of children into cognitive groups should be viewed as a rough initial approximation of a more refined description of capacity.

2. For students struggling with basic ideas (students in our Groups 1 and 2), a more careful analysis of inappropriate strategies needs to be done.

3. The most interesting data are those on the strategies that children use, not on performance. Longitudinal data on change of strategies by specific children should be gathered.

4. To be more effective, curricula must be organized and sequenced differently. Although the ideal organization and sequence for teaching addition and subtraction skills is not yet clear, more instruction on writing sentences and counting strategies is called for. One possible alternative would be to teach specific routines such as addition and subtraction facts or algorithmic procedures without trying to relate them to problems until students have mastered them. Students could build the bridge from verbal problems to use of algorithms later.

5. Students need more opportunities to work with verbal problems and to represent such problems with mathematical expressions. This procedure of modeling a problem situation with a mathematical sentence is a very important skill throughout all mathematics.

6. Although we believe that routine procedures are important, they only become important in the eyes of children when they see them as efficient and feel confident in using them to solve problems.

7. Children differ in their capacity to solve a variety of mathematical problems. Instruction should begin where children are. Teachers should take into account the strategies and procedures children use to solve problems and build upon those capacities.

In conclusion, our intent was to incorporate data from different perspectives to study how children learn to add and subtract. The picture that emerges is of children struggling to learn a variety of important concepts and skills. Some children are limited by their capacity to handle information. Most are able to solve a variety of problems by using invented strategies that have not been taught. They dismiss or fail to see the value of the taught procedures for solving problems. The capacity of children for processing information, the procedures students invent to solve a variety of problems, and the way in which instruction in schools is carried out are not consonant. The challenge in the future is to change this fact. Our goal is to make instruction compatible with children's capacities and the strategies they use.

References

Bachelder, B. L., & Denny, M. R. (1977). A theory of intelligence: I. Span and the complexity of stimulus control. _Intelligence_, _1_, 237-256.

Baldwin, M. M. (1895). _The development of the child and of the race_. New York: Macmillan.

Bauersfeld, H. (1979). Research related to the mathematical learning process. _New trends in mathematics_, Vol IV, (pp. 199-213). UNESCO.

Berliner, D. (1975, March). _Impediments to the study of teacher effectiveness_. Paper presented at the meeting of the National Association for Research in Science Teaching, Los Angeles.

Biggs, J. B., & Collis, K. F. (1982). _Evaluating the quality of learning: The SOLO taxonomy_. New York: Academic Press.

Bloom, B. S. (1974). Time and learning. _American Psychologist_, _29_, 682-688.

Bruner, J. S. (1966). _Toward a theory of instruction_. New York: Norton.

Buchanan, A. E., & Romberg, T. A. (1983). _Instrumentation for the achievement monitoring component of coordinated study no. 1_. Madison: Wisconsin Center for Education Research.

Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research on teaching. In N. L. Gage (Ed.), _Handbook of research on teaching_ (pp. 171-246). Chicago: Rand McNally.

Campione, J. C., & Brown, A. L. (1978). Toward a theory of intelligence: Contributions from research with retarded children. Intelligence, 2, 279-304.

Carpenter, T. P., & Moser, J. M. (1979). An investigation of the learning of addition and subtraction (Theoretical Paper No. 79). Madison: Wisconsin Research and Development Center for Individualized Schooling. (ERIC Document Reproduction Service No. ED 188 892)

Carpenter, T. P., & Moser, J. M. (1982). The development of addition and subtraction problem-solving skills. In T. P. Carpenter, J. M. Moser, & T. A. Romberg (Eds.), Addition and subtraction: A cognitive perspective (pp. 9-24). Hillsdale, NJ: Erlbaum.

Carpenter, T. P., & Moser, J. M. (1983). The acquisition of addition and subtraction concepts. In R. Lesh & M. Landau (Eds.), The acquisition of mathematical concepts and processes (pp. 7-44). New York: Academic Press.

Carpenter, T. P., Moser, J. M., & Romberg, T. A. (Eds.). (1982). Addition and subtraction: A cognitive perspective. Hillsdale, NJ: Erlbaum.

Carroll, J. (1963). A model for school learning. Teachers College Record, 64, 723-733.

Carroll, J. (1973). Fitting a model of school learning to aptitude and achievement data over grade levels (Research Bulletin No. 51). Princeton, NJ: Educational Testing Service.

Case, R. (1972). Learning and development: A neo-Piagetian interpretation. Human Development, 15, 339-358.

Case, R. (1975). Gearing demands of instruction to the developmental capacities of the learner. Review of Educational Research, 45, 59-97.

Case, R. (1978a). A developmentally based theory and technology for the design of effective instruction. Review of Educational Research, 48, 439-463.

Case, R. (1978b). Implications of developmental psychology for the design of effective instruction. In A. M. Lesgold, J. W. Pellegrino, S. D. Fokkema, & R. Glasser (Eds.), Cognitive psychology and instruction (pp. 441-463). New York: Plenum.

Case, R. (1985). Intellectual development, birth to adulthood. New York: Academic Press.

Case, R., & Globerson, T. (1974). Field independence and central computing space. Child Development, 45, 772-778.

Case, R., & Kurland, D. (1978). Development of the counting span test. Unpublished manuscript, Ontario Institute for Studies in Education.

Case, R., Kurland, M., Daneman, M., & Emmanuel, P. (1979). Operational efficiency and the growth of M-space. Paper presented at the biennial meeting of the Society for Research in Child Development, San Francisco.

Collis, K. F. (1971). A study of concrete and formal reasoning in school mathematics. Australian Journal of Psychology, 23, 289-296.

Collis, K. F. (1973). A study of children's ability to work with elementary mathematical systems. Australian Journal of Psychology, 23, 121-130.

Collis, K. F. (1974a). <u>Cognitive development and mathematics learning</u>. London: Chelsa College, University of London.

Collis, K. F. (1974b). The development of a preference for logical consistency in school mathematics. <u>Child Development</u>, <u>45</u>, 978-983.

Collis, K. F. (1975). <u>A study of concrete and formal operations in school mathematics: A Piagetian viewpoint</u>. Melbourne: Australian Council for Educational Research.

Collis, K. F. (1976). Mathematical thinking in children. In V. P. Varma & P. Williams (Eds.), <u>Piaget, psychology, and education</u> (pp. 144-154). London: Hodder and Stoughton.

Collis, K. F. (1978). Operational thinking in elementary mathematics. In J. Keats, K. F. Collis, & G. S. Halford (Eds.), <u>Cognitive development</u> (pp. 221-248). Chichester, England: Wiley.

Collis, K. F. (1980a). School mathematics and stages of development. In R. Modgil & S. Modgil (Eds.), <u>Toward a theory of psychological development</u> (pp. 635-671). Slough, England: National Foundation for Educational Research.

Collis, K. F. (1980b). Levels of cognitive functioning and selected curriculum areas. In J. R. Kirby and J. B. Biggs (Eds.), <u>Cognition, development, and instruction</u> (pp. 65-89). New York: Academic press.

Collis, K. F. (1982). The structure of learned outcomes: A refocusing for mathematics learning. In T. P. Carpenter, J. M. Moser, & T. A. Romberg (Eds.), <u>Addition and subtraction: A cognitive perspective</u> (pp. 171-182). Hillsdale, NJ: Erlbaum.

Collis, K. F., & Biggs, J. B. (1979). Classroom examples of cognitive development phenomena: The SOLO taxonomy (Technical Report prepared for ERDC). Hobart: University of Tasmania.

Cookson, C., & Moser, J. M. (1980). Coordinated study individual interview procedures (Working Paper No. 290). Madison: Wisconsin Research and Development Center for Individualized Schooling.

DeAvila, E., & Havassy, B. (1974). Intelligence of Mexican American children. Austin, TX: Dissemination Center for Bilingual and Bicultural Education.

DeCorte, E., & Verschaffel, L. (1981). Children's solution processes in elementary arithmetic problems: Analysis and improvement. Journal of Educational Psychology, 73, 765-779.

DeCorte, E., Verschaffel, L., Janssens, V., & Joillet, L. (1984). Teaching word problems in the first grade: A confrontation of educational practice with results of recent research. Leuven, Belgium: University of Leuven.

Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. Psychological Review, 87, 477-531.

Flavell, J. H. (1971). What is memory development the development of? Human Development, 14, 272-278.

Fullerton, T. (1968). An approach to number readiness by scalogram analysis. Unpublished honours thesis, University of Newcastle.

Ginsburg, H. (1977). Children's arithmetic: The learning process. New York: Van Nostrand.

Greeno, J. G. (1980). Trends in the theory of knowledge for problem solving. In D. T. Tuma & F. Reif (Eds.), Problem solving and education: Issues in teaching and research (pp. 9-23). Hillsdale, NJ: Erlbaum.

Greeno, J. G., & Riley, M. S. (1981). Processes and deveopment of understanding. Pittsburgh: Learning Research and Development Center.

Guttman, L. (1954). A new approach to factor analysis: The radex. In P. Lazarsfeld (Ed.), Mathematical thinking in the social sciences (pp. 258-348). Glencoe, IL.: Free Press.

Halford, G. S. (1980). Toward a redefinition of cognitive developmental stages. In J. R. Kirby & J. B. Biggs (Eds.), Cognition, development and instruction (pp. 39-64). New York: Academic Press.

Harnischfeger, A., & Wiley, D. E. (1975). Teaching-learning processes in elementary schools: A synoptic view (Studies of Educative Process Report No. 9). Chicago: University of Chicago.

Hiebert, J. (1979). The effect of cognitive development on first-grade children's ability to learn linear measurement concepts (Technical Report No. 506). Madison: Wisconsin Research and Development Center for Individualized Schooling. (ERIC Document Reproduction Service No. ED 178 361)

Inhelder, B., & Piaget, J. (1958). The growth of logical thinking from childhood to adolescence. London: Routledge and Kegan Paul.

Klahr, D. (1984). Transition processes in quantitative development. In R. Sternberg (Ed.), Mechanisms of cognitive development (pp. 101-140). San Francisco: Freeman.

Klausmeier, H. J. (1977). Origin and overview of IGE. In H. J. Klausmeier, R. Rossmiller, & M. Saily (Eds.), Individually guided elementary education: Concepts and practices (pp. 1-24). New York: Academic Press.

Kouba, V. L., & Moser, J. M. (1979). Development and validation of curriculum units related to initial sentence writing (Technical Report No. 522). Madison: Wisconsin Research and Development Center for Individualized Schooling. (ERIC Document Reproduction Service No. ED 183 382)

Kouba, V. L., & Moser, J. M. (1980). Development and validation of curriculum units related to learning of addition and subtraction algorithms (Working Paper No. 287). Madison: Wisconsin Research and Development Center for Individualized Schooling.

Moser, J. M. (1979). Young children's representation of addition and subtraction problems (Theoretical Paper No. 74). Madison: Wisconsin Research and Development Center for Individualized Schooling. (ERIC Document Reproduction Service No. ED 178 329)

Moser, J. M. (1980). What is the evidence that children invent problem-solving strategies. In R. Karplus (Ed.), Proceedings of the fourth International Conference for the Psychology of Mathematics Education (pp. 312-321). Berkeley, CA: University of California.

Nesher, P., Greeno, J. G., & Riley, M. S. (1982). The development of semantic categories for addition and subtraction. Educational Studies in Mathematics, 13, 373-394.

Pascual-Leone, J. (1970). A mathematical model for the transition rule in Piaget's developmental stages. Acta Psychologica, 32, 301-345.

Pascual-Leone, J. (1976). A view of cognition from a formalist's perspective. In K. F. Riegel & J. Meacham (Eds.), The developing individual in a changing world (Vol. 1, pp. 89-100). The Hague: Mouton.

Pascual-Leone, J. (1984). Attentional, dialectic, and mental effort. In M. L. Commons, F. A. Richards, & C. Armon (Eds.), Beyond formal reasoning (pp. 32-110). New York: Plenum.

Piaget, J. (1974). Foreword. In B. Inhelder, H. Sinclair, & M. Bovet (Eds.), Learning and the development of cognition (pp. ix-xiv). London: Routledge & Kegan Paul.

Resnick, L. B. (1978, September). The role of invention in the development of mathematical competence. Paper presented at the workshop on Children's Mathematical Learning. Learning Research and Development Center, University of Pittsburgh, Pittsburgh, PA.

Romberg, T. A. (1977). Developing mathematical processes: The elementary mathematics program for Individually Guided Education. In H. J. Klausmeier, R. Rossmiller, & M. Saily (Eds.), Individualally guided elementary education: Concepts and practices (pp. 77-109). New York: Academic Press.

Romberg, T. A. (1982). An emerging paradigm for research on addition and subtraction skills. In T. P. Carpenter, J. M. Moser, & T. A. Romberg (Eds.), Addition and subtraction: A cognitive perspective (pp. 1-7). Hillsdale, NJ: Erlbaum.

Romberg, T. A., & Braswell, J. (1973). Achievement monitoring via item sampling: A practical data-gathering procedure for formative evaluation. Journal for Research in Mathematics Education, 4, 262-270.

Romberg, T. A., & Carpenter, T. P., & Moser, J. M. (1978). Studies in mathematics (Technical Proposal, 1978-1979). Madison: Wisconsin Research and Development Center for Individualized Schooling.

Romberg, T. A., & Collis, K. F. (1980a). The assessment of children's M-space (Technical Report No. 540). Madison: Wisconsin Research and Development Center for Individualized Schooling. (ERIC Document Reproduction Service No. ED 195 331)

Romberg, T. A., & Collis, K. F. (1980b). The assessment of children's cognitive processing capabilities (Technical Report No. 539). Madison: Wisconsin Research and Development Center for Individualized Schooling. (ERIC Document Reproduction Service No. ED 195 330)

Romberg, T. A., Collis, K. F., & Buchanan, A. E. (1981). Performance on addition and subtraction problems: Results from individual interviews--Sandy Bay Study (Technical Report No. 580). Madison: Wisconsin Research and Development Center for Individualized Schooling. (ERIC Document Reproduction Service No. ED 215 892)

Romberg, T. A., Collis, K. F., Buchanan, A. E., & Romberg, M. N. (1982). Classroom learning: Results from observations--Sandy Bay Study (Working Paper No. 326). Madison: Wisconsin Center for Education Research.

Romberg, T. A., Harvey, J., Moser, J. M., & Montgomery, M. (1974, 1975, 1976). Developing mathematical processes. Chicago: Rand McNally.

Romberg, T. A., Small, M., & Carnahan, R. (1979). Research on teaching from a curricular perspective (Theoretical Paper No. 81). Madison: Wisconsin Research and Development Center for Individualized Schooling. (ERIC Document Reproduction Service No. ED 193 232)

Romberg, T. A., Small, M., Carnahan, R., & Cookson, C. (1979). Observer's manual, coordinated study #1, 1978-1980. Madison: Wisconsin Research and Development Center for Individualized Schooling.

Siegler, R. S. (1981). Developmental sequences within and between concepts. Monographs of the Society for Research in Child Development, 81.

Sternberg, R. J. (1984). Mechanisms of cognitive development. San Francisco: Freeman.

Vergnaud, G. (1982). A classification of cognitive tasks and operations of thought involved in addition and subtraction problems. In T. P. Carpenter, J. M. Moser, & T. A. Romberg (Eds.), Addition and subtraction: A cognitive perspective (pp. 39-66). Hillsdale, NJ: Erlbaum.

Wohlwill, J. (1960). A study of the development of the number concept by scalogram analysis. Journal of Genetic Psychology, 97, 345-377.