IDENTIFICATION OF NUMERICAL PRINCIPLES PREREQUISITE TO A FUNCTIONAL UNDERSTANDING OF PLACE VALUE

by

Drusilla Frey Gotow

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Educational Research

APPROVED:

Robert M. Todd, Co-Chairman

Lawrence H. Cross, Co-Chairman

Lawrence McCluskey

Cherry K. Houck

Harold W. Mick

June 1985

Blacksburg, Virginia

IDENTIFICATION OF NUMERICAL PRINCIPLES PREREQUISITE
TO A FUNCTIONAL UNDERSTANDING OF PLACE VALUE

by

Drusilla Frey Gotow

Committee Co-Chairman: Lawrence H. Cross
Educational Research
Committee Co-Chairman: Robert M. Todd

Mathematics Education

(ABSTRACT)

The purpose of this study was to find some remedy to frustrations engendered when children fail to grasp the essential principle of place value after several attempts at reteaching. It was hypothesized that these children must have failed to acquire understanding of some numerical principle(s) prerequisite to understanding the place value aspect of the numeration system. Four plausible prerequisite principles were identified: (1) synthesis of ordinal and cardinal properties of the numeration system, (2) both the addition and subtraction operations, (3) understanding of counting by groups, and (4) understanding of exchange equivalences such as one ten for ten ones, etc. It was hypothesized that understanding of analog clock reading was also dependent upon understanding of the same four prerequisite principles.

By conducting four pilot studies, six interview protocol instruments were developed to measure levels of understanding for the four prerequisite principles and the place value and clock reading criterion principles. Three levels of understanding: no understanding, transitional understanding, and competence were designated to correspond with Piagetian stages in the development of a new operation. Forty-eight children, twenty with second grade completed and twenty-eight with third grade completed, were tested on all six instruments.

Hypotheses tested were: (1) if the four identified prerequisite principles are necessary to understanding of place value, then subjects will demonstrate a level of understanding on the place value measure no higher than their lowest level of understanding achieved on the four prerequisite measures; and (2) if the four identified prerequisite principles are necessary to understanding of clock reading, then subjects will demonstrate a level of understanding on the clock reading measure no higher than their lowest level of understanding achieved on the four prerequisite measures.

The findings were that both hypotheses were supported at the .01 probability level. Analysis of the research design and examiner observations suggested possible explanations for anomalous aspects of the obtained data. Limitations, directions for further research, and implications for teachers were also discussed.

ACKNOWLEDGEMENTS

Special thanks and appreciation are offered to the following:

,
for my introduction to cognitive developmental research on the concept of number and for urging me to enter the doctoral program in Educational Research;

Robert M. Todd,
for extensive assistance at all stages of this project, including numerous suggestions for improving many aspects of the design and procedures and for devoting a large portion of the Advanced Curriculum and Instruction in Mathematics course time, in the Summer of 1982, to the examiner training and data taking interviews that made completion of this project possible;

Lawrence H. Cross,
for the many hours devoted to assistance on research design, statistical analysis, editing and procedural details to complete this project;

Cherry K. Houck and Harold W. Mick,
for their many hours and professional expertise devoted to reading and commenting on this manuscript;

and Lawrence McClusky,
for their time and expertise providing research design and statistical analysis advice on this project;

, , , ,
, , , , and
,
for all the time, care, and effort devoted to training for and performing as examiners for this study;

, , , and ,
for their generous contributions of time and professional knowledge to evaluating the face validity of the instruments;

Montgomery County and Giles County School personnel,
who provided their cooperation and assistance to my conducting the pilot studies in their schools; and

,
for many years of patience, support, encouragement and professional advice.

TABLE OF CONTENTS

Page

LIST OF TABLES vii
LIST OF FIGURES..viii

Chapter
1 THE RESEARCH QUESTION..1
THE PREREQUISITE PRINCIPLES..4
Prerequisite A: One Equal in Value to Many.........................5
Prerequisite B: Counting by Groups.................................6
Prerequisite C: N More than M is N up from M in Counting...7
Prerequisite D: The Addition and Subtraction Operations Represent the Joining and Separating of Disjoint Subsets Comprising the Whole Set..........................8
Prerequisites E, F, G, etc...9
THE HYPOTHESES...10
The Primary Hypothesis...12
A Secondary Hypothesis...13
Levels of Understanding as Measured by Performance.........15
2 A SURVEY OF PREVIOUS WORK ON THE PRINCIPLES BEING INVESTIGATED.....18
READINESS TO LEARN...18
PLACE VALUE PREREQUISITES..23
Hierarchies of Skills Preceding Place Value: Empirical Studies.......26
Advice on Skills and Concepts Preceding Place Value........29
MEASUREMENT FACTORS CONSIDERED.....................................33
Levels of Understanding..33
The Roles of Concrete Materials and Language in Learning...37
SUMMARY 41

3 DEVELOPMENT OF THE INSTRUMENTS...................................43
TESTING FACTORS INVESTIGATED.......................................44
OVERALL DESIGN OF THE INSTRUMENTS..................................46
The Addition-Subtraction Instrument................................47
The Counting By Groups Instrument..................................57
The One Equal in Value to Many Instrument..........................60
The N More and Less Instrument.....................................62
The Place Value Instrument...67
The Clock Reading Instrument.......................................73
Explanation Tasks for Each Instrument..............................76
DEVELOPMENT OF SCORING PROCEDURES..................................77
RELIABILITY OF THE INSTRUMENTS.....................................81
VALIDITY OF THE INSTRUMENTS..84
Teacher-Instrument Agreement.......................................85
Age and Grade Appropriateness......................................89
Face Validity Assessment...91

4 METHODOLOGY 96
SUBJECTS...98
TESTING PROCEDURES...100

PROTOCOLS..102
THE EXAMINERS..102
ANALYSIS OF THE DATA...103
LIMITATIONS OF THE STUDY.....................................104
Reliability..105
Validity...105
Research Design..106

5 FINDINGS AND DISCUSSION......................................108
DATA OBTAINED..108
TESTS OF HYPOTHESES..114
DISCUSSION OF RESULTS..117
SUMMARY 123

6 INTERPRETATION AND CONCLUSIONS...............................124
CONSIDERATIONS PERTAINING TO CLOCK READING...................126
SUMMARY OF CONCLUSIONS.......................................129

7 SUMMARY..131

LITERATURE CITED...135

APPENDIX A: THE INSTRUMENTS.....................................146

APPENDIX B: PILOT STUDY FORMS...................................224

VITA...229

LIST OF TABLES

Table..Page

1. Test-Retest Scores Obtained in the Spring 1980 Pilot Study, Accompanied by the Spearman's Rho Reliability Statistic and the Simple Proportion of Agreement..................................83

2. Contingency Tables Comparing Teacher Estimates and Tested Levels of Understanding for each of the Six Principles............87

3. Raw Scores and Associated Performance Level Classifications......109

4. Crosstabulation of Performance Levels Between the Prerequisite Measure and the Criterion Measure of Place Value.................111

5. Crosstabulation of Performance Levels Between the Prerequisite Measure and the Criterion Measure of Clock Reading...............112

6. Bionomial Test of the Difference between Expected and Observed Proportions of Subjects Whose Performance Level on the Criterion Exceeds Their Performance Level on Each of the Prerequisite Measures..................................116

7. Number of Prerequisite Principle Scores Equal to the Lowest Score Measured on all Four Prerequisites (N = 48)................119

LIST OF FIGURES

1. One Set of Picture Part and Whole Equations........................54

2. Examples of Picture Part and Whole Equations.....................152

3. A-S Comp. 3 Set of Equations.....................................158

4. A-S Comp. 4 Set of Equations.....................................159

5. A-S Comp. 5 Set of Equations.....................................160

6. Grp-Cnt Comp. 1 Task Card..171

7. Grp-Cnt Comp. 2 Task Card..172

8. Grp-Cnt Comp. 3 Task Card..173

9. Grp-Cnt Comp. 5 Task Card..174

10. Fifteen of the Dot Strips arranged in Order.....................196

11. Ordered Dot Strips as They Appear after Being Covered............197

12. Magnetic Number Board with Initial Placement of Numbers..........212

13. Card One for the Trans. 5 Task..................................213

14. Card Two for the Comp. 1 Trial..................................214

15. Card Three for the Comp. 2 Trial................................215

16. Problems for the Trans. 5, Comp. 3, Comp. 4, and Comp. 5 Tasks...216

17. Profile Sheet for Reporting Subject Responses....................225

18. Evaluation Form for Teacher Estimate of Subject's Levels of Understanding of the Six Principles Investigated.................227

CHAPTER 1

THE RESEARCH QUESTION

After four years of observing student teachers (and their pupils) in elementary classrooms and helping them to cope with the myriads of problems encountered when teaching elementary mathematics, this researcher has become more and more convinced that the lack of a functional understanding of the place value aspect of our numeration system is at the root of a large proportion of math learning problems encountered in Grades 3 through 7. Underhill (1972:128-130), Payne and Rathmell (1975:126-160), and Ginsburg (1977:150-168) have all stressed the importance of understanding the place value aspect of numeration for successful learning in elementary mathematics.

The term "place value" refers to an organizational aspect of our number system which is essential to assigning specific quantitative values to each digit in a numeral. The integers, 0,1,2,...,9 symbolize cardinal values representative of the empty set, a set of one item, a set of two items,...; and a set of nine items, respectively. However in the numeral 20, the digit 2 does <u>not</u> represent the quantity in a set of two items. In this case, the two represents the quantity in a set of two <u>groups</u> of items with each group containing ten items. Furthermore, in the numeral 200, the digit 2 represents the cardinal value of a set of 2 <u>groups</u> of items, where each group contains one hundred items. Thus in our numeration system, the quantity represented by each digit in a numeral is the cardinal value, symbolized by the digit, multipled by the position value, symbolized by the position that digit holds in the numeral. That is, in the numeral XYZ, Z symbolizes a quantitative value of Z times one, while Y symbolizes a quantitative value of Y times ten, and X symbolizes a quantitative value of X times one hundred. This position multiplier (i.e. one, ten, one hundred, etc.) for each digit in a numeral, then, is what is referred to by the term "place value."

Since place value plays an essential role in so many areas of the middle and upper elementary mathematics curriculum (for example, in the vertical algorithms for addition, subtraction, multiplication and division, in work with other bases and numeration systems, in all the operations with decimal fractions, the metric system and money, and in the use of calculators and digital clocks), failure to acquire a functional understanding of place value when it first appears in the math curriculum can be a crippling deficiency.

Knowing why some children grasp the place value principle so readily when it is initially presented, while other children continue to flounder after numerous "reteaching" efforts on the part of the teacher is of interest. Since learning mathematics is quite distinctly hierarchical in nature, in that concepts and principles must be learned in a logical order, each successively building on previously established principles, it seems quite plausible to suspect that children who continue to flounder after several "reteachings" of the place value principle of numeration are missing some essential concept or principle. For these children the place value aspect of our numeration system just does not exhibit any logical organization. In fact, Brownell (1972:72-80), Ginsburg (1977:121-129, 150-168) and Kennedy (1975:11-12) all stress the hierarchical nature of mathematics learning and the wisdom of looking for "gaps" in the child's understanding when persistent errors are encountered.

Therefore, it is the purpose of this research effort to identify the most likely prerequisite principles which may explain the place value difficulties many children encounter. If such prerequisite principles could be identified, the path for moving these children towards a successful encounter with the place value principle could be defined.

After identifying plausible prerequisite principles, the empirical test for determining whether each is a probable prerequisite to a functional understanding of place value would be possible. For a particular principle (A) to be identified as a prerequisite to place value understanding, it is necessary to classify each child's performance as falling into one of three patterns of association with regard to his or her performance on explaining and applying the place value principle. These three performance patterns are:

1) failure on Foundation A tasks is associated with failure on place value tasks,
2) success on Foundation A tasks is associated with success on place value tasks, or
3) success on Foundation A tasks is associated with failure on place value tasks.

The pattern of association which would <u>not</u> support Foundation A as a prerequisite to place value understanding would be:

4) failure on Foundation A tasks associated with success on place value tasks.

Furthermore, if a complete set of N prerequisites (i.e. Foundations A, B,...,N) has been identified, then

1) failure on place value tasks is associated with failure on at least one of the Foundation A, B,...,N tasks, and
2) success on place value tasks is associated with success on all of the identified Foundation A, B,...,N tasks.

Diagnostic snapshots recording elementary children's explanations of and attempts to apply the principles in question, including the place value principle, provide the empirical data for the statistical decision making process of identifying place value prerequisites.

THE PREREQUISITE PRINCIPLES

The initial task was to identify what appear to be the essential prerequisite principles to be tested. The vast majority of research studies on the place value topic seem aimed at identifying and analyzing the resultant error patterns and other symptoms associated with place value problems, rather than looking for specific causes of the deficiency. For example, Ginsburg (1977:79-149) cited numerous examples of clinical interview data setting forth children's systematic errors and hypothesized that incomplete understandings of essential principles (such as the place value aspect of numeration) was the cause. However, he does not attempt to probe further the reasons for the incomplete understandings. Weaver (1971, 1972, 1973) and Grouws (1972, 1974) completed several research studies for the purpose of identifying factors in addition, subtraction, multiplication and division open sentences which affect pupil performance, but have only speculated as to why one type of problem poses more difficulty than another. Cox (1975) conducted a two year study to identify the most frequently occurring systematic errors in the vertical algorithms for the four arithmetic operations and tested 744 children in grades 2 through 6. Again, while many consistent but faulty error patterns were catalogued, the underlying causes of the faulty procedures were not investigated. Ashlock (1972), Ellis (1972), and Roberts (1968) have all published similar studies identifying persistent erroneous computation strategies.

A review of mathematics educators' suggestions of numerical principles essential to the understanding of place value proved more fruitful.

Prerequisite A: One Equal in Value to Many

A substantial number of recognized authorities (Payne and Rathmell (1975:137-141), Reisman (1977:104-117), Moser (1971), Schminke, Maertens and Arnold (1973:114)) have suggested that the principle of one of something being equivalent in value to many of something else (such as a dime equals ten pennies) is a crucial element of place value "readiness." This idea represents the very essence of the place value construct. One might argue whether it should be viewed as a "readiness" requirement or as the central goal for understanding the place value aspect of our numeration system. However, numeration is a symbolic system for organizing and recording a particular property—number. As Frege (1884:67-81) explained, this property number is quite an abstraction that can represent things and non-things. For example, one can count objects, but one can also count spaces between objects or intervals of time or a list of ideas. Furthermore, the number property can be quite unstable. It can change by simply redefining the spatial constraints of the collection in question. For example, the number of spoons on the table and the number of spoons in the room could easily be different, the first being a subset of the second, but both are "the number of spoons." Number can only be defined for a specified space and time.

In addition to this instability and abstractness of the number property itself, the principle of "one equal in value to many" involves the notion that a numeral can represent a value other than the cardinal value the child has previously been trained to associate with that symbol through concrete experiences with sets of objects. Therefore, accepting Piaget's assertion (Ginsberg and Opper, 1969:169) that children learn by generalizing from concrete experiences, it seems reasonable to hypothesize that children would need active concrete experiences with the notion of "one of something carrying a value equivalent to many of something else" before attempting the more abstract concept that the <u>position</u> of a numeral in a symbolized quantity determines the value of that numeral in combination with its previously ascribed cardinal value.

<u>Prerequisite B: Counting by Groupings</u>

A closely related numerical principle that may precede the development of "one equal in value to many" principle, is the idea of counting a set of items by grouping them in subsets of equal size and then counting by groups rather than counting each separate item. This process tends to increase the speed with which one can obtain a cardinal value for a set and also tends to reduce the chances for committing counting errors. Counting by twos or by fives would be examples of applying this principle. This is hypothesized to developmentally precede the "one equal in value to many" principle since the latter involves counting groupings abstractly by assigning the value of the whole grouping to one substitute item, while "counting by groups" is a procedure for counting concrete items. Thus the place value aspect of numeration can be construed as a four step abstraction from 1) counting equal groupings, to 2) counting individual items carrying the value of a grouping, to 3) applying the groupings value idea to subsequent groupings of groupings, etc., and to 4) symbolizing the values of the different levels of groupings by the position of the digit representing the number of groupings to synthesize a complete system of numeration. Many authorities (Underhill (1972:128-131), Dienes (1969:83-87), Payne and Rathmell (1975:137-141) and Moser (1971)) have cited the principle of counting equal sized groups rather than individual items as one of the most elemental for understanding the place value principle.

Prerequisite C: N More than M is N Up From M in Counting

Another basic principle of numeration was first brought to twentieth century educators' attention by Piaget (1952:96-157). Piaget asserted that the synthesis of the ordinal and cardinal aspects of number is essential to a functional understanding of the abstract construct, number, prior to superimposing the principle of place value to elaborate the whole system of numeration. This principle is more simply stated as: the quantity which is N more than a given number of items is expressed by the N^{th} number above the given number in the ordered sequence of counting numbers; and conversely, N less than a given quantity is expressed by the N^{th} number down in the ordered sequence from that quantity's number. Thus 8 which is the third number after five in the counting order represents an amount which is quantitatively 3 more than 5. To understand place value functionally, this principle must be generalized from one more than, two more than, etc. to one group of ten more than, two groups of ten more than, etc. Thus, N groups of ten more than any number will be represented by the number whose tens place numeral is the N^{th} one above the tens numeral in the given quantity with the one's place numeral remaining the same. For example 20 more than 34 is 54 since 5 is the second number after 3 in the counting sequence and 20 is 2 tens. Since this principle must ultimately be generalized to the relationship between the cardinal value of two and three place numbers and their ordinal position in the counting sequence, it seems likely that understanding of the less complex version of the ordination-cardination synthesis principle, applied to the counting of individual items rather than groupings, would be a likely prerequisite to a functional understanding of the place value aspect of numeration.

Prerequisite D: The Addition and Subtraction Operations Represent the Joining and Separating of Disjoint Subsets Comprising the Whole Set

A fourth principle which Payne and Rathmell (1975:136-137), Schminke, Maertens and Arnold (1973:113-114) and Kennedy (1975:98) have stressed as necessary before commencing the study of place value is a proper perception of the operations of addition and subtraction. That is, the addition process must be perceived as the joining of disjoint subsets (rather than as a process of counting up all the given items which ignores the subset components of the process) since quantities greater than nine in our numeration system are symbolized as separate tallies of specific component subsets. That is, 123 is the sum of one group of one hundred items, two groups of ten items and three individual items. Seeing a sum as comprised of two or more subset components rather than as just the totality of many individual items is quite indispensable to perceiving the component structure of two and three digit numbers, which is in essence perceiving the place value aspect of numeration.

Likewise the perception of the subtraction process as the separation of a whole set into disjoint subsets (rather than as a counting backwards exercise or a removal of items followed by a counting tally of the remaining items) is equally essential since many functions involving quantities greater than nine require separating the quantity into its place value components. In fact, a functional understanding of place value can be equated with an ability to perceive a quantity greater than ten as the combined tally of its component power-of-ten groupings. Thus, the perception of both the addition and subtraction operations as mental actions needed in dealing with a whole set as the combination of its component subsets is a likely prerequisite to a functional understanding of place value.

Prerequisites E, F, G, etc.

While numerous other numerical concepts and principles can be logically supported as prerequisite to an understanding of place value, others considered by researchers (i.e. relating cardinal number of a set of objects, ordering sets according to quantity and conserving number in spite of perceptual transformations) can all be demonstrated logically to precede some or all of the four prerequisite principles already proposed.

THE HYPOTHESES

Two hypotheses were tested in this research, a primary hypothesis relating to the Place Value criterion measure and a secondary hypothesis relating to a Clock Reading criterion measure. Both hypotheses define an expected relationship between levels of performance on a criterion measure and levels of performance on measures of the four prerequisite principles. The four prerequisite principles are defined below, accompanied by the title used for subsequent reference:

Prerequisite A (One Equal to Many):

> One of something can be considered equivalent in value to many of something else;

Prerequisite B (Counting By Groups):

> It is possible and often quicker to determine the cardinality of a set by counting equivalent groupings rather than counting each separate item;

Prerequisite C (Plus or Minus N):

> N more than M items is the n^{th} number above M in the counting sequence and conversely, N less than M is the n^{th} number below M in the counting sequence; and

Prerequisite D (Addition-Subtraction):

The addition and subtraction operations represent the joining and separating of disjoint subsets comprising the whole set. Three ordinally related levels of performance (N: No understanding, T: Transitional understanding, and C: Competence) were distinguished on measures of understanding of each of the prerequisite and criterion principles included in this study.

For each of the prerequisite principles, the expected relationship with the Place Value principle is that the level of performance on the Place Value measure will not exceed the level of performance on the prerequisite measure.

This expected relationship can be illustrated by considering the crosstabulation of score levels across prerequisite measures and the Place Value measure as shown below:

		Place Value Performance Level		
		N	T	C
Prerequisite Performance Level	N	a	c	c
	T	b	a	c
	C	b	b	a

In this diagram, the diagonal entries (a) represent the numbers of subjects scoring at the same performance level on both the prerequisite and Place Value measures. The entries below the diagonal (b) represent numbers of subjects whose score level on the prerequisite is lower than their performance level on the Place Value measure. The entries above the diagonal (c) represent the numbers of subjects whose performance level on the Place Value measure exceed their performance level on the prerequisite measure.

To the extent that the data support the hypothesized relationship, the performance of all subjects should be represented by the a and b entries in the above diagram with no subjects represented by the c entries. This is because performance on the Place Value criterion measure is permitted to be equal or lower than performance on the prerequisite but cannot exceed the performance on a prerequisite measure if the understanding tested by the prerequisite measure truly is prerequisite to understanding of Place Value.

The Primary Research Hypothesis

The primary research hyothesis is that the performance level on the criterion measure of Place Value will not exceed the lowest performance level recorded on any of the four prerequisite measures. The rationale for considering the lowest performance level across all four prerequisite measures is that if all four are equally necessary to a functional understanding of place value, then the level of performance on the Place Value measure should not exceed any of the performance level scores on the prerequisite measures. Since the lowest score recorded will be the limiting score level which is not expected to be exceeded, selecting this score level for testing the hypothesis is the statistical procedure required.

The Secondary Hypothesis

Since the clock represents another type of numerical system presented in our mathematics curriculum which would logically require the understanding of the same four foundation principles as prerequisites to its understanding, it would seem fruitful to include the clock system in this research. Although it is usually approached as a specific measuring device rather than as a modular numerical system, it includes the relevant components.

The one equal to many principle is involved in the aspect of 60 minutes being equivalent to one hour, 60 seconds being equivalent to one minute, and two groups of 12 hours being equivalent to one day. Thus while our decimal system has one scale of equivalences throughout (that is, ten in any digit position is equivalent to one in the digit position to its immediate left), the clock system is a multi-scale system having different equivalences for minutes, hours, and days. The number of different scales increases considerably if one considers the calendar part of our time measurement, but in this study, only the clock system will be studied.

Continuing the contrast with our number system, the clock system "value indicator" is not the position of a numeral in an ordered sequence. Instead the "value indicator" is the length of the "hand" pointing at the same set of numbers. Thus, the same numbers, 1 through 12, carry different values depending on the length of the hand doing the pointing. The clock system has an added complexity over our numeration system in that the same numeral can take on two different values simultaneously, if both hands of the clock happen to point to the same position.

The counting by groups principle is needed to count the minutes since each number on the clock is counting groups of five minutes as well as single hours. Moreover, the synthesis of ordination and cardination principle must be utilized whenever the elapsed time before or after a given time is to be determined; while the addition and subtraction principles are used whenever time is represented precisely in hours and minutes, or even more so, in hours, minutes and seconds.

Although the theoretical justification could not be found in educators' curriculum recommendations, an analysis of elementary textbook series found that most math curriculum designers schedule teaching of clock reading (to exact minute accuracy) during the same school year that place value is most commonly introduced. In an attempt to obtain empirical confirmation for this relative scheduling of the teaching of clock reading and to identify the "readiness" requirements for a successful encounter with time measurement, the following secondary hypothesis will also be tested in this research endeavor. Children will not achieve a higher level of performance on a measure of their clock reading understanding than their lowest level of performance on measures of understanding of the four hypothesized prerequisite principles enumerated for the primary hypothesis.

For most educators, clock reading is considered a measurement skill to be learned rather than the application of a mathematical principle. However, the emphasis in this study is on the fact that the analog clock is a representation of a modular numeration system. ("Modular" refers to the recycling of the minute count from one to sixty, then back to one again. Similarly, the hour count recycles back to one after reaching twelve.) Therefore, reference is made to the Clock Reading principle during this discussion of this research study, rather than to clock reading skills. Understanding of this Clock Reading principle is assumed to encompass the synthesis and application of all the "rules" about lengths and positions of hands, direction of movement of hands, and interpretation of minute and hour values for the numerals that appear on the clock face, as well as the understanding that the minute count returns to one after reaching sixty at the end of each elapsed hour.

Levels of Understanding as Measured by Performance

Since these research hypotheses make very specific predictions about relative levels of understanding of certain principles as measured by tasks designed for the purpose, a clear definition of the term "levels of understanding" is essential. While many research studies base comparative levels of learning attainment on the number of questions answered correctly, this appears to be both an arbitrary and presumptive method for determining levels of learning attainment. Arbitrary, because it assumes that each correct response represents an equal increment of attainment; and presumptive, because it assumes that each and every question will elicit a response indicative of the subject's true attainment level.

According to Piagetian learning theory as described by Ginsburg and Opper (1969:161-180), the acquisition of any generalization is characterized by three stages or levels of attainment. A person either (1) lacks knowledge of the generalization completely, (2) has become initially aware of the generalization, but lacks some substantial knowledge about the extent of its applicability, or (3) has a broad understanding of the generalization and its range of applicability (although, perhaps, still limited in minor ways by a finite variety of experiences with its use) and can explain the essential aspects of the general rule clearly, although not necessarily eloquently. Therefore, it would be useful to construct instruments to measure the extent of understanding of these six principles in such a way as to determine which of these three levels of understanding has been achieved. Thus, for all six principles described in the two research hypotheses (the four hypothesized prerequisite principles and the two predicted criterion principles), the instruments of measurement were designed to determine which of these three levels of understanding has been attained by each subject. Then, these measured levels of understanding, (1) no understanding, (2) transitional recognition, and (3) competence, were the data base for the statistical tests of the two hypotheses previously stated.

To state the research predictions more explicitly, a child who demonstrated the competence level (3) of understanding on the place value tasks (and similarly, for the clock reading tasks) would be expected to demonstrate the competence level (3) of understanding on all of the four prerequisite tasks. However, a child who was measured to be at the transitional recognition level (2) of understanding on the place value tasks could be expected to demonstrate either the transitional recognition level (2) or the competence level (3) of understanding on each of the prerequisite tasks. Extending this, a child who demonstrated no understanding of place value might display any of the three levels of understanding on each of the prerequisite tasks.

CHAPTER 2

A SURVEY OF PREVIOUS WORK ON THE PRINCIPLES BEING INVESTIGATED

This literature review includes discussions on a number of aspects of this research study. Concerning the research question, the topics of "readiness to learn" and "Place value prerequisites" need to be investigated, and specific "prerequisites" to understanding the place value principle need to be identified. Next, factors influencing how understanding of each principle is measured need to be investigated. This will encompass discussions of (1) Piaget's three stages in the acquisition of a new logical operation, and (2) the roles of language development and experiences with concrete materials in the learning process.

READINESS TO LEARN

The concept of readiness for the learning process is not new. Paschal (1968) and Callahan and Glennon (1975:58) discussed three types of readiness for mathematics learning. Paschal calls them "process readiness," "affective readiness," and "product readiness." His "process readiness" refers to the level of cognitive functioning of the learner, while "affective readiness" refers to interests, appreciations, attitudes, and values of the learner. "Product readiness" refers to subject-matter readiness in terms of needed concepts, skills, and understandings for a specific learning task.

Cognitive readiness is most extensively elaborated by Jean Piaget (Richmond, 1970:99-108) in his theory of cognitive development. Richmond and Suydam and Weaver (1975:46-48) point out that Piaget identified four important factors interacting to bring about intellectual change. These include (1) neural maturation or physiological development of the brain, (2) interaction or encounters with the concrete environment, (3) social interaction with peers and adults, and (4) equilibration or the synthesis of the previous two interactions with the mental structures and language of the learner to produce adjustments or changes in language and or mental structures. Dienes (1969:xvii-11) also emphasized the process of abstracting from concrete experiences to form concepts such as number, equal, more, less, etc. and then further generalizing from abstractions to formulate hypotheses about the real world.

As Suydam and Weaver pointed out, numerous research studies have confirmed Piaget's age-range predictions for children's transition from the preoperational stage to the concrete operational stage of cognitive development during the primary grades of elementary school. Two operations which Piaget (1952) identified as markers of the onset of concrete operational thinking are conservation of number and synthesis of ordination and cardination. Of these two, conservation of number has been the most extensively researched in relation to mathematical readiness in the primary grades. Almy, Chittenden, and Miller (1966), Dadwell (1961), Robinson (1968), and Wheatley (1968) have all found a significant positive relationship between a young child's ability to conserve number and his level of achievement on beginning mathematics tasks in grades K or 1. Two other studies found that success on conservation of numerousness was an excellent predictor of success on addition problems (Steffe, 1967) and on subtraction problems (LeBlanc, 1968) for children entering first grade. In contrast, Hiebert et al. (1982) found that success on Piagetian tasks was not necessary for success in solving verbal arithmetic problems although appropriate solution strategies were used more often by children succeeding with Piagetion tasks. Kingrea and Koops (1983) found that Piagetian tasks were about as effective as intelligence tests in predicting success on a variety of primary school arithmetic tasks. Thus it would appear that "cognitive readiness" as defined by Piaget's concrete operational stage of development is to some extent a predictor of the degree of success children have with beginning arithmetic, but not necessary for some success to be achieved.

On the other hand, Almy and associates (1970) and Simpson (1971) conducted conservation studies on early primary children in which they found that learning style (reflective or impulsive), attitudes and interests of the children as well as the type of materials used by the teacher in the classroom appeared to affect children's cognitive development. These findings would seem to support Piaget's assertion that the four interacting factors of maturation, experience with concrete materials, social interactions, and equilibration all have a bearing on cognitive development. They also seem to support Callahan and Glennon's and Paschal's contentions that "readiness" for mathematical learning is a composite factor involving "affective readiness" as well as "cognitive readiness" and "subject-matter readiness."

On the issue of subject-matter readiness, Brownell (1970) and Gagné (1975:101-106) both stress the necessity of assuring that children have acquired prerequisite skills and knowledge when commencing with a new topic in the mathematics curriculum. In a research study on introducing division with two-place divisors, Brownell found that nearly half the students deemed "ready" by their teachers did not have mastery of the requisite skills and facts basic to success in this learning endeavor. He stressed the need for improved diagnostic procedures to determine readiness for a specific task before introduction.

In a study of learning success on several math topics where programmed learning materials were used presenting a hierarchy of tasks sequentially, Gagné (1970) found that children effectively stopped achieving subsequent objectives in the hierarchy once they were unable to master a subordinate knowledge task. Thus this study supports the need for subject-matter readiness for successful learning and Gagné's assertion that task analysis and careful sequencing of hierarchically related objectives is a very important component of the instructional process.

Another study of a task analyzed and hierarchically sequenced set of objectives in an introductory math curriculum was conducted by Resnick, Wang and Kaplan (1973). This curriculum sequence developed number concepts through numeration up to 20 as well as addition and subtraction equations with sums up to 10. While a very detailed curriculum sequence was developed, the data presented on its use in a kindergarten for two successive years was anecdotal in nature. Profiles for individuals were presented indicating the number of objectives mastered each month. While there were almost no violations of the orders permitted within the hierarchies, rates of mastery were somewhat uneven for most individuals. Thus the hierarchical sequences seemed to be validated with some reservations about possible additional steps needed where mastery rates were consistently slow. Since alternative mastery sequences were not attempted, possible violations of the order would not necessarily have been detected during the program. Nonetheless, the results of the study did lend additional support to the importance of subject-matter readiness for successful math learning.

Several other representative studies of "readiness" factors generally supported the hypothesis that higher scores on "readiness" measures correlate with greater success on mastery of academic objectives. Chissom, Collins and Thomas (1974) reported correlations between perceptual-motor measures and academic readiness in preschool children. Murray (1973) reported the effect of teaching "readiness" activities on the success levels with first-grade math topics. Hamrick (1976) found that facility with the verbal terminology of the addition and subtraction processes preceded success with the symbolization of addition and subtraction. Hirstein (1978) conducted a longitudinal study of children's acquisition of number concepts through the first and second grade. He found that children who could conserve number at the beginning of the study had a tremendous advantage in the acquisition of arithmetic concepts throughout first and second grade. He found also that rational counting strategies of conservers could form the basis for children's acquisition of whole number addition and subtraction concepts, but not their acquisition of base ten numeration concepts. The variety of "readiness" factors represented in these studies discussed collectively supports the assertions by Paschal, Callahan and Glennon that academic readiness is a multi-dimensional variable affecting outcomes of the learning process.

PLACE VALUE PREREQUISITES

In reviewing the literature for studies specific to place value "readiness" or identifying hierarchies leading up to learning about 2- and 3-digit numeration or addition and subtraction with regrouping, two findings dominated the results of the search. Firstly, place value, both as an aspect of numeration and as an important property affecting understanding of the vertical algorithms for addition and subtraction, is a topic of concern to many educators. This follows from the vast number of articles on these subjects and the variety of approaches to these topics. Underhill (1972:128) states it very succinctly: "Place value is a rogue! Many children in grades one, two, three and four do not understand place value concepts even after repeated exposure. Many teachers do not fully understand place value either." The second finding, however, was that most articles could be categorized as (1) advice on remediating difficulties with the learning of 2- and 3-digit numeration or the aforementioned algorithms, (2) research studies cataloguing the patterns of errors children make on these topics and thus, the relative difficulty of different categories of problems on these topics, or (3) research studies comparing different teaching procedures for introducing the topics to children in the classroom. Very few studies were directed at determining prerequisite understandings necessary to success in learning about the place value aspect of our numeration system.

While an enumeration of the many studies identifying problems with the understanding of place value would be unproductive, one study

(Flournoy, Brandt and McGregor, 1963) may be representative. This study tested pupils' understandings of 15 basic concepts related to understanding decimal numeration. Pupils were tested in Grades 4, 5, 6 and 7. Even in the seventh grade, more than 50% of the subjects missed ten out of 25 test items. The item, "2. Which means 25 hundreds and 4 tens?" was missed by 75% of seventh graders, even though the correct answer, 2540, was given as one of four choices from which to select and all questions were read aloud by examiners to assist pupils with weaker reading skills! The additive principle whereby 444 means 400 + 40 + 4 was identified as one of the less understood concepts of numeration at all four grade levels tested. The equivalence of 1000, 100 tens and 10 hundreds was another poorly understood numeration concept. While these results are somewhat dated, more recent studies of performance on numeration tasks by Bednarz and Janvier (1982) and Grossman (1983), demonstrate that more current pupils in the same grades do not do much better. These studies all describe children's weaknesses with understandings of numeration system concepts, and suggest that problems in the lower grades tend to persist even into the middle school grades. Thus the apparent extensive concern about children's difficulties with understanding the place value aspect of our numeration system would seem to be well-founded.

Several studies were reviewed which investigated alternative methods of introducing two-digit numeration to first or second graders. Studies by Baker (1977), De Flandre (1974), Edge (1980), and Ziegenbalg (1981) found introductions using manipulative materials such as Dienes' base-ten blocks, chip trading activities, and bundled coffee sticks were more successful than more abstract presentations using an abacus or paper-and-pencil activities. One similar study by Duncan (1980) found no difference between concrete and expanded notation introductions, however his treatment period was only one week. Another study by Warner (1978) found that teaching place value understanding to second graders with base three, five, and ten blocks was no more successful than using base ten blocks alone when twelve 30-minute lessons were employed. In general it would appear that the use of concrete representations of the numeration system in introducing the principle of place value has some merit over more abstract introductions, and that a variety of materials can be effective.

One other study (Barr, 1976) comparing methods of introducing two-digit numeration to kindergarten children was of special interest. Three methods were compared, where all three involved counting concrete objects in sets greater than ten. One group was taught to count objects one-by-one; a second group was taught to group the objects by groups of ten and then assign the numeral giving the number of groups of two followed by the leftover ones; and the third group was taught to count the objects one-by-one and then check their count by using the grouping by tens method. While immediate post-test results did not show significant differences among groups, retention tests four weeks later confirmed trends evident in the post-test results, and indicated that there was a significant difference among groups. The third method was superior to the first, with the second method rating last. Prior abilities in writing two-digit numerals, conserving number, and proficiency with counting by tens were found to be good predictors of success on the numeration application measures used.

Hierarchies of Skills Preceding Place Value: Empirical Studies

One other group of studies was of considerable interest to this investigation. These studies all used Gagné's (1965) method of task analysis to develop a hierarchy of skills leading up to an understanding of place value in two- or two- and three-digit numerals. Two of the studies also investigated the application of place value understanding to the vertical addition and subtraction algorithms.

Matlow (1976) explored the effects of token reinforcement and a task analyzed program on the acquisition of the place value concept for learning handicapped children. The task analyzed program was compared with a regular textbook program using the appropriate sections of the Stanford Diagnostic Arithmetic Test and an experimenter-designed test of the place value concept. Significant differences were found between groups in favor of the task analyzed program, but no significant effect was found for the token reinforcement groups or interactions between programs and token reinforcement. Thus, this study demonstrated that a task analyzed program was more effective for teaching the place value concept to these children with low mathematical abilities and minimal understanding of the place value concept.

A second study by Smith (1973) first identified skills prerequisite to place value mastery using task analysis, and then determined which of these prerequisite skills had not been mastered by primary children using item error rates to compute difficulty indices. While Smith cited problems with the vertical addition and subtraction algorithms as motivation for the study, only the place value aspects of numeration, including hundreds, tens, and ones, were task analyzed and tested. It was found that naming the same number in several ways, such as "1 ten, 7 ones = ____ tens, 17 ones" was the most difficult type of item analyzed. Counting by tens in situations like "65, 75, ____" and "13, 23, ____" was found to be the next most difficult type of item. Other items asking for equivalent amounts in ones given tens or vice versa were also found to be especially difficult. Thus renaming amounts in other denominations as needed in regrouping in the vertical algorithms was found to be especially difficult. This finding confirmed the results of Flournoy, Brandt, and McGregor study cited earlier, although this time, with primary grade children. Furthermore, counting by tens was found to be a more difficult critical skill as was found in the Barr study cited earlier.

A third study by Daugherity (1978) used ordering theory to validate the ordering of early primary mathematics topics. Gagné's hierarchical prerequisite analysis was used by the curriculum designer to produce the initial hierarchy of prerequisites. The hierarchy encompassed beginning addition and subtraction through three-digit vertical addition and subtraction, and one-digit multiplication and division. The Primary II Metropolitan Achievement Test was given to 121 second graders to validate the ordering with each item then categorized to correspond to the appropriate topic in the initial heirarchy. Those orderings found to be supported at the .05 confidence level were: (1) 1-digit results precede 2-digit results, (2) 1-digit operands precede 2-digit operands, (3) 2 operands precede 3 operands, and (4) all 2-operand vertical additions precede all 3-operand vertical additions. Also noteworthy, addition preceding subtraction was supported at the .061 confidence level, and all operations without regrouping preceding all operations with regrouping was supported at the .081 confidence level. While there were some limitations to the study, especially the use of the Metropolitan Achievement Test which did not have a desirable number of items for each topic appearing in the initial hierarchy, the findings in support of some specific orderings of topics do suggest that the validation method does merit further investigation. Specifically, for this investigation, the identified ordering of specific topics could be relevant to item selection for the diagnostic instruments needed. In addition, simple addition and subtraction of one-digit operands is shown to precede the same operations with two- and three-digit operands with regrouping. Thus, this study suggests that simple addition and subtraction could be considered prerequisite to the application of place value understanding to the regrouping problem in the corresponding vertical algorithms.

Thompson (1982) conducted a study to create a theoretical framework for explaining the development of children's concepts of numeration. His framework was analyzed with respect to the products of eight case studies of first- and second-graders. He found that flexibility in numerational thinking did not develop until linguistic routines and meanings were established as mental operations. He also found that reading and writing numerals bears little relationship to understanding numeration. While his findings seemed to be derived somewhat subjectively, they suggest that understanding of numeration would be demonstrated by performing operations on numerals rather than by accurate counting, reading and writing of numerals.

Advice on Skills and Concepts Preceding Place Value

One other hierarchical study by Bidwell (1969) is of interest for its completeness, although this study must be categorized as "advice" since no empirical testing of the ordering was reported. The author endeavored to construct the ordering of Gagné-defined "subordinate concepts" in order to identify previous learning necessary to "readiness" for introduction of the vertical algorithm for subtraction. Bidwell developed three distinct hierarchical sets of subordinate concepts hypothesized to precede readiness for learning the vertical subtraction algorithm. These three hierarchical sets were (1) Whole-Number System concepts, (2) Whole-Number Addition concepts, and (3) Whole Number Subtraction concepts. It should be noted that a number of concepts were included in all three hierarchies such as: concepts of sets and set properties, equivalent sets by 1-1 correspondence, concept of cardinal number, and grouping by tens. The Whole Number System set also included: rote counting, recognizing and writing numerals, ordering the numerals from 0-99, and grouping by 100's, all subordinate to a general place value concept. The Whole Number Addition set included concept of addition, properties of addition, and addition equations to 18, all subordinate to the vertical addition algorithm. The Whole Number Subtraction set was analygous except for the inclusion of: addition and subtraction as inverses, and regrouping a 2-digit number. Bidwell suggested that teachers use his learning structures analysis and, teaching with concrete materials, assure that pupils master subordinate concepts before commencing to learn superordinate concepts.

For the purposes of this study the completeness of these hierarchies is especially helpful. Consolidating the many concepts into more inclusive groupings of mathematical principles then, one could summarize by concluding that: (1) both ordinal and cardinal properties of the numeration system, (2) both the addition and subtraction operations, including the recognition that they are inverses, (3) the understanding of counting including using groupings of ten, and (4) the understanding of regrouping or exchanging one ten for ten ones, etc., are principles subordinate or prerequisite to the learning of the place value principle in order to successfully apply it to the use of the vertical algorithms for addition and subtraction.

In conclusion then, the studies by Matlow (1976), Smith (1973), and Flournoy, Brandt and McGregor (1963) all support Bidwell's inclusion of ordinal and cardinal properties of the numeration system as prerequisite to successful understanding of place value, while Barr's finding that prior acquisition of conservation of number is a predictor of success in learning place value suggests that Piaget's concrete operational stage of cognitive development could also be considered prerequisite. Almy (1966), Howlett (1973), LeBlanc (1968), Overhalt (1965), and Steffe (1969) have all conducted studies supporting the contention that children who have demonstrated successful performance of various tasks associated with Piaget's concrete operational stage of development have greater success on measures of early primary mathematics achievement. In addition, Copeland (1974:97-110), Dienes (1966:3-5) and Lovell (1971:32-40) have stressed the importance of developing children's understandings of cardinal and ordinal number including conservation of numerousness and coordinating cardinal and ordinal aspects of numeration before commencing study of place value and symbolization of addition and subtraction operations. Therefore, selecting the principle of the synthesis of ordination and cardination as one of the four principles prerequisite to a functional understanding of place value encompasses subordinate understandings of ordinal and cardinal aspects of numeration as well as assuring that the child has achieved Piaget's concrete operational stage of development, as well.

In reference to the second consolidated grouping of mathematical concepts, Schminke, Maertens and Arnold (1973:113) stressed the importance of understanding addition and subtraction as joining and separating actions prior to the study of number greater than nine in reference to their place in the system of numeration. Similarly, Payne and Rathmell (1975:137) emphasized the importance of recognizing the subset or inclusion relation in reference to concepts of addition and subtraction, and Kennedy (1975:98) asserted that the ability to express a number as the sum of two addends (such as 11 = 10 + 1) is essential, before learning about place value in the system of numeration. Also, Underhill (1972:128-131) included experience with the use of the associative property of addition for regrouping quantities as an introductory activity for understanding place value in numeration, thus implying that the addition concept must already have been acquired. Thus, along with Flournoy and others' findings on the critical importance of the additive principle to the mastery of place value, these other authorities confirm the selection of the addition-subtraction principle as an essential prerequisite to a functional understanding of place value.

Looking at the third grouping of concepts, namely ability in counting, and especially counting by groups, many authorities stress the importance of giving children concrete experiences with grouping things in groups of ten, and some recommend using smaller groups of 2, 3, 4, etc., as well, however few authorities put much emphasis on extending counting skills specifically to counting-by-groups. May (1974:42-44) labeled this "sequence counting" and recommended including counting by twos, fives, and tens, using both concrete groupings and representational sequences such as number lines and arrays such as a 10 x 10 hundreds chart in the place value readiness curriculum. Current elementary school mathematics textbooks all include varying amounts of emphasis on "sequence counting" activities, supporting May's (1974) assertion of its importance along with Barr's (1976) findings cited earlier.

The fourth inclusive grouping of mathematical concepts consolidated from the hierarchies encompasses the One Equal to Many principle, the understanding of exchanging equivalent amounts such as one ten for ten ones. While Payne and Rathmell (1975:137-158) and Reisman (1977:91-117) emphasized the importance of activities involving groups of ten, Dienes (1969:80-111), Kennedy 1975:128-182) suggested that activities using concrete materials to work with numeration systems having smaller bases as well as base ten are important to acquiring a good conceptualization of the place value aspect of our numeration system. Kennedy (1975) and May (1974) both suggested that money exchange activities are especially effective in conveying the one equal many principle since it is a real-world model of the place number equivalences. Thus this principle appears to have almost unanimous backing, both from educators and from empirical studies, as an important subject-matter prerequisite to functional understanding of place value.

To summarize, extensive supporting evidence for the four hypothesized prerequisite principles has been found in empirical studies, hierarchical curriculum design studies, and mathematics educator's recommendations. While other mathematics concepts have also been suggested as additional subordinate concepts to place value understanding, all of them are encompassed in the four consolidated principles, or were shown to be subordinate to one of these four principles in the hierarchical studies cited.

MEASUREMENT FACTORS CONSIDERED

In order to make an empirical determinaton of whether and to what degree a child has acquired a functional understanding of place value, clock reading, and the four hypothesized prerequisite principles, some choices about measurement procedures were made. First of all, it was necessary to decide what sort of scoring procedure would be used to designate understanding or the lack thereof. A second related decision was needed concerning the type of test item and testing procedure to be used. For both of these decisions, the fact that measurement of conceptual understanding rather than skill acquisition was the objective was of considerable importance.

Levels of Understanding

Three scoring possibilities were considered: (1) an all or nothing score, (2) a percentage of items correct score, and (3) a Piagetian stage score designating no understanding, transitional understanding, or competence level understanding. As explained in Chapter 1, the three levels of understanding designation favored by Piaget (1952) was selected as being most appropriate for measuring understanding of basic mathematical principles. Copeland (1974:84-95), Ginsburg and Opper (1969:164-168), and Van Engen and Steffe (1970:101-102) have all discussed the rationale behind the designation of three levels or stages in the development of understanding of a basic mathematical generalization.

Essentially, a child starts by being completely ignorant of the principle and, when presented with a situation upon which to make a judgment, at this first stage, the child makes a decision based on the most salient perceptual attributes of the situation. For example, in judging which of two linearally arranged sets of objects has more members, a child at this first stage is apt to base the decision on the observed lengths of the two set arrangements and ignored other relevant attributes such as relative size of objects and density or spacing of objects. Similarly, in judging which of two glasses has more juice, the first stage responder is most likely to compare the heights of liquid in the two glasses while ignoring the relative diameters of the glasses. In another situation such as candy in bowls, the first stage child might find diameter the more salient dimension and judge that the wider bowl has more, without considering the height of the candy in the bowls. Thus the first stage of no understanding of the principle is characterized by attending to whatever feature is most salient in the situation and inability to coordinate dimensions from more than one attribute in order to make an accurate judgment.

At the second stage of partial understanding, the child has recognized that attending to one salient attribute can lead to erroneous judgements. However coordinating data on more than one variable to make an accurate judgment is still a less than secure ability. Feigenbaum (1963) in a study on number conservation including 90 children found that the number of objects in a collection could affect children's ability to ignore salient perceptual features. Dodwell (1960) obtained similar results on a study involving 250 children. It would seem that even when children have experienced an initial recognition of the targeted principle, extremely convincing perceptual features can still dissuade the child from sticking to the newly found principle. Thus while the second stage child compares sets of 5-10 objects correctly, considering all relevant features, that same child may revert to the immature perceptual judgment when larger numbers are involved. The second stage child vacillates between his newfound understanding and his former immature mode of making judgments based on salient perceptual attributes.

In the third stage of complete understanding, the child has a thorough and unshakeable grasp of the principle and applies it consistently in appropriate situations. The third stage child can explain what the relevant variables are in regard to applying the principle and can justify judgments involving the principle's application to a concrete situation.

Acceptance of this three stage scoring method for designating levels of understanding of basic mathematical principles is widespread. Virtually all research on Piagetian cognitive developmental tasks involving preschool to adult subjects includes the three levels of understanding in the analysis of the results. Since this research study is concerned with measuring functional understanding of the targeted principles, this three stage scoring method seems consistent with the way children are likely to acquire understanding of each principle.

On the role of explanation or justification for determining whether the third (or competence) stage of understanding has been achieved, Piaget (1952) was unequivocal. He believed that a child could not be judged to have left the second stage, characterized by vacillation in application of the principle, unless the child could give a clear justification for his correct judgments in applying the principle. Green (1966) discussed the role of explanations in conservation research in detail. Considerable controversy had surfaced as to whether explanations were necessary to judge a child a conserver. Researchers found that children were judged to be conservers at considerably younger ages if explanations were not required. However, a study by Biskin and others (1975) found that children who gave conserving responses but were not able to justify their decisions were more easily dissuaded from their conserving responses when presented with contradictory perceptual evidence. Children who were able to justify their conserving responses were not so easily dissuaded. This suggested that children who are unable to explain their judgments might be more appropriately classified as being stage two responders since vacillation can be very easily induced. Therefore Piaget's criteria for the three levels of understanding were used in this study and explanations for correct judgments were necessary to earn the third stage or competence level designation.

The Roles of Concrete Materials and Language in Learning

After reviewing studies related to the research objectives, it became evident that there were two distinctly different types of test items and procedures employed to obtain data on primary children's numerical understandings. One type included multiple-choice or fill-in-the-blank written items such as the Flournoy and others' (1963) item shown earlier: "2. Which means 25 hundreds and 4 tens?" with four selections for answers from which to choose. This type of diagnostic item was either read by the subject alone, or read in unison by each subject while an examiner read aloud to that group of subjects. For this type of testing procedure, the reliability and validity of the responses is dependent on the subject's reading skills and levels of attention and motivation. While reading skill can be measured and treated as a variable in the study, attention and motivation are less easily monitored in a procedure where groups of subjects are tested simulataneously. However, testing time and examiner requirements can be kept to a minimum using this type of testing procedure, which is a decided advantage in allocating research study resources.

The second type of test items and procedures employed in empirical studies reviewed was the individual interview procedure using some kind of written protocol to standardize the items administered. This type of procedure is necessarily much more time consuming at least for examiner time devoted to data gathering. In addition, there is a much greater potential for variability in the presentation of the task items to different subjects. The specificity of the protocols and the training of examiners become important factors in the overall research design. However, close monitoring of a subject's attention and motivation is possible with an opportunity available to adjust testing procedures immediately to remedy attention and motivation problems, within the constraints of the specified protocol. The opportunity for interchange between examiner and subject offers the opportunity to probe unexpected or incomplete responses further where protocol and research objectives dictate.

In general, the Gagné inspired hierarchical studies reviewed earlier employed the first type of written items administered in a group setting, while most Piagetian cognitive development studies employed the individual interview procedures. However, a study by Brownell (1967) including 90 Scottish and 928 English primary school subjects employed both types of testing procedures in different ways. Group administered tests were used to measure achievement and ability variables while individual interviews were conducted to measure conceptual maturity. Likewise, a number of Piagetian cognitive development studies use group administered tests to measure co-variables such as ability and achievement in specific curriculum areas while using individual interviews to measure the primary dependent variable.

A major factor influencing the selection of the individual interview style of testing procedure for this research endeavor was a consideration of the characteristics of primary-age children. A number of researchers of learning and cognitive development (Dienes (1969:1-11), Lovell (1971:1-22), Richmond (1970), and Siegel (1982:123-155)) have studied and discussed Piaget's theory about how children develop concepts and acquire a repertoire of numerical principles for use in the continuing process of living and learning. Concrete materials and manipulative actions on these materials are thought to be the initiator in this process leading to generalizations for further use. Piaget asserted that children abstract properties common to a class of objects being manipulated.

For example, children abstract the concept of number as an indicator of the numerousness of members of sets of objects. Number is not an attribute of any object within the set. It is a property only of a collection of objects and only as that collection is perceived by the child. Number is not a property intrinsic to a chair or a collection of chairs, but a property abstracted to represent the numerousness of the collection of chairs. Thus according to Piaget (1952) number is purely a logico-mathematical construction of the mind, abstracted from the child's perceptions of collections of objects manipulated by him. The numeral 2 is a symbol which stands for the concept of "two," the concept itself abstracted from manipulative experiences involving collections having two objects each. Likewise, the written word "two" and the spoken word "two" are symbolizations of this concept representing the numerousness of a pair of objects. Therefore, language is a form of symbolic representation of concepts abstracted from concrete experiences.

In the years of early childhood children are abstracting from interactions with their environment constantly to develop concepts at an astonishing rate. Language is concurrently developed through social interaction with other human beings to provide symbolic representations for these concepts being developed. At a later stage in early childhood, prelogical associations of concepts and actions represented through language gives way to the emergence of mental operations on classes and relations which are truly manipulations of the symbols, rather than symbolic representations of concrete actions performed by the child. At this point the child is capable of constructing "generalizations" symbolizing perceived relationships between concepts. Nevertheless, as indicated by Piaget's use of the term concrete operations to represent this stage in the child's logical development, these mental operations on symbols within the mind are still dependent upon the abstraction process from concrete experiences. Actions not specifically experienced can be imagined and operated on within the mind as long as they can be constructed from concepts and relations previously abstracted from concrete experience.

Researchers dealing with primary children just entering or recently having entered the stage of concrete operations in their logical thinking, must consider this factor of language as symbolization of concrete experiences in terms of the testing process. To use the written or spoken language, unelaborated by reference to concrete manipulation or pictorial materials, would seem to run the risk of making unsupported assumptions about each child's conceptual understandings symbolized by the testing language used. While any mode of communication employed in the testing situation must necessarily make the assumption that both parties are symbolizing very similar concepts when the same word is used, it would seem to be prudent to use whatever means possible to ensure as accurate communication as possible where conclusions are going to be drawn about the understandings of the subjects based on these communications. Copeland (1974), Dienes (1969), and Lovell (1971) as well as many other educators mentioned herein have stressed the role of manipulative materials in communicating and developing numerical concepts with children. Therefore, the individual interview procedure coupled with verbatim protocols employing the use of manipulative and pictorial representational materials to improve communication of task questions and responses was selected as most likely to ensure as accurate communication as possible between subject and examiner.

SUMMARY

Several conclusions can be drawn from the findings of this review of literature. Empirical studies involving "affective readiness," "cognitive readiness" and "subject-matter readiness" factors have been shown to be positively correlated with success levels on a variety of numerical learning measures. Furthermore, a number of empirical studies identifying "subject-matter readiness" factors for understanding various aspects of the place value principle have been reported. Consolidating the many concepts identified into more inclusive groupings, four numerical principles were suggested by these findings to be prerequisite to a functional understanding of place value. These consolidated groupings are (1) both ordinal and cardinal properties of the numeration system, (2) both the addition and subtraction operations, including the recognition that they are inverses, (3) the understanding of counting, including using groupings of ten to count, and (4) the understanding of regrouping or exchanging one ten for ten ones, etc.

The use of Piaget's three stages in the acquisition of new numerical generalizations for defining the scoring procedure is supported by the logical construction of the three stages and their relation to the research being conducted. Another aspect of Piaget's cognitive development theory is the identification of the roles of manipulating concrete materials and acquisition of language in the learning process. Considering these roles led to the use of manipulative and representational materials for the measurement process in this study.

Further consideration of findings from published research studies will be incorporated into the development of the instruments to be discussed in Chapter 3. In the process of developing the protocols for measuring levels of understanding of each of the six principles included in the research hypotheses, related research was considered in the selection of tasks to be used.

CHAPTER 3

DEVELOPMENT OF THE INSTRUMENTS

After determining the kind of data needed to test the hypotheses, the next step was to obtain or develop the necessary instruments. The decision to indicate the degree of understanding of each principle with a three level score system presented a problem for using published tests, although it might have been possible to adapt selected instruments to this scoring requirement. The need for instruments each designed to measure the understanding of one of the six specific math principles identified in Chapter 1 made obtaining appropriate instruments even more difficult. Furthermore, the variety of principles to be tested made it impossible to find six instruments of similar formats and similar numbers of test items which were desirable for appropriate statistical evaluation of the data obtained.

For these reasons, the development of new instruments designed purposefully to meet the requirements enumerated above was preferable. In total, four pilot studies were conducted during the development of these instruments. The first two pilot studies were conducted to evaluate the suitability of a variety of possible diagnostic tasks. Besides selecting task items appropriate to measuring understanding of the targeted math generalizations, it was important to find items which generated reliable responses. In addition, time required for administration as well as attractiveness and clarity of direction for the early primary subjects being tested were equally important factors.

The first study was conducted during the first few months of 1976 with sixty first and second graders in a rural county school system in southwest Virginia. The second study was conducted one year later with twenty second and third graders in one school of another southwest Virginia county school system.

TESTING FACTORS INVESTIGATED

In terms of administration, several considerations were important. It was necessary to limit the number of different styles of tasks using different equipment and manipulative materials for each of the six instruments in order to keep total administration time short. One and one-half hours was set as the upper limit since schools generally preferred that students not be kept out of the classroom for more than thirty to forty minutes during any one day. Having to do additional interview sessions with the same subject on separate days presented two problems. Absenteeism could result in fairly long periods between testing sessions and ongoing instruction in the classroom could produce an intervening history factor where teachers were giving instruction on one or more of the principles being tested. By contrast, longer sessions would have presented the difficulty of maintaining attention with young children. This made the use of more than one session per subject necessary unless all six principles could be tested in thirty to forty minutes.

It was found, not surprisingly, that the more colorful and toy-like the testing materials were, and the more the subjects were involved in manipulating the test materials, the easier it was to maintain concentrated attention on the tasks at hand. It was also noted that task materials not in immediate use needed to be stored out-of-sight of subjects as the attention-getting properties of some materials became a distraction during other tasks. Shelves below table level on the examiner's side of the presentation surface became a necessary part of the testing environment. A closed room where all distractions could be avoided was found to be desirable after low-traffic hallway areas were tried and found to introduce considerable inconsistency in responses.

In order to put children a little more at ease when being interviewed by a stranger, outside of their regular classroom, an introductory explanation was devised and perfected throughout the four pilot studies. The straightforward introduction was found to be unexpectedly successful in encouraging fairly uninhibited responses from the children. Subjects were informed that they were participating in research to help teachers learn more about how children think about numbers and math so that teachers can find more ways to help children when they are having problems with their math. It was emphasized that researchers wanted to see many children's ways of approaching these tasks whether or not they were successful since even "wrong" procedures would help researchers to better understand children's thinking. The children were told that examiners would be asking lots of "why" questions and also watching how they performed activities since how solutions were obtained was of special interest.

The explanation also stressed that to do good research and to be able to compare different subjects' responses, it is necessary to be sure all questions are exactly the same for all subjects and that all answers were remembered exactly as the subject gave them. Therefore examiners would be reading questions exactly as written and writing down each response given in a booklet to be sure everything was done accurately. It was also explained that each child would be assigned a number which would be recorded on the booklet rather than their name so that they would be accorded privacy from being talked about personally as researchers studied the answers later. This explanation seemed very effective in getting subjects to "help" researchers by showing off their skills and ideas for the research.

OVERALL DESIGN OF THE INSTRUMENTS

In developing the task items, questions were needed that focused as directly on the generalization as possible yet which were somewhat unconventional so as to avoid rote responses devoid of real understanding of the principle in question. Thus, novel ways of approaching these generalizations, not likely to be encountered in the public school classroom, were devised for each instrument.

Observations of language difficulties associated with immature language skills for some subjects prompted the introduction of a high degree of redundancy between verbal and visual information throughout the instruments. Most explanations and questions include verbal redundancy as well as simultaneous reference to visual or manipulative materials in order to communicate in both visual and verbal modes as consistently as possible. The implementation fo this policy throughout the six protocols was found to eliminate almost all instances of difficulty in communicating directions and questions.

After completing the second pilot study in the Spring of 1977, the format for all six instruments was stabilized. It was decided to include ten items plus one explanation or justification for each of the six instruments. Of the ten items, five would be aimed at detecting initial recognition of the principle in question while the second five items would be aimed at application of the principle or at a more flexible and extensive view of the generalization involved.

The following discussion of specific item selection and modification combines the results of all four pilot studies, the first two dealing with item development and the last two testing versions similar to the final products. The items for each of the six principles will be discussed separately. The third pilot study was conducted in the Spring of 1980 with 35 subjects including five kindergarteners, five fifth graders, eleven second graders and fourteen third graders. All subjects were attending the same elementary school in a southwestern Virginia town of somewhat more than 25,000 population. The fourth study included 48 beginning third graders, 20 attending one rural school and 28 attending a town school in the same southwestern Virginia school system.

The Addition-Subtraction Instrument

Using the order in which the instruments appear in Appendix A, the protocols for testing the Addition-Subtraction principle will be discussed first. This principle was stated in Chapter 1: "The addition and subtraction operations represent the joining and separating of disjoint subsets comprising the whole set."

It is generally recognized that most children can perform simple addition (sums under 20) and simple subtraction fairly reliably by the end of second grade, if not sooner. However, the diagnositc instrument had to distinguish between two different conceptualizations of the addition process, the subtraction process, and their relationship with one another. One conceptualization involves "counting up all the pieces" for addition and "taking away the specified number of pieces" for subtraction. The relationship between the two is just perceived as "counting in the other direction" (i.e. up or down) for the other operation. This conceptualization fails the user when faced with missing addend, missing minuend, or comparison subtraction problems since addition has only one direction, "up," and subtraction has only one direction, "down." Lindvall and Ibarra (1980), Matthews (1981), Underhill (1977), and Weaver (1971) have conducted research studies in which the just mentioned types of open addition and subtraction sentences were shown to be especially difficult for many first and second graders, as well as much older low math achievers.

The other conceptualization, called the "part-whole" generalization perceives the addition process as "part plus part equals whole" and the subtraction process as "whole minus part equals part." Furthermore one process is perceived as the inverse of the other and both are perceived as reversible, so that any of the three components of either process can be the missing component in a problem. This second conceptualization of the addition-subtraction process, then, does serve the user in solving missing addend, missing minuend, and comparison subtraction problems. Therefore, these types of problems were chosen as one type of item to be used in the Addition-Subtraction instrument.

In the third pilot study two forms of problems were tried for these types. The "story" problems which appear in the finalized instrument in the appendix, accompanied by blocks, animal cards and bird cards were retained with only minor modifications in wording. A study by Larson and Trenholme (1978) investigating the effects of syntactic complexity on problem responses helped to guide the initial selection of wording in the stories although subject responses were used to make final decisions on wording.

Another type of problem used for comparison subtraction was eliminated because of presentation time and because of possible confounding with a history variable. Originally, two dot strips (shown and used in Instrument 4 in Appendix A) having 17 dots and 11 dots in color-contrast groups of fives were presented side-by-side for a visual subtraction comparison. The subject was asked to tell how many more dots were on the longer strip. Many subjects resorted to counting all the dots on each strip—often more than once—and some still did not know what to do with their counts. Thus, the item consumed an inordinate amount of time for many subjects.

Even more of a problem was the confounding factor. Since the dot strips are used for Instrument 4, subjects who were tested on Instrument 4 before Instrument 1 had considerable prior experience with the dot strips while subjects who were tested on Instrument 1 prior to Instrument 4 did not have this earlier experience. The third pilot study demonstrated that this factor did affect performance on the comparison subtraction item, so it was deleted in favor of an additional "story problem" item comparing two towers of blocks. This solved the confounding problem, although a substantial proportion of subjects had difficulty with the comparison subtraction in this form also as would be expected from research reports on this subject by Matthews (1981) and Underhill (1977).

Generally, the "story problems" with manipulatives were found in all pilot studies to be attractive to the subjects and very effective in displaying the "disequilibrium" expected when children were faced with finding an unexpected component of the addition or subtraction process rather than the usual end component. Children who had ineffective generalizations of the Addition-Subtraction principle tended to move quickly through the warm-up problem asking for a sum of blocks after adding a second set to a tower, but then hesitated considerably when the sum of the blocks was given and the missing number of added blocks was asked for. In many cases, the problem had to be reread (always in its entirety) once or twice before the subject even made an attempt to answer. It was evident that the subjects "expected" to be asked for the sum and were confused when the process was clearly addition and the sum was not the missing component. These items were found to be very effective in distinguishing whether or not subjects had the more functional generalization of the Addition-Subtraction principle.

Since the "story problem" items serve as a measure of application of the Addition-Subtraction principle, items which focus more directly on the generalization itself were needed. Because initial understanding of the Addition-Subtraction principle depends on recognition of the part-whole relationship as shown in a study by Masse (1979), class inclusion items were chosen to measure the subjects' capabilities to make part-whole comparisons. Study of the extensive literature on class inclusion research provided potential test items, but did not bring attention to one major problem encountered in the first two pilot studies.

In keeping with findings in previous studies by Dodwell (1968), Inhelder and Piaget (1969:59-118), Markman and Seibert (1976), and Piaget (1952:161-184), two items were chosen. One presented the whole set with the parts distinguishable but combined (many paper clips and a few nails in a plastic bag). The second item presented separated parts (a few spoons and many forks separated by about 8 to 10 inches of space on the table) so that the subject was required to mentally construct the combined set (of silverware). As predicted by previous research, both early pilot studies demonstrated that more subjects were successful with presentation of the whole than with presentation of separated parts. However, in both of the first two studies, a significant number of false negatives were obtained on these two items. Typically these were obtained by subjects who performed well above average on all instruments but who had a very hasty response style. They tended to anticipate questions before the examiner completed asking them and responded after little or no reflection. Clearly subjects who answered all other Addition-Subtraction items correctly should not miss the class inclusion items.

When the second study confirmed the problem fairly early in the examination period, a modification was made to the presentation procedure in an attempt to overcome this difficulty. A third item presenting separated parts (a few cookies and many pretzel sticks) was added and the item presenting one whole set was relegated to being a warm-up item. The subjects were warned to listen very carefully because the question was going to be "tricky" before giving the warm-up item. Then if the subject missed the warm-up item, he was asked to repeat the question as it was asked. This might appear to cue the subject to the fact that he answered the warm-up item incorrectly, however it seemed that the subjects assumed that the examiner followed the same procedure for all subjects. More than one subject commented that the "tricky" aspect of the item was that the examiner was interested in knowing if the subject could remember the question rather than being interested in the answer. Thus the procedure did not appear to cue the subject to an erroneous answer, but only his inability to remember the question accurately if that was the case. Whenever the question was not repeated correctly, the examiner reread the question to the subject, allowing him to answer it again if he so wished. While one subject did change her answer to the warm-up question and then missed the next two class inclusion items suggesting a cueing response, it did not improve the score of the subject since the warm-up item was not scored. Further analysis on the last two pilot tests confirmed the evaluation that the class inclusion item warm-up procedure averted false negatives from hasty responders and did not improve scores of subjects who did not have class inclusion. These item response analyses were aided by examiner comments on observed behavior during administration as well as by comparisons with responses and observation comments on the "story problem" items.

Another factor prompting concern for the reliability of class inclusion responses was the fact that the questions were not open-ended, but had only two possible responses. These are the only questions in all six instruments which are not open-ended. The large body of published studies on class inclusion research along with its relevance to the part-whole generalization of the Addition-Subtraction principle supported the retention of the class inclusion items in this instrument. As previously explained, item response analysis indicated class inclusion responses were consistent with responses on other items in the instrument thus alleviating concern for their reliability.

The other type of item used to evaluate the subjects' addition-subtraction generalization directly is necessarily unconventional. Finding a means of discriminating between the part-whole perception of the addition-subtraction process and other perceptions which prove useful some of the time but not always was not easy. Anything involving the adding or subtracting of numbers of items was not effective in making the discrimination since in all instances, alternative solution methods could be employed. Finally, a procedure using parts of pictures (or figures) and whole pictures was devised employing plus, minus and equal signs to depict addition and subtraction equations in a non-numerical presentation. Figure 1 which follows gives an example of the final version of these equations.

Each set of five equations includes one correct example and four distractors. The distractors each include a different type of error. One displays proper disjoint subsets but an incomplete total set. Another displays a proper total set but subsets which are not disjoint. A third distractor has disjoint subsets which, when combined, comprise less than the total picture. A fourth distractor shows a

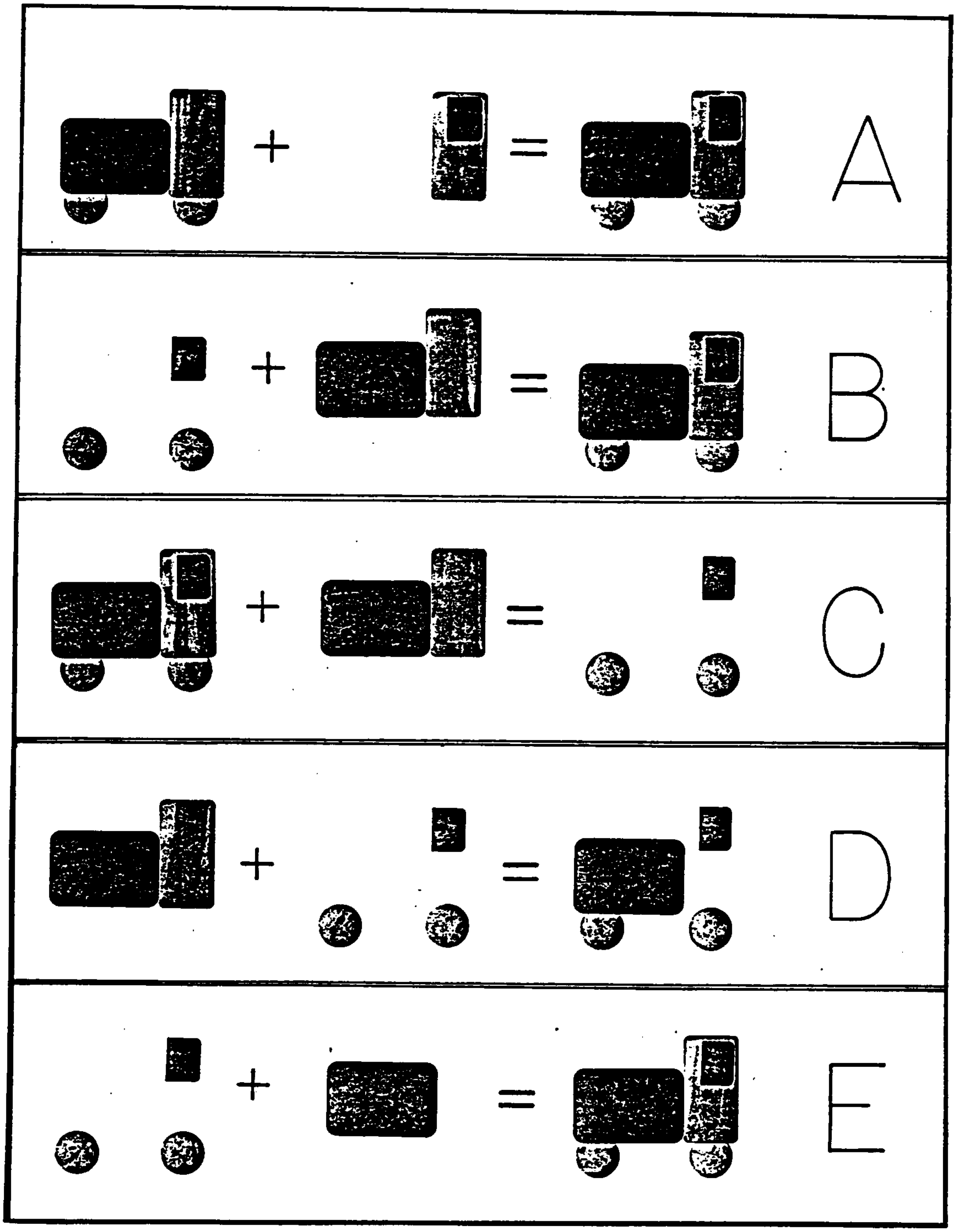

Figure 1. One Set of Picture Part and Whole Equations

correct equation of the inverse process but with the wrong operation sign. That is, if addition is being depicted, the equation has a plus sign between the first two figures in the equation, but the actual process shown is whole minus part equals part. Children were told that there were one or more correct examples in each set and to indicate whether each one was a good or bad example.

During the development of these items, a warm-up procedure was used initially in anticipation of difficulties with the unconventional nature of the items. It was found that the warm-up was unnecessary since children adapted very readily to this portrayal of addition and subtraction equations without numerals. The subjects who had difficulties with these items were not helped by the warm-up procedure and consistently had difficulties with at least some of the story problem items, suggesting that their difficulties were due to deficiencies with the addition-subtraction generalization and thus, the items were discriminating as desired.

Another change was made after the first two pilot studies with these items. Initially, line drawings of figures and parts of figures were developed using red and blue ink for contrast. Some children seemed to be slowed in their consideration of the sets of equations by the perceptual demands of the task. An article by Cunningham and Odom (1978) concerning how perceptual salience affects performance on certain tasks and another article by Lombard and Riedel (1978) reporting a study on the effect of color on the coding subtest of the Wechsler Intelligence Scale for Children-Revised (WISC-R) led to development of a new set of picture part-and-whole equations. These new items have figures constructed of geometric shapes with each shape being a different solid color. Colors were chosen so that adjacent shapes would present clear contrast to enhance perceptual salience of the parts within each figure.

When the new sets of picture part-and-whole equations were piloted, the contrast in time for scanning was noticeable. (The sets of equations are shown in Figures 3, 4 and 5 in Appendix A.) In fact, scanning was facilitated to the point that hasty responses became a problem unnoticed in earlier pilot studies. The final solution to this new problem was to ask the subjects to explain why each equation was judged to be either a good or bad example of the addition or subtraction process in order to slow them down long enough to give each equation careful consideration.

While responses by professional mathematicians to this unorthodox style of addition and subtraction equations has ranged from surprised delight to absolute horror, children seem to accept the novelty with enthusiasm. The violation of usual symbolic math conventions does not seem to bother them at all. Although mathematics educators tend to aid violating established rules so as to avoid confusion and reinforce the dependability and consistency within our system of mathematics, for this diagnostic procedure, the violation appeared to be the only way found to approach the Addition-Subtraction principle directly without eliciting rote responses conditioned in the classroom. The effectiveness of these items in obtaining clear explanations from young children about what the addition process is and is not seem to justify the transgression. Directions to subjects had to be carefully worded to avoid use of the terms "part" and "whole" entirely so that when subjects used these words in their explanations of the addition and subtraction processes, the examiner could be sure they were not "cued."

The Counting by Groups Instrument

The second instrument in Appendix A is the counting by groups diagnostic procedures. In Chapter 1, this principle was stated: "It is possible and often quicker to determine the cardinality of a set by counting by equivalent groupings rather than counting each separate item."

Initially, the statement of the principle included: "It is possible and often quicker and more reliable to determine..." However, the second, third and fourth pilot studies all demonstrated that even the subjects most competent with the counting by groups tasks mostly felt that counting one-by-one was more likely to give an accurate count. Reasons cited included that: "All the groups might not have the same number (of items);" "I know how to count one-by-one better;" and "Lots of kids don't know how to count by groups very well." Thus, confidence in the more familiar, first-learned counting procedure seemed to be an important factor. Restudy of Elkind's (1974) discussions on Piaget suggested that a judgement on the relative reliability (or error-potential) of the two counting procedures required more mature logical thinking than could be expected of primary school children. Elkind suggested that the ability to make such comparisons of relative probabilities does not usually appear until adolescence. Therefore, the "more reliable" judgement for counting by groups was dropped from the generalization to be studied.

The 1977 pilot study demonstrated that having children simply count in a rote fashion by twos, threes, fives and tens does not really assess competence. When subjects were asked to count buttons by 2's and cards containing 5 buttons each by tens, to find the total number of buttons, several of the subjects who did the rote counting correctly, produced button counts that were double the actual number of buttons because they moved only one button or one card of buttons for each count of two or ten. These subjects clearly did not understand the meaning of counting by groups. A few other conceptual errors were manifested with smaller frequencies. It was decided that specified group counts of a requested total of buttons, given more than enough objects to count, would be appropriate to assess initial recognition of the principle.

To assess more mature application of the principle, it was decided that task items should determine whether subjects would employ a counting-by-groups strategy of their own choice when the opportunity was presented. Items presenting natural groupings such as pairs of shoes, wheels on cars and fingers on hands were presented. Directions had to be developed to convey to subjects that both speed and accuracy were important in this counting activity.

The most difficult problem was found to be how to word the request to count the objects since many subjects believed that "count" meant only one-by-one. Eventually, telling subjects that we were investigating different ways of counting and wanted to see how they would count things to be both fast and accurate proved effective. Overt timing of each counting performance as well as visible recording of the obtained count were employed to remind subjects of the criteria of speed and accuracy for the activity. Even then, it was necessary to provide examiners with a standard response to subjects' queries as to whether they were "allowed" to count by groups to count faster even with the "cue" in the directions. Many subjects felt that counting by groups might be "illegal" for the activity. Some expressed the feeling that they were somehow "cheating" by taking such a shortcut. Occasionally subjects asked also if they could use multiplication to obtain the total count of objects. Multiplication was deemed to be a more mature form of counting by groups for the purposes of this study. Therefore examiners were instructed to repeat the original directions if asked permission to count by groups or multiply. (They were instructed to add to the directions that the subject should get the total count in whatever way he feels is best for that purpose.)

One other consideration was necessary for this instrument. In most cases, the items requiring less mature responses were presented first in each instrument in an attempt to put subjects at ease. However, in order to determine if subjects would choose to use counting by groups of their own volition without cueing them to the desirable response, the items in which the subject is asked to count buttons by twos, fives and tens were placed after the volitional counting items. The counting by threes items were dropped after pilot studies indicated that very few second and third graders could comfortably count by threes. The groups of four wheels were no problem as subjects could count by twos if not by fours. Asking for the more mature performance first was no threat to less mathematically mature subjects since they could simply count presented groupings one-by-one if that seemed best to them. All efforts were praised so that no stigma was perceived, whatever the counting schemes.

The One Equal in Value to Many Instrument

The third instrument in Appendix A is designed to evaluate conceptual understanding of the principle: One of something can be considered equivalent in value to many of something else. Any number of different models are possible. Chip trading models, measurement models, and money counting or exchanges are all good possibilities. However, chip trading is just playing a game by the rules. Although the metric measurement system best mirrors our base ten number system, primary children are generally unfamiliar with this system. Therefore it was decided to use the model most familiar and meaningful to primary children--our money system.

In devising two levels of tasks with money, it was felt that initial recognition of the principle would be demonstrated by simply finding the total value of progressively larger sets of coins and bills. To make sure the responses were not just rote, for each money collection, the subject would also be asked how many pennies would make up the same total value. The first two pilot studies established the order in which children usually acquire facility with counting coins and bills. First they learn simple equivalences starting with the smallest values although competence with one dollar bills sometimes precedes facility with using quarters. Afterwards children begin to develop grouping strategies to improve counting efficiency, starting with substituting two nickels for one dime and five pennies for one nickel. Four quarters for one dollar sometimes precedes substituting two dimes and one nickel for a quarter. Facility with fifty cent pieces was less consistent with substantial recognition problems, perhaps because of less frequent use in everyday money transactions, so fifty cent pieces were eliminated from the tasks entirely.

The more mature level of performance tasks was designed to require versatility and flexibility in the use of coins and bills. The early pilot studies had demonstrated that some subjects had very rigid ways of counting out needed amounts of money and encountered considerable difficulty or outright failure if substitutions for accustomed coin denominations were necessary. It was felt that true competence with money values and equivalences would include the ability to make up a requested amount with whatever denominations were available as long as the total value was sufficient for the need. Thus, the last five tasks required subjects to count out (or "pay") the same amount two or three times from a set of coins and bills provided for the purpose. The provided sets of cash were arranged to necessitate using different coin collections for each payment, thus requiring flexibility for success. Each task was scored as correct only if the requested amount was made correctly as many times as designated. As in the first part of the instrument, each subsequent item escalated in difficulty by adding to the number of denominations provided and/or the total amount to be paid. The last task required the addition of dimes to one quarter, which was consistently found to be an especially difficult combination for subjects who were otherwise competent. Subsequent pilot studies led to some minor modifications in the coin collections presented. No substantial difficulties with this instrument were encountered after selecting the basic design.

The N More and Less Instrument

The fourth instrument in Appendix A is designed to assess the level of competence with the principle: N more than M items is the N^{th} number above M in the counting sequence and conversely, N less than M items is the N^{th} number below M in the counting sequence. This principle is essentially Piaget's ordination-cardination synthesis. Accordingly, Piaget's (1952) classical staircase task was explored. However, success was found to be very dependent upon understanding the verbal expressions for ordinal numerals and many of the subjects displayed unfamiliarity with verbal expressions for ordinal numerals. Therefore an alternative procedure was needed for this instrument.

Two different approaches were initially explored in the first two pilot studies. One of the approaches involved incorporating length and number conservation trials on the same arrays of blocks so that the two conservations had to be coordinated and discriminated at the same time. Essentially subjects were asked to compare the overall lengths and the number of cars in two trains, one with orange cars and one with green cars, placed on parallel tracks. Equidistant pairs of crossing lines perpendicular to the tracks were provided to facilitate length comparisons. This procedure was effective in that there were no problems with directions and subjects sustained attention to the task throughout the series of unrelated modifications in number of cars and lengths of the two trains with repeated queries about length and number equality or inequality. Performance on the task also correlated rather well with place value performance. However, the task was only related to ordination-cardination synthesis in a general way, and it was also a fairly time consuming procedure.

The other approach explored was a procedure which does require judgements coordinating understandings of ordination and cardination quite directly. The procedure was begun with the subject ordering, from the least number to the greatest number, a collection of sets representing the cardinal numbers from one to eighteen. The sets were formed by two colors of dots in color groups of five, for easy discrimination of larger numbers. (See Nichols (1977).) The dots were evenly spaced on three-quarter inch wide strips of cardboard about fifteen inches long so that all of the strips could be comfortably arranged side-by-side on a table in front of the subject. (Fifteen of the strips are pictured in Figure 10 at the end of Instrument 4 in Appendix A.) Having the subject order the strips ensured that the subject was fully cognizant that the strips were completely ordered according to quantitative values.

After the strips were ordered, the subject was given a set of nine small plastic cubes sized to fit comfortably on the dots of the strips. The subject was asked to find the strip which had the same number of dots as he had cubes, and to place the cubes on the dots of the selected strip to show that the numbers of dots and cubes were indeed equal. Correction was made if the cubes were not on the correct strip. At this point, the actual scored trials began. Another cube was added to the cubes on the strip and the subject was told to predict which strip all the cubes would fit on exactly without counting all of the cubes or the number of dots on any of the strips. Nine more additions or subtractions of cubes followed with the number of cubes exchanged increasing to a subtraction of seven cubes on the last move. The subject was scored on the accuracy of his prediction before actually moving the cubes to the selected strip. Thus the subject was forced to predict the correct strip based on the relationship between their ordinal positions and their cardinal values.

The procedure appeared successful for testing the principle in a game-like procedure totally unfamiliar to the subjects. However, a few problems were encountered with this initial version of the procedure. The biggest problem was the fact that the procedure was a "learning" activity as much as it was a "testing" activity because the subject was able to confirm or reject his prediction at each step by moving the cubes to the selected strip. Since the objective was to diagnose what the subject did know when he came to the interview, "learning" during the diagnostic process would only have confounded the measurement. Another problem was that subjects could "cheat" by counting the cubes silently to themselves and then count up the strips from the number one strip to find the correct match rather than predicting from the number of cubes added or subtracted. Since this solution procedure circumvented the use of the principle being tested, it was a severe problem.

To solve both of these problems, a modification to the procedure was developed. A warm-up (unscored) item adding one cube was used to introduce the addition of cubes to the inital cube placement so that subjects would have a clear understanding of the "game" procedure before the complication was introduced. Then the cubes were removed and the ordered strips were covered by a piece of cardboard extending sideways far enough to cover the whole sequence, but allowing the tops and bottoms of each strip to be visible extending beyond the upper and lower edges of the cover. (See Figure 11 at the end of Instrument 4 in Appendix A.) Next, the examiner removed a few strips off the upper and lower ends of the sequence simultaneously, using a thumb sweep with each hand and storing the removed strips out of sight immediately so that the number of strips removed could not be ascertained by the subject. At that point then, the subject could no longer count from either end to determine correct placement since end values were no longer known. Furthermore, the cover would prevent confirmation after each prediction, eliminating the "learning" problem as well. In order to allow additions and subtractions of from two to eight cubes after removal of several strips from each end, it was necessary to expand the initial sequence from eighteen to twenty-seven strips.

Besides extensive modifications of verbal directions, two plastic arrow markers were found to be needed to mark the ends of the strip currently occupied and the strip selected for the next move to compensate for the perceptual difficulty introduced with partial covering of the strips. Several successive modifications were made to the verbal directions during the last two pilot studies to refine the explanation of assuming the cubes were placed correctly on the partially covered strip, even though the correctness could not be confirmed by matching cubes to dots. The final version of the instructions was found to be more effective than any previously tried, but still necessitated encouraging a few subjects to just try the first addition of cubes after the cover-up to clarify the verbal directions. The act of covering and removing strips seemed to upset the expectations of subjects who had just decided to count from the end to play the game successfully. Some subjects could see no alternative strategy and never really recovered from the coverup even though they did clearly understand the procedures of the game after commencing the scored trials of adding or subtracting cubes and predicting the new strip placement each time.

Since there was a question about the validity of using the conservation of length and number procedure to test for the ordination-cardination synthesis principle and because the "set-up" time for the number strips "game" was considerable, it was decided to use the same "game" procedure for both levels of questioning for this principle. Two levels of difficulty were clearly identified in the earlier pilot studies in that less mathematically mature subjects tended to be more secure about their predictions when one, two or three cubes were added or removed, but clearly were less secure or extremely doubtful when five or more cubes were added or removed. This dichotomy seemed to confirm that children are reluctant to generalize beyond their experience. That is, children are willing to generalize about adding or subtracting one, two and three much sooner than with larger numbers because they have more experiences of this nature that are easily tested and confirmed with the smaller numbers. As mentioned earlier in Chapter 2, Dodwell (1960) and Feigenbaum (1963) both found this small-large number dichotomy with children in the transitional stage of acquisition for an concrete operational task. After confirming the dichotomy again on the third pilot study, changes of two or three cubes were used for the initial recognition level, while changes of five to eight cubes were used for the more inclusive or applied generalization level.

The Place Value Instrument

The 5th Instrument in Appendix A tests understanding of the Place Value principle: The position of a digit in a numeral determines its value such that, in the numeral 369, the 9 has a value of nine, while the 6 has a value of six tens, and the 3 has a value of three hundreds. Three areas of investigation were chosen as relevant to the diagnositc task: the counting sequence of numerals, direct questioning about digit values in a numeral, and application of the principle to addition and subtraction with regrouping. These three areas were the dominant areas of investigation for determining place value understanding in the many place value studies reviewed in Chapter 2. The biggest problem encountered was distinguishing understanding from classroom conditioned rote responses. The Place Value principle is taught so intensely in second and third grade, that finding unique approaches to test understanding of the principle proved to be quite a challenge.

For testing understanding of the counting sequence, a ten by twenty magnetic number grid was developed for positioning the numerals from 20 to 219 in proper sequence after the first two pilot studies. (A facsimile is pictured in Figure 12 at the end of Instrument 5 in Appendix A.) The early pilot studies demonstrated that children were familiar with placing numerals up to 100 in rows of ten, usually putting 1 to 10 in the first row, 11 to 20 in the second row, etc. However, they were accustomed to filling in all missing numerals, in sequence, and seldom had placed numerals above 100 in such a grid.

For this diagnostic activity, "guide" numerals as shown in Figure 12 were placed on the grid before the subject was brought to the testing site. The first row began with 20 and ended with 29, the second with 30 to 39, etc. so that all numerals in the first row had a 2 in the tens place, in the second row, a three in the tens place, etc. The "guide" numerals displayed in their correct boxes consistent with the counting sequence were 20, 22, 26, 29, 33, 38, 41, 44, 47, 74, 80, 85, 89, 92, 137, 163, 166, 171, 178, and 185. These numerals were selected so that they bunched in rows 1-3, 6-8, 12, and 15-17. There were two "guide" numerals in each column of the chart so that a careful perusal would confirm that numerals in the same column had the same digit in the ones place, and that columns could logically be labeled from left to right, the zeros column, the ones column, the twos column, etc. in reference to the ones place pattern of numeral placement.

For the interview procedure, the row and column patterns of the number grid were <u>not</u> pointed out in the introduction. The subjects were told that the numerals already found on the grid were placed so that counting each box from 20 by ones would find them all in the right boxes for when their number would be counted. The examiner demonstrated by counting up to 33 while pointing successively to each box counted. Then, one at a time, the subject was asked to place the numerals 67, 116, and 202 in the correct boxes for the counting sequence. These numerals all fall in rows which have no "guide" numerals. Various arrangements of "guide" numeral and task numeral placements were investigated in the first two pilot studies. The third and fourth studies used the placements described here. The numerals 116 and 202 were designed to expose the most prevalent difficulty regions found in the two decades just above one hundred and two hundred. The 67 was found to be most difficult of the pilot study items under 100 and served to give the subjects an initial item which was not quite so upsetting for the less adept.

While the examiner could usually discern the subjects' methods for finding the correct numeral placement, such observations could not be wholly reliable. Since there was an extensive hierarchy of different placement strategies identified, scoring performance by strategy would have been both complex and less than perfectly reliable. Therefore, it was decided to score the subjects' performances simply by whether the numeral was placed correctly or not. Correct placement certainly indicated a recognition of the essential elements of how the digits in the numeral placed it in the number sequence. On the other hand, recognizing the row and/or column patterns assisted greatly in numeral placement and in checking to assure correct placement had been accomplished. The pilot studies showed that errors of placement were much less likely to occur when these more mature placement strategies were used.

For the second area of investigation, direct questioning about digit values in a numeral, considerable ingenuity was required to avoid any facsimile of the usual classroom examples. It was decided to try to elicit correct explanations of digit values by prompting the subject with incorrect models of introductory explanations. These examples were eventually all somewhat standardized in size and form, somewhat, by presenting them on unlined five-by-eight-inch laminated index cards.

The first example displayed the numeral 24 in the upper left section of the card with twenty-four "candy" hearts drawn to the right of the numeral in two rows of ten hearts and a third row of four hearts. Below the numeral 24, two blue circles were placed vertically below the digit 2 and four red circles were lined up vertically below the digit 4. (See Figure 13 at the end of Instrument 5 in Appendix A.) The subject was asked to explain what the digit 2 meant and what the digit 4 meant when they were put together in a numeral to show the number 24. Beforehand it was pointed out that the circles showed the counts of the digits 2 and 4 respectively, but that the six circles were much less than twenty-four "things" such as the twenty-four candies pictured at the right. The suggestion that the circles could be thought of as coins was also included in the introduction to the question. It was found that juxtaposing the abstract "coin" representation with the twenty-four hearts and mentioning that there were "only six circles in all" were effective for making this question sound different from the usual classroom presentation. The question of what the 2 and 4 are "counting" is a very common question in the subjects' classrooms but was seldom recognized as such immediately in this presentation.

The second example in this section is a simple model for the digits in the numeral 24. The numeral is centered at the top of the card with two dimes lined up vertically under the 2 and four dimes lined up vertically under the 4. The subject is asked if this card is "a good way to show what the 2 means and what the 4 means in the numeral 24." They are also asked to explain why they think so and to suggest alternatives if they think something else would be better for this purpose. The third example is of the same type. The numeral 124 is modeled with one dime, two nickles, and 4 pennies and the question to the subject is the same. These examples are pictured in Figure 14 and 15 at the end of Instrument 5 in Appendix A.

This use of some inappropriate subsets of coins on each card was found to be very effective in eliciting explanations of what numerical values each digit represents. Typically subjects would say that the model was "okay" as presented, "but...". And then, they went on to suggest improvements in the model and to explain why the improvements were better. Only improvements which precisely model the digit values as well as the total amount were considered successful. In the early pilot studies, more appropriate models were tried, both with coins, and more abstractly, with colored circles. However, inconsistencies with other responses about digit values and with explanations of regrouping procedures indicated that some subjects were scoring one or more false negatives. The problem seemed to be that the examples were not reliable in prompting subjects to relate the questions to what they knew about the place values of numerals. The incorrect models using familiar coins proved to be the most reliable in eliciting explanations when they could be expected to be forthcoming.

In order to further guard against false negatives, all three examples were kept visible to the subject on the table until all had been answered to the subject's satisfaction. Additional comments on any of them were permitted until all were finished. Sometimes it was the collection of examples that jogged their perceptions to realize that they did know of better models and why they were better models. The subjects' explanations of hundreds, tens, and ones place values appeared to be both spontaneous and lucid suggesting genuine understanding of the place value principle of our system of numeration.

The third type of Place Value task investigated was the application of the principle to addition and subtraction with regrouping. The early pilot studies demonstrated a hierarchy of skill acquisition in this area. The pilot study results confirmed the difficulty ordering and systematic error patterns reported by Cox (1975). Addition of two digit numerals, with regrouping required just once from the ones place to the tens place, was the type of problem with the highest frequency of successes. Next came subtraction with two digit numerals requiring regrouping of one ten to ten ones. Another stage of difficulty was represented by a problem adding a three digit numeral to a two digit numeral where regrouping was required twice with a zero in one addend and in the sum. The hardest example selected was subtraction with two three digit numerals with one zero in the minuend and two regroupings required. With this last problem, instead of asking the subject to do the problem, the problem was shown completed, but with the most common error included for the answer. The subject was asked to check the problem to see if the student who did it was right or wrong. If the subject found it wrong, he was asked to show how it should be done correctly. The exact problems selected for use in Instrument 5 are shown in Figure 16 in Appendix A.

The problem types are representative of the range of regrouping problem types taught in second and third grades. The success rates on the final selection of regrouping problems ranged from 91% on two digit addition to a low of 17% on three digit subtraction. The three-digit subtraction performance was clearly affected by when the testing was done, with Fall testing of second and third graders producing a much lower success frequency than end of the school year testing. (The 17% success rate is an average of these two different response rates.) Thus, while a few second graders could do all types of regrouping problems, subjects who had finished third grade were much more likely to have success on the more difficult types.

The Clock Reading Instrument

The sixth instrument in Appendix A tests understanding of the Clock Reading principle: The hands tell what is being counted and what values are taken by each numeral on the clock. For the short hand, hours are being counted and each numeral counts one hour. For the long hand, minutes are being counted and each numeral counts five minutes with the short lines between the numerals marking the one minute intervals.

An extensive search of the research literature on clock reading performance of primary children produced no studies identifying any kind of a learning progression or skill hierarchy. For that reason, classroom textbooks and lists of curriculum objectives from a local school system were used to design a range of items for the first pilot studies.

After assessing responses to different types of Clock Reading tasks on the early pilot studies, several stages of progressive acquisition of understanding were identified. The first successes identified were for reading exact hour and half-hour times correctly, followed by quarter-hour readings. However, "quarter-past" and "quarter-of" readings clearly did not consistently indicate an understanding that fifteen minutes had elapsed since the exact hour or that it would be fifteen minutes until the next hour, respectively. In many cases, verbalizations corresponding to certain configurations of the hands on the clock had been memorized without any understanding of the dynamics of how changes in clock hand positions corresponded to changes in time. Therefore, scoring was commenced with reading times to exact five minute intervals such as 4:45 and 5:25 so that some degree of understanding of minute, as well as hour, measurement would be indicated.

For the transitional level of understanding, it was decided that any one correct reading of clock settings would be judged sufficient for success. Two warm-up questions of 3:30 and 4:15 settings were used to make less adept subjects more comfortable. Then three settings to exact multiples of five minutes were included as the first three tasks to be scored as successes or failures. These settings were 4:45, 5:25, and 7:40.

The next three settings of 7:38, 10:18, and 11:44 could earn two different scores. Two success points were scored if any two different times were read correctly one way. That is, they were scored correct if read correctly as time "before the hour" or as time "after the hour." Thus, 7:38 could be read as "seven thirty-eight" or as twenty-two minutes before eight." Correct verbal variations of these such as "thirty-eight minutes after seven" and "twenty-two 'til eight" were also scored as successes.

After the subject gave one reading for an "exact minute" setting such as 7:38, the subject was asked if he could read the time another way--as time "before the hour" if he had first given an "after the hour" reading or vice versa. An additional success score was given for each of the three settings--up to two successes for reading correctly one way, and up to three successes for reading the settings correctly both ways.

The pilot studies demonstrated that children first learn to read times to exact minutes in a fairly inflexible manner. They are unable to give both "time before" and "time after" the hour readings on the same clock setting, at first. Therefore, this inflexibility was felt to be an appropriate dividing line between a "complete competence" level of understanding and an initial level of understanding of all the components of the Clock Reading principle.

For the final two items of the ten item procedure, subjects were asked to set times as well as read the times set. However, the subject was asked to set the time that the clock would show X minutes before or after an initial setting presented. The two requested settings selected were forty-one minutes earlier than 12:08 and twenty-five minutes earlier than 1:22. These were found to be good indicators of how well children understood the dynamics of clock hand movements for indicating the measurement of elapsed time. On the pilot studies, facility with this last type of question tended to follow pretty closely upon the acquisition of flexibility in reading clock settings both ways. This seemed to indicate that when children finally put all the components of the Clock Reading principle together, they do then have the ability to apply their understanding to different time measurement situations in a fairly flexible manner.

Interestingly, in light of the debates on the use of digital time pieces, success rates on comparable clock reading tasks did appear to diminish with each successive pilot study from the Winter of 1977 to the Fall of 1980. Some subjects in each study volunteered the explanation that they preferred digital clocks and watches for telling time because they were so much easier to read. Nonetheless, these selected Clock Reading items did appear to achieve the objective intended--to evaluate the subject's degree of understanding of the principle for reading an analog clock.

Explanation Tasks for Each Instrument

In addition to including ten tasks in each instrument, it was decided to require an explanation of the principle to award a competence level rating for each principle tested. Since second and third graders are often not consciously aware of the explicit generalizations they are using to solve numerical tasks, it is not reasonable to simply ask them to state the "rule" they are using. Instead, for each instrument subjects are asked to explain the reason behind their procedure used on a specific task item--usually the tenth task item for each instrument. By asking the subject to justify their choices on a specific task, the "rule" can be stated as it relates to the details of the task, rather than as a true generalization.

During the piloting of the explanation tasks, it was expected that most subjects who answered at least eight out of ten task items correctly on an instrument would have been able to give a satisfactory explanation. On the other hand, most subjects who answered far fewer items correctly should not have been successful with the explanation. These two criteria were used together to develop appropriate explanation tasks for each instrument.

Several problems were encountered with the initial explanation trials piloted. With the Addition-Subtraction and Place Value instruments, presenting mistakes to be corrected by the subject was the solution for too few successful explanations on the early trials. As discussed earlier, subjects were much more likely to verbalize an explanation of the correct procedure when presented with an incorrect procedure to elicit their protests. With the Counting By Groups instrument, the problem was that too mature a generalization was being expected. As explained in the development of that instrument, the adjustment was made in the statement of the principle, in order to adjust expectations to those appropriate to the maturity of the subject. The final versions of the explanation tasks are shown at the end of the text for each instrument in Appendix A.

DEVELOPMENT OF SCORING PROCEDURES

The text of each instrument is followed by a score sheet for recording responses during the subject interviews. These were designed primarily to aid the examiner during the interviews. In addition to providing a place to record the details of each response, the format also provides reminders of item content for the examiner. The reminders assist both in presentation and in judging correctness of responses where such judgements are needed to determine how to proceed further.

Each score sheet format evolved during the pilot studies to facilitate fast and accurate recording of all pertinent details of the subjects' responses and observed behaviors. Extensive corroborating comments and behaviors were helpful in assessing effectivenss and validity of task items during the pilot studies. The score sheet information also provided a means for rechecking the examiner's judgement of the correctness of each task response.

The score sheets for all six instruments were compiled in a recording booklet for each subject in a study. However this booklet, while excellent for detailed analysis of specific task responses, was rather unwieldy for reporting subject performance to teachers and parents--a necessary byproduct for obtaining volunteer subjects for the research studies. Therefore, a one page profile sheet was developed which indicated the proportion of successful responses on each type of task included within each of the six instruments. A copy of this profile sheet, followed by an explanation sheet to aid in interpretation of the profile sheet, is shown in Figure 17 in Appendix B.

For the purposes of the research analysis, each subject's response profiles had to be converted to scores indicating levels of understanding as defined in Chapter 1. Initially, it was decided to designate "initial understanding" (also called transitional) for at least four successes on the first five trials for each instrument, and to designate a "competence" level for at least eight successes out of the total of ten trials as long as a successful explanation was also obtained for each instrument. However, a non-negligible number of subjects produced considerable scatter in their performances, achieving success on two or three of the first five trails but also achieving one or two successes on the sixth through tenth trials. After analyzing subject comments and examiner observations, it appeared that these subjects were demonstrating an "initial understanding" of the principle being tested so a modification of the scoring procedure was devised.

Subsequently, it was decided that successes on any four trials out of the ten trials would indicate the "transitional understanding" level (T). The same criterion of at least eight correct items out of ten plus an explanation were required for the "competence" level of understanding (C). Less than four successes was classified as the "no understanding" level (N). Analysis of score frequencies for the Spring and Fall 1980 pilot studies confirmed the selection of the "breakpoints" of four and eight successes with low frequencies on either or both scores of three and four, and likewise seven and eight successes. Lack of breakpoints between seven and eight successes on the Addition-Subtraction instrument on the Fall 1980 pilot study was attributed to the "hasty response" problem discussed earlier in this chapter. Introduction of the picture part equations using colored block designs resulted in a substantial increase in hasty response errors for that particular pilot study. The problem was corrected late in that study, but too late to avoid high frequencies of scores of seven and eight successes which mostly should have been nines and tens according to previous performance patterns. The desired breakpoint was observed in the 1982 study, confirming that the problem had been properly diagnosed and corrected.

Two types of reliability checks were made on the scoring for the Fall 1980 pilot study. This was the first study in which examiners other than the principle researcher were used. Two other examiners, one undergraduate and one graduate level education major, conducted all of the subject interviews for this study. The principle researcher independently rescored all of the response booklets. Virtually no discrepancies were found in scoring the 10 task responses on any of the six instruments. A policy of terminating trials on the place value instrument because of low success rates on the first half was judged to be an undesirable procedure resulting in fewer scores in the four to seven successes range than otherwise would have been expected. This policy was changed to prevent termination until all ten tasks had been presented except on the clock reading instrument when the subject clearly was unable to succeed at any subsequent tasks.

A separate reliability checking procedure was used for judging the success of explanations for each of the six principles. All explanations were tape recorded as well as being recorded in writing on the score sheets. The principle examiner independently scored all the written explanations and then checked the tape recorded explanations of a random sample plus all explanations where any scoring discrepancies or questions about procedures were encountered.

The two examiners hesitated to make a judgement of success or failure on a few explanations. The principle researcher was able to score a few of these definitively after checking the tape recording. However, in most cases, it was determined that the examiner probably should have probed further to try to get a more definitive explanation from the subject. Modifications in directions to the examiner were made to try to correct this problem. In essence, one can question whether any failed explanation is merely incomplete or was the best that the subject could muster. However, in over 90% of explanations no judgement discrepancies were encountered, since subjects clearly had succeeded or were unable to give a correct explanation. With the remaining explanations, they were scored as unsuccessful unless all required elements were given by the subject at some point during the administration of the instrument. Extra care was taken on the final study to insure that examiners probed explanations until they were satisfied that they had obtained the "best justification" the subject could give.

RELIABILITY OF THE INSTRUMENTS

The design of the instruments was definitely not planned to conform to conventional reliability measurement. Since only three different scores—no understanding (Z), initial understanding (T), and competence (C)—are possible, only non-parametric measures applicable to test-retest data are applicable. Therefore, during the Spring 1980 pilot study, retest data was obtained. Because of the extensive time requirements (approximately one and one-half hours) to administer the six instruments to each subject, retesting each subject on all six instruments was impractical. Furthermore, the second time through is not nearly as exciting as the first time, so that attention span became a real concern for the retests. For both reasons, it was decided to do only a partial retest on each subject, rotating the instruments being retested in order to get a reasonable sample on each.

Because instruments were grouped for the retests to equalize time requirements and because a few subjects were unavailable for retests within one and one-half weeks of the completion of the original administration of the six instruments, the number of retests per instrument varies from seven to nine. The total number of subjects in the pilot study was 37, so the retest samples ranged from approximately one-fifth to one-fourth of the total sample.

Of the non-parametric reliability statistics available, Spearman's Rho was selected as the most promising because Siegel (1956:202-213) has developed a statistical procedure for adjusting Spearman's Rho for ties. With only three different possible scores--Z, T, and C--ties in rank are the rule rather than the exception. Even with the tie adjustment procedure, the resulting reliability statistics did not always seem to correspond logically with the simple proportions of agreement on the test-retest data. Therefore, both the reliability statistic and the simple proportion will be given for each instrument, as shown in Table 1.

The Counting by Groups instrument had a test-retest agreement proportion of six out of eight or 75%. The reliability coefficient is .74. While the Addition-Subtraction and One Equal to Many instruments both had test-retest agreement proportions of seven out of eight or 87.5%, their reliability coefficients are .77 and .76 respectively. The N More and Less instrument proportion was six out of nine or 67%, with a reliability coefficient of .75. The Place Value and Clock Reading instruments both had Spearman's Rhos of 1.0 since the test-retest agreements were 100% at seven out of seven and nine out of nine, respectively.

Table 1

Test-Retest Scores Obtained in the Spring 1980 Pilot Study, Accompanied by the Spearman's Rho Reliability Statistic and the Simple Proportion of Agreement (SPA)

SUBJECT NUMBER	TEST SCORE[b]	RETEST SCORE	RHO/ SPA	SUBJECT NUMBER	TEST SCORE	RETEST SCORE	RHO/ SPA
	ADDITION-SUBTRACTION				COUNTING BY GROUPS		
320	T	T		220	N	N	
362	T	T	$\rho = .77$	290	T	T	$\rho = .74$
352	T	T		232	T	T	
262	T	C		322	T	C	
370	C	C		282	T	C	
332	C	C	SPA = 7/8	540	C	C	SPA = 6/8
030	C	C		520	C	C	
360	C	C		240	C	C	
	ONE EQUAL TO MANY				PLACE VALUE		
220	N	N		010	N	N	
240	T	C		252	T	T	
232	C	C	$\rho = .76$	230	T	T	$\rho = 1.0$
322	C	C		222	T	T	
520	C	C		260	C	C	
290	C	C	SPA = 7/8	342	C	C	SPA = 7/7
540	C	C		532	C	C	
372	C	C					
	PLUS OR MINUS N				CLOCK READING		
242	N	T		022	N	N	
050	T	N		242	N	N	
340	T	T	$\rho = .76$	042	N	N	$\rho = 1.0$
022	T	T		050	N	N	
042	T	C		340	T	T	
350	C	C	SPA = 6/9	512	C	C	SPA = 9/9
560	C	C		330	C	C	
512	C	C		350	C	C	
330	C	C		560	C	C	

[a] Spearman's Rho procedure was adjusted for ties.

[b] N: No understanding, T: Transitional understanding, C: Competence

The lower agreement proportion on the N More and Less instrument was probably partially due to a learning effect since two out of the three disagreements had higher retest levels of understanding. As discussed earlier, the testing activity was an effective learning activity as well, although the procedure had been modified to avoid the learning effect as much as possible. However, redoing the activity one week later may have given subjects some opportunity for reflection. The closeness of the four non-unitary reliability coefficients (.74 to .77) is somewhat surprising since the agreement proportions range from 67% to 87.5%. Analysis of the computation procedure suggests that the large number of ties tends to affect the coefficient somewhat erratically. It does seem strange that the reliability coefficient for Counting by Groups is lower than the coefficient for N More and Less although the Counting by Groups proportion of agreement is higher. It appears that the larger number of subjects for the N More and Less retest outweighed the larger number of ties.

Nonetheless, considering the small size of the test-retest samples and the small number of test items in each instrument, the reliabilities obtained were judged acceptable. The scoring procedures for assigning levels of understanding did seem to have the expected result of damping the effects of less than optimum performances.

VALIDITY OF THE INSTRUMENTS

Three types of validity evaluations were conducted during the development of the instruments in addition to the thorough literature search which was conducted to assist in item design and selection. The first type of evaluation was a measure of agreement between levels of understanding scores obtained by administration of the instruments and levels of understanding estimated by the subjects' classroom teachers. The second type of validity evaluation was a study of age and grade appropriateness of the instrument items. The third type of validity evaluation conducted was a face validity assessment by authorities. Each of these will be discussed in turn.

Teacher-Instrument Agreement

The first type of evaluation was administered with the cooperation of the subjects' classroom teachers during the Spring 1980 pilot study. A one page evaluation form (Figure 18 in Appendix B) was developed to be completed by classroom teachers within a few days of the completion of each subject's testing on the instruments. The evaluation form was organized to facilitate teacher estimations of each subject's current levels of conceptual development on each of the six principles being investigated. While each teacher was given an extensive verbal explanation of the research design and objectives, a brief list of behavioral criteria was included for each principle on the evaluation form to remind teachers of the research objectives at the time the developmental estimates were being made. Since the interviews were conducted over a period of four months, with an Easter vacation intervening, these reminders were important to produce consistency in the estimation process.

In spite of efforts to obtain a 100% response by teachers and excellent cooperation on their part, five out of the thirty-five evaluations were not completed within the time frame designated. These omissions were primarily due to teacher absences and the intervening school vacation. Nevertheless, the thirty evaluations obtained appeared to be a representative sample with all five teachers fairly evenly represented. The biggest problem with the evaluation procedure seemed to be that teachers were much more accustomed to assessing skill performance of students than to assessing conceptual levels of understanding. Each teacher expressed reservations about making such estimates.

Certainly, the most difficult problem encountered in the teacher assessment procedure was communicating what the three levels of understanding were meant to imply without describing the specific tasks being administered in the research interviews. It was felt that any communication of specific information about the actual instrument tasks might compromise the research process by prompting modifications to classroom instruction during the interview process. Therefore brief descriptions of the levels were included as headings for a check-sheet response but more definitive description was not given.

Contingency tables comparing teacher estimates and tested levels of understanding for each of the six principles are displayed in Table 2. Comparison of the frequencies in the triagles of cells above and below the diagonals of each table results in the following observations. For all instruments except the Plus or Minus N instrument, teacher estimates tended to exceed measured performance levels, with overestimates greatly exceeding underestimates. For the Plus or Minus N instrument teacher underestimates slightly exceeded overestimates, perhaps because teachers were less familiar with this principle. In all but one contingency

Table 2

Contingency Tables Comparing Teacher Estimates and Tested Levels of Understanding for each of the Six Principles

Addition-Subtraction

Test Score	Teacher Estimate N[a]	T	C
N	0	5	0
T	0	3	3
C	0	3	16

Counting By Groups

Test Score	Teacher Estimate N	T	C
N	3	3	0
T	2	1	9
C	0	1	11

One Equal to Many

Test Score	Teacher Estimate N	T	C
N	4	4	1
T	0	1	3
C	0	0	17

Plus or Minus N

Test Score	Teacher Estimate N	T	C
N	1	1	2
T	0	6	2
C	2	5	11

Place Value

Test Score	Teacher Estimate N	T	C
N	2	5	0
T	1	5	5
C	1	0	11

Clock Reading

Test Score	Teacher Estimate N	T	C
N	5	4	0
T	1	3	5
C	0	1	11

[a] N signifies no understanding, T signifies transitional understanding, and C signifies competence.

table, the number of teacher-test agreements (the sum of on-diagonal cell frequencies) exceeded the number of teacher-test disagreements (the sum of the off-diagonal frequencies).

Overall teacher estimate-subject score agreement was 62% on the 180 different estimations rendered. By instrument, the Counting by Groups came in with the lowest agreement at 50%. The Place Value instrument had 60% agreement, while the Addition-Subtraction, N More and Less, and Clock Reading instruments all received 63% agreement, and the One Equal to Many (money) assessment had 73% agreement. The poor agreement level on the Counting by Groups instrument was probably due partly to the behavioral description given. No mention of correct application was made. Overwhelmingly teacher estimates exceeded actual performance scores on the disagreements for this instrument (80% overestimated). In general, teachers tended to be overly optimistic about their student's performance on the 38% of estimates which did not match actual scored performance levels. In all, 76% of disagreements were overestimates rather than underestimates by the teacher.

Since the problem was roughly equal across all six instruments given the crudeness of the comparison procedure, it was felt that the validity of the instruments was not seriously impugned. There was substantial evidence that the instruments were measuring understanding of the principles as intended, given the considerable agreement between teacher estimates and scored performances and taking into account the vagueness of the definition of levels of understanding. (Only 5 out of the 68 instances of diagreement were discrepancies of more than one level.) In addition, teacher comments when results were given and explained at the end of the study were overwhelmingly positive in expressing confidence in the validity of the obtained scores. In many instances, identified weaknesses in individual performances on specific interview tasks were found helpful to teachers in explaining unexpected inconsistencies in performance by individual students on classroom tasks.

Age and Grade Appropriateness

A second type of validity assessment was conducted in the Spring 1980 pilot study to evaluate the age and grade appropriateness of the instrument items selected. Along with twenty-seven second and third graders, five fifth graders and five kindergarterners were interviewed as subjects. While all second and third graders in the participating classes were invited to participate, and all those receiving parental permission were included, the kindergarteners and fifth graders were preselected by their teachers in order to obtain a representative sample with only five subjects per grade. The teachers were instructed to select one of the strongest students in math, one of the weakest, and three representing a range of performance between the first two selected.

The objective for including these additional subjects was to demonstrate that the test items on the instruments were specifically targeted at understandings of math principles usually acquired at least in part, in second and third grades. If this acquisition period has been accurately pinpointed by the test items selected, then even the weaker fifth graders should achieve competence level scores on all six measures. On the other hand, even the strongest of the kindergarteners would be expected to achieve no better than initial understanding levels on one or two of the six instruments. Typically some kindergarteners are beginning to develop understanding of the addition and subtraction processes, especially where manipulative problems are involved.

The performance of the kindergarteners and fifth graders corroborated expectations completely. The most exceptional kindergarten subject achieved a competence level score on the Addition-Subtraction instrument and an initial understanding level on the N More and Less instrument but scored no understanding on all other instruments. A second kindergarten subject achieved initial understanding levels on three instruments--Addition-Subtraction, N More and Less, and Counting by Groups. This subject explained that she had an older brother in second grade who taught her to count by groups. Two other kindergarten subjects achieved initial understanding scores on the N More and Less instrument but no understanding levels on everything else. It was interesting to find that kindergarten subjects as a group related addition and subtraction situations to the number system (the N More and Less instrument) more readily than to story problems with small sets presented manipulatively. All kindergarten subjects scored no understanding levels on the One Equal to Many, Place Value, and Clock Reading instruments.

Also, as expected, all five fifth graders, even the weakest math student, achieved competence level scores on all instruments. In fact, only one item was missed by one fifth grade subject on one of the six instruments. The other four fifth grade scores were perfect scores with no mistakes. Thus, the age-grade validity of the instruments was strongly supported by these data.

Face Validity Assessment

The third type of validity assessment conducted was a review of the instruments by authorities done in the Spring of 1982 when the composition of the instruments had been tentatively finalized and preparations were being made to take the actual research study data that Summer. Ten authorities in the fields of math education or educational measurement were asked to review three instruments each. Their names and positions at the time are listed in Appendix B with an indication of their response. Because of the length of the text of the instruments, it was felt that asking one person to review all six instruments would be a major imposition. Response to the requests for reviews was disappointing, although those who did respond gave some very helpful comments. Two complete reviews for the Addition-Subtraction, Counting by Groups, and One Equal to Many instruments were received as well as one review of just the Addition-Subtraction instrument. Only one review of the N More and Less, Place Value, and Clock Reading instruments was received, however, it was found to be a quite thorough review with very insightful and helpful comments.

Generally, the responses were positive, judging that each item on each instrument did relate to the concept being tested and would be a valid indicator of the subject's understanding of the concept. Therefore, only the reservations and other comments will be discussed in detail.

On the Addition-Subtraction instrument, one of the three reviewers felt that the first scored story problem, a comparison subtraction example might confuse subjects and therefore would lack validity. As discussed earlier in this chapter, while comparison subtraction was found to be more difficult and less familiar to our pilot study subjects overall, it is one type of subtraction problem and therefore has validity in assessing a child's understanding of the subtraction principle. The other two reviewers did not discern any problem with the inclusion of the comparison subtraction example. One reviewer suggested that the variety in types of addition and subtraction problems would limit conclusions that could be drawn from the results, but since the objective was to measure the overall utility of the subjects' generalizations about the addition and subtraction processes, the variety of example types seems especially appropriate in determining how well their generalizations work.

On the class inclusion questions, two reviewers approved them with no reservations while the third reviewer expressed substantial reservations. The fact that the class inclusion items required a choice of two answers rather than being open ended was felt to compromise reliability. It was also felt that the warm-up procedure for preventing false negative responses would produce bias towards correct responses. These issues had already been carefully considered and were discussed earlier in this chapter. The final objection, by one of the three reviewers was that plus and minus signs should not be used with the picture parts and wholes since these were representing union of sets rather than addition or subtraction. While these objections certainly had merit, they had been considered carefully and decisions made as described earlier in the chapter. A few comments about directions to examiners and precise wording of story problems were heeded and appropriate modifications were made.

For the Counting by Groups instrument, only one reservation was expressed by one of the two reviewers. This related to the explanation of the principle at the end of the administration of the ten items. One of the reviewers felt unsure as to whether a successful response by the subject would be a valid indication of the subject's understanding of the principle. This reservation was considered carefully and subjects' explanation responses on the pilot studies were reviewed again. It was found that responses did state the principle targeted, and that, coupled with the responses to the ten previous items, the response did indicate understanding of the principle.

For the One Equal to Many instrument, the only reservation expressed by one of the two reviewers was that since only one model (our money system) was used, researchers should not generalize the findings to assume global understanding of the One Equal to Many principle. This point is valid. As discussed earlier in the chapter, the concern is primarily with how understanding of the One Equal to Many principle relates to understanding of the Place Value principle so that a model representing the base ten system was deemed most appropriate and global understanding of the principle is not an issue. Thus, it is simply necessary to recognize that the instrument diagnoses only a limited understanding of the principle stated.

On the N More and Less instrument, the reviewer expressed the same reservations discussed earlier about preventing subjects from cheating on the game by using counting of the cubes and dots. The cover-up procedure was successful in preventing the use of this strategy. This was confirmed both by observation and subject comments. The other reservation was about the directions given to the subject. They were reworded again before the final data were taken, but examiners were instructed to simply push the subjects into commencing the game if the directions were not clear after one repeat. The directions were revised at least four times in attempts to improve calrity, but no version tried eliminated all confusion. Just starting the game seemed to solve the confusions that were unresolved earlier. The subjects quickly related the directions to the game process once they got started.

The reviewer felt that the number board items on the Place Value instrument were not valid measures of the subject's understanding of the principle because success could be achieved through ability to count, only. As discussed earlier one numeral in the 60's and two numerals greater than 100 were selected because these were identified on the early pilot studies as being especially difficult. While rote counting is the first step in acquiring understanding of the numeration system, a much more extensive recognition of the organizational patterns of the number system must be, at least intuitively, recognized before higher decade numeration and counting above one hundred is understood well enough for insertion of missing numerals as required in the number board tasks. While total competence with the principle is probably not required for success on these tasks, the pilot studies suggest that initial understanding probably is required for the items selected.

On the Clock Reading instrument, the validity of the tasks for measuring the principle was confirmed by the reviewer. But, it was felt that the items might be too difficult for the targeted age group. The pilot studies confirmed that success scores were lowest on this instrument. Observations from the studies suggest that while most subjects were cognizant of the different elements of the Clock Reading principle, many of them had difficulty in applying them all simultaneously to obtain correct clock readings. Since these elements are all taught during the second and/or third grade, the test items do seem appropriate, although, as the research hypothesis suggests, it may be too early to expect a high incidence of competence with their application. Thus, this particular reservation seems to relate more to the validity of the hypothesis than to the face validity of the instruments.

In conclusion, while the instruments were found to be less than perfect by the reviewers, no serious problems with their face validity were uncovered. The series of four pilot studies used to ensure a high degree of validity and reliability did much to uncover and resolve potential problems.

CHAPTER 4

METHODOLOGY

The two null hypotheses to be tested both involve levels of understanding on each of the four prerequisite principles:

(a) one of something may be counted as equal in value to many of something else;

(b) one may determine the number of objects in a set by counting by equivalent subgroups more quickly than by counting each individual item;

(c) N more than a quantity M items is given by the number which is N steps up from M in the sequence of counting numbers, and similarly for N less than M objects being N steps down from M in the counting sequence; and

(d) the addition and subtraction processes deal with the joining or separating of component subsets and the whole set which they comprise.

The null for the primary research hypothesis is that the proportion of children having a higher performance level on the Place Value principle than on the lowest prerequisite performance will equal the proportion expected by chance if there is no relationship. The null for the secondary research hypothesis is that the proportion of children having a higher performance level on the Clock Reading principle than on the lowest prerequisite performance will equal the proportion expected by chance if there is no relationship. The alternative hypothesis for these null hypotheses is that the proportion of children whose performance scores fall above the diagonal on the corss tabulation of criterion and prerequisite performance levels will be less than expected by chance.

The procedure to be used for testing these null hypotheses is to measure individual children's levels of understanding of the four prerequiste principles and their levels of understanding on the two criterion measures. Then, for each child's profile, the lowest performance level recorded for any of the prerequisite principles is recorded as the prerequisite score. If either of the null hypotheses is rejected, it is appropriate to investigate whether the relationship between performance on the prerequisite principles and performance on the criterion measure is as postulated.

On the other hand, if either of the two null hypotheses is found to have insufficient evidence to reject it, then it would seem worthwhile to conduct a similar analysis using the individual profile data for only three of the four principles predicted to be requisite in order to determine if the null would be rejected for this reduced set of possibly prequisite principles.

For all planned statistical tests, the data needed are a collection of individual profiles giving a measured level of understanding for each of the six numerical principles in question. It is important that the complete set of profile data for any one individual be obtained in as short a span of time as is practicable, to avoid any intervening history factors, such as classroom instruction or incidental learning outside of the classroom, changing the individual's level of understanding on any of the measured principles during the process of measurement.

SUBJECTS

In order to make a valid test of the null hypotheses, it was important that the subjects whose profiles were to be used had been exposed to instruction on all the principles to be measured. In this way, any measured lack of understanding could not be attributed primarily to lack of exposure to the experiences necessary for acquiring understanding of that particular principle.

Since subjects were obtained from three elementary schools within the same Southwestern Virginia county school system, the subjects had all received instruction from the same math textbook series. In addition, instruction was directed at the same set of goals and objectives adopted by the county school system to conform with statewide Virginia Standards of Quality Goals and Objectives for Mathematics Learning. Thus considerable uniformity of instructional content was assured even though, collectively, more than two dozen teachers had delivered math instruction to subsets of the subjects during the two or three years of instruction preceding the data collection for this study.

It was also desirable to obtain a sample of subjects in which a complete range of levels of capabilities on each of the four prerequisite measures and two criterion measures was represented so as to ensure that all naturally occurring combinations of levels of performance would be observed. Pilot studies demonstrated that mid-year third graders could be expected to provide a fairly complete range of levels of performance, while end-of-the-school-year second graders skewed towards the lower levels of performance and late Spring third graders tended to perform at higher levels on the six measures.

However, testing subjects during on-going instruction with severe constraints on the amount of time the subjects could be taken from the classroom on any one day led to difficulties in controlling intervening exposure to relevant math experiences during the testing of any one subject on pilot studies. Therefore, it was decided to test some subjects who had completed second grade and some who had completed third so as to obtain the desired distribution of performance levels while testing during the summer months when there would be no intervening instruction.

All second and third graders in the three schools selected were given letters to take home explaining the research study to parents. Seventy-eight students returned slips from the letters giving written permission to participate in the study and indicating times of day and dates during the summer when participation would be most convenient. The available pool of seventy-eight subjects represented the three schools in numbers proportional to the second and third grade enrollments in the respective schools. However, the pool of prospective subjects contained approximately three students who had just finished third grade for every two students who had just finished second grade. This proportion was maintained for subjects actually tested.

Appointments for the interview testing procedure were made by telephone, scheduling children from the same school for the same session so that familiar faces would give reassurance to the young subjects. All prospective subjects (except one, with no phone) were called at least twice to make an appointment between June 28 and August 6. The first subjects contacted and available were scheduled for each interview date. Thus, all prospective subjects were given an equal opportunity to participate with availability being the only factor affecting selection.

Since examiner availability limited the total number interviewed during the six weeks of testing, in all, 51 out of the 78 prospective subjects were tested. One subject's data had to be dropped because one of the six concept measures was mistakenly discontinued before completion, and two subjects' data were dropped because the subjects had recently entered the school system from Brazil and did not have a sufficient command of English to assure a valid test of their ability to perform on all six measures. Their unfamiliarity with our money system was an additional confounding factor on the One Equal to Many measure. Altogether, 48 subjects' performances were included in the final data.

Demographically, the schools serve a population including families residing in a town of about 25,000 along with families living on farms and in rural residences near the town. The community includes a collection of small service and manufacturing industries as well as a university. Thus, a fairly broad range of socio-economic lifestyles was represented in the sample of subjects tested. The final sample of 48 subjects included 23 girls and 25 boys, with 20 of them having just finished second grade and 28 of them, third grade.

TESTING PROCEDURES

Two to six subjects were picked up at their homes for each interview session and brought to a Math-Science Laboratory classroom on the University campus. Thirty-six inch high lab counters were distributed around two sides of the classroom making it possible to place up to six interview stations between counters around the periphery while preventing any eye-contact or sound distractions between subjects. For each interview station, two 36-inch wide, flat-topped student desks were pushed together with the examiner facing the subject while the subject had his or her back to the center of the room. Testing materials were kept in a compartment under the desk top on the examiner's side of the desk, presented when needed, and then put away again so as to avoid distracting the subject from the immediate task. Fluorescent lighting was excellent in all parts of the room and no distracting sound penetrated from outside the room.

Since all of the subjects were accustomed to spending several hours per day in an instructional setting during the school year and since almost no problems with attention-span had been encountered in the pilot studies, each subject was tested on all six instruments during the same interview session. Actual testing time ranged from one and one-quarter to one and three-quarters hours, but one or two "juice and cookies" breaks were included to break up the length of the sessions. Each examiner administered either two or three instruments and then took the subject to a separate but nearby classroom for the snack. Short exploration trips in the building were available during the break, also. When all subjects had completed the initial block of tests and had had at least a ten minute break, subjects returned to a different interview station and examiner to complete a different set of tasks, until all six instruments were completed.

Using an appropriate counterbalancing procedure, the order of administration for the six measures was rotated so that each instrument was given eight times in each of the six ordinal positions of adminstration (i.e. first, second, third, etc.). Two complete sets of virtually identical testing materials were available so that each station had a complete set of the instruments to be administered there. All materials were designed to be colorful, easily manipulated, and durable. Their attractiveness to the subjects appeared partly responsible for the lack of attention-span problems, along with the novelty of the interview tasks. Protocols were used verbatim by each examiner and responses were recorded in a corresponding booklet of forms designed to facilitate complete, quick, and unambiguous recording of all task responses.

PROTOCOLS

Because of the considerable detail needed to describe precisely all the manipulative materials needed and the extensive verbatim dialogue necessary for administering the sixty test questions the complete instruments have been placed in Appendix A. Each instrument is prefaced with a general description of the testing procedure. Since the theoretical as well as empirical development and justification for each type of question selected has been included in Chapters 1, 2, and 3, the reader is referred to these sources and Appendix A for more explicit information on the measurement procedures.

THE EXAMINERS

Eight different examiners participated in the subject interviews. All but one of them were certified, practicing teachers of elementary classes or of mathematics at the middle school level. The one other examiner was an undergraduate college student who had followed the development of the protocols over several years and had a special interest in children's development of math concepts. The teachers were all satisfying requirements for a graduate school course by performing as examiners while the undergraduate student was paid for his time.

All examiners received at least ten hours of training including a demonstration of the use of all six protocols interviewing an appropriate elementary school student followed by detailed discussion of the observed demonstration. Supervised practice of the interview protocols was followed by more discussion to resolve any difficulties encountered. Supervision continued throughout the data collection phase of the research in order to resolve any unanticipated problems immediately. Since only one subject's data had to be dropped because of a procedural discrepancy, the training and supervision procedures for the examiners were deemed effective.

ANALYSIS OF THE DATA

Since the scale for the scores on the concept measurement protocols was judged to be ordinal in nature rather than equal interval, the more common parametric statistical procedures are inappropriate. For a test of the null hypotheses, the Binomial Test of proportions as discussed in Hinkle, Wiersma and Jurs (1979:183-187) was used.

In addition, since the alternative hypotheses predict an ordinal assocation between the three levels of performance measured on the four prerequisite principles and the three performance levels measured on each of the criterion principles of place value and clock reading, it was desirable to use an appropriate organization of the data so that the strength of this relationship could be assessed. The contingency table organization diagrammed in Chapter 1 was found useful for this purpose.

The three levels of performance: no understanding (N), transitional understanding (T), and competence (C) were the score classifications used to organize the data for the contingency tables. It should be noted that for each principle tested, raw scores of 0-3 were classified (N), and raw scores of 4-7 were classified (T). Raw scores of 8-10 were classified (C) if and only if a satisfactory explanation of the principle was given. For the small number of instances where raw scores of 8-10 were recorded, but a satisfactory explanation of the principle could not be elicited, a classification of (T) was given. The rationale for these performance level classifications was discussed in Chapters 1, 2, and 3.

As discussed in Chapter 1, the contingency table format provides a visual assessment of the support for the reserach hypotheses. Zeroes or very small frequencies above the diagonal in the contingency table indicate support for the expected relationship that criterion levels of performance will not exceed prerequisite levels of performance.

LIMITATIONS OF THE STUDY

The most severe limitation of this research is one which plagues all educational measurement. We wish to infer something about the competence of individuals by measuring their performance on related tasks. Every care was taken to ensure a physical and emotional climate conducive to maximal performance and the scoring procedure for indicating levels of understanding allows for a slightly less than maximal performance.

Reliability

One could argue that to define competence with less than perfect performance as a criterion is likely to impair the reliability of the instruments. However, since almost all questions are open-ended, correct guesses on as many as three or four out of five is very unlikely. By imposing the requirement of an explanation in order to be judged competent, the risk of accepting correct answers arrived at by guessing is reduced. A repeated performance reliability measure was obtained on the third pilot study of the instruments and was discussed in Chapter 3. An acceptable degree of test-retest reliability in levels of understanding was obtained for each subject.

Validity

In assessing the validity of the instruments, thoughtful consideration of the theoretical foundations for each instrument's design and of the construction of the actual interview questions was undertaken. Chapter 3 discusses the four pilot studies which together involved the testing of more than 150 children. Some prospective test items were discarded and replaced or substantially changed in order to improve both validity and reliability. In addition, modifications were made to make procedures more understandable to every subject, less time consuming to administer, or to prevent the wily subject from finding alternatives to the principle in question for solving the tasks presented.

In the third pilot study, profiles of each subject were shared with the subject's regular classroom teacher after the teacher had written predictions of probably performance levels on the measures. In all cases, actual performance levels were found to correspond well with the teachers' independent assessments of the subjects' levels of understanding. In several cases, the profiles were found to provide corroboration for weaknesses strongly suspected but unconfirmed on regular classroom work due to the many possibilities for producing answers without real understanding. After the fourth and final pilot study, the written instruments were submitted to a group of experts in appropriate areas of the field of education, for their assessments of the validity of the protocols. Collectively, the group found the face validity of the instruments to be acceptable. Chapter 3 includes a more detailed description of responses and comments from this group.

Research Design

The research design itself presents considerable limitations on conclusions that can be drawn from the results of the data analysis. As discussed by Keppel (1982:6-9) and Underwood (1957:112-125), this research design is called a "correlation study" and cause and effect conclusions are not generally warranted. The design itself simply demonstrates an association of the variables identified (or a lack of association). The only basis for a suggestion of a cause and effect relationship comes from the theoretical and empirical observations which prompted the study. That is, the plausibility of a cause and effect relationship depends on the plausibility of the hierarchical relationships between the math concepts being investigated based on logical arguments and observations of educational results from curriculum sequences aimed at teaching these concepts. Thus the discussion in Chapters 1 and 2 asserts a probable relationship between the hypothesized four prerequisite principles and the hypothesized dependent principles of Place Value and Clock Reading. If a strong directional association is demonstrated by the data with scores on Clock Reading and/or Place Value measures equal to or lower than scores on the four hypothesized prerequisite concepts, the predicted association will have been demonstrated. Moreover, if the predicted directional relationship is demonstrated for both Place Value and Clock Reading, then the predicted pattern of association of Place Value and Clock Reading principles as both being dependent upon the same four hypothesized prerequisites is substantiated, strengthening the plausibility of a cause-and-effect interpretation.

In essence, all this research study can conclude, at best, is that the predicted cause and effect relationship has been supported rather than refuted. Stronger conclusions could be drawn if follow-up studies further substantiated the hypothesized relationship. Repeated replications of the results and/or follow-up remediation and retesting of subjects with identified prerequisite concept deficits would lend considerable strength to more ambitious conclusions.

CHAPTER 5

FINDINGS AND DISCUSSION

The purpose of this study was to identify prerequisite principles for a successful outcome to instruction on the place value aspect of our number system. Four hypothesized prerequisite principles were selected for study: Addition-Subtraction, Counting By Groups, one of something equivalent to many of something else (such as one dime equal to ten pennies), and N more than M is N integers up from M in the counting sequence (and vice versa for less than). A second criterion principle--Clock Reading--was hypothesized to be dependent upon the same four prerequisites.

It was hypothesized that subjects would not score higher on measures of understanding of the two criterion principles than their lowest score across all four of the hypothesized prerequisites. Conversely, if understanding of the Clock Reading and Place Value principles is not dependent on understanding of the other four identified principles, one would expect to find no significant relationship between performance on the criterion and performance on prerequisite measures across subjects.

DATA OBTAINED

Performance levels on each of the four hypothesized prerequisites and the two hypothesized criterion measures were obtained for the 48 subjects described earlier. Table 3 contains the master data chart displaying the raw scores (0 to 10 plus an E if the explanation was satisfactory) for each of the six principles tested, as well as a

Table 3

Raw Scores Associated Performance Level Classifications[a]

Rank	Subject Number	1982 Concept Test Response Scores and Concept Ratings: Add.Subt.		Cnt. Grps.		One=Many		± N		Plc.Value		Clk. Rdg.		Total Score
1	2304	10+E	C	10+E	C	10+E	C	10+E	C	10+E	C	10+E	C	60+6E
2	1308	10+E	C	10+E	C	10+E	C	10+E	C	7+E	T	9+E	C	56+6E
3	2306	10+E	C	9+E	C	10+E	C	10+E	C	9	T	8	T	56+4E
4	1359	10+E	C	10+E	C	10+E	C	10+E	C	9+E	C	6+E	T	55+6E
5	3353	8+E	C	9+E	C	10+E	C	10+E	C	9+E	C	7+E	T	53+6E
6	1351	10+E	C	10+E	C	9+E	C	9+E	C	8	T	7	T	53+4E
7	1354	9+E	C	10+E	C	9+E	C	10+E	C	8+E	C	6	T	52+5E
8	1310	10+E	C	9+E	C	10+E	C	10+E	C	7	T	6	T	52+4E
9	2355	8	T	10+E	C	10+E	C	9+E	C	8+E	C	6+E	T	51+5E
10	1257	10+E	C	7	T	9+E	C	5+E	T	8	T	10+E	C	49+4E
11	1205	10+E	C	6	T	9+E	C	10+E	C	8	T	5+E	T	48+4E
12	1300	8+E	C	10+E	C	9+E	C	7+E	T	7	T	6+E	T	47+5E
13	2302	9+E	C	5	T	9+E	C	10+E	C	9+E	C	5	T	47+4E
14	3352	10+E	C	9+E	C	9+E	C	9+E	C	4	T	5	T	46+4E
15	1956	9+E	C	5	T	9+E	C	10+E	C	7+E	T	5	T	45+4E
15	1311	9+E	C	9+E	C	10+E	C	10+E	C	3	N	4	T	45+4E
17	1355	10+E	C	9+E	C	10+E	C	4	T	7	T	5	T	45+3E
17	1253	8+E	C	10+E	C	10+E	C	6	T	8	T	3	N	45+3E
19	2308	9+E	C	10+E	C	9+E	C	4	T	6	T	5+E	T	43+4E
20	2305	9+E	C	10+E	C	9+E	C	5	T	5	T	5	T	43+3E
21	1258	10+E	C	5	T	10+E	C	10+E	C	3	N	4	T	42+3E
21	2256	5	T	5	T	10+E	C	10+E	C	5	T	7+E	T	42+3E
23	1252	8+E	C	9+E	C	10+E	C	5	T	6	T	3	N	41+3E
24	2307	5	T	6	T	10+E	C	10+E	C	6	T	4	T	41+2E
25	2357	10+E	C	9+E	C	4	T	10+E	C	5	T	2	N	40+3E
25	1207	8+E	C	6	T	9+E	C	8+E	C	5	T	4	T	40+3E
27	3204	6+E	T	10+E	C	10+E	C	6	T	2	N	5	T	39+3E
28	3203	8	T	5	T	10+E	C	10+E	C	4	T	2	N	39+2E
29	2251	5	T	5	T	10+E	C	7	T	7	T	5	T	39+E
30	1262	5	T	8	T	7	T	10+E	C	5	T	3	N	38+E
31	1303	6+E	T	5	T	9+E	C	5	T	6	T	6+E	T	37+3E
31	1260	5	T	4	T	9+E	C	8+E	C	5	T	6+E	T	37+3E
33	1312	7	T	5	T	4	T	10+E	C	8	T	3	N	37+E
33	2301	8+E	C	6	T	7	T	7	T	8	T	1	N	37+E
35	2258	7	T	8+E	C	5	T	10+E	C	5	T	1	N	36+2E
36	1361	3	N	3	N	10+E	C	10+E	C	3	N	6+E	T	35+3E
37	1304	8+E	C	7	T	10+E	C	2	N	3	N	5	T	35+2E
37	2360	3	N	0	N	9+E	C	10+E	C	8	T	5	T	35+2E
39	3251	1	N	6	T	5	T	10+E	C	6	T	6	T	34+E
40	2259	7	T	4	T	9+E	C	7+E	T	5	T	0	N	32+2E
40	2253	5	T	7+E	T	7+E	T	4	T	2	N	7	T	32+2E
42	2252	5	T	4	T	9+E	C	7	T	3	N	3	N	31+E
43	1313	6	T	5	T	10+E	C	1	N	4	T	3	N	29+E
44	3302	3	N	0	N	9+E	C	9+E	C	2	N	4	T	27+2E
45	2354	4	T	5	T	5	T	0	N	7	T	5	T	26
45	2203	6	T	5	T	6	T	1	N	4	T	4	T	26
45	3201	6	T	5	T	5	T	7	T	3	N	0	N	26
48	1206	3	N	3	N	1	N	0	N	3	N	0	N	10

[a]Performance classifications are N: No understanding, T: Transitional understanding, and C: Competence. The E after raw scores indicates that a satisfactory explanation was given.

performance level rating (N: No understanding, T: Transitional or initial understanding, and C: Competence) for each of the six principles tested. The total scores listed were used for comparisons of individual students' performances to the overall group, (a) for the use of parents, and (b) for the use of teachers. A maximum score of 60 (10 for each concept) plus 6E (one explanation for each concept) could be obtained.

The prerequisite-criterion performance data are summarized in the contingency tables as shown in Table 4 for the Place Value criterion measure and shown in Table 5 for the Clock Reading criterion measure. The topmost contingency table on each page compares the lowest performance across all four of the prerequisite measures to the criterion measure. Also shown are the contingency tables comparing performance on each of the four prerequisite measures with performance on the criterion measure.

It was explained earlier that if the expected predictor-criterion relationship was supported perfectly by the research data, the three cells above the diagonal of each contingency table would contain only zeroes. This would indicate that criterion levels of performance did not exceed prerequisite levels of performance for any subjects. In fact, the One Equal to Many prerequisite measure displays exactly this expected pattern of zeroes on both the Place Value and Clock Reading criterion measures as shown in the lower left contingency crosstabulations of Tables 4 and 5.

Table 4

Crosstabulation of Performance Levels Between the Prerequisite Measure and the Criterion Measure of Place Value

N[a]		Place Value N[a]	T	C
Lowest	N	4	5	0
Prerequisite	T	5	22	2
Score[b]	C	1	5	4

		Place Value N	T	C
Addition	N	3	2	0
Subtraction	T	4	13	1
Score	C	3	17	5

		Place Value N	T	C
Counting	N	3	1	0
By Groups	T	5	18	1
Score	C	2	13	5

		Place Value N	T	C
One Equal	N	1	0	0
to Many	T	2	8	0
Score	C	7	24	6

		Place Value N	T	C
Plus or	N	2	3	0
Minus N	T	4	11	0
Score	C	4	18	6

[a]The levels of performance are designated by N: No understanding, T: Transitional understanding, C: Competence.

[b]The lowest score of the four scores obtained for Addition-Subtraction, Counting By Groups, One Equal to Many, and Plus or Minus N.

Table 5

Crosstabulation of Performance Between the Prerequisite Measure and the Criterion Measure of Clock Reading

		Clock Reading		
		N[a]	T	C
Lowest	N	2	7	0
Prerequisite	T	11	17	1
Score[b]	C	0	8	2

		Clock Reading		
		N	T	C
Addition -	N	1	4	0
Subtraction	T	8	10	0
Score	C	4	18	3

		Clock Reading		
		N	T	C
Counting	N	1	3	0
By Groups	T	8	15	1
Score	C	4	14	2

		Clock Reading		
		N	T	C
One Equal	N	1	0	0
to Many	T	6	4	0
Score	C	6	28	3

		Clock Reading		
		N	T	C
Plus or	N	2	3	0
Minus N	T	6	8	1
Score	C	5	21	2

[a]The levels of performance are designated by N: No understanding, T: Transitional understanding, C: Competence.

[b]The lowest score of the four scores obtained for Addition-Subtraction, Counting By Groups, One Equal to Many, and Plus or Minus N.

Further perusal of the cells above the diagonal in each contingency table yields some additional observations. First, looking at just the four lower contingency tables for the Place Value criterion measure, it can be seen that no more than three out of the forty-eight subjects recorded a Place Value level of understanding higher than the level for any one prerequisite measure. Similarly, for the Clock Reading criterion measure, exactly four subjects scored criterion levels of understanding which exceeded the prerequisite level of understanding on each of the three prerequisite measures other than the One Equal to Many measure. Looking at the cells below the diagonals associated with the four hypothesized prerequisite measures, it is apparent that many more subjects (ranging from twenty to forty) demonstrated prerequisite performance levels higher than the level of performance on the criterion measures. Thus a strong ordinal relationship between levels of understanding on prerequisite and criterion measures is demonstrated, with prerequisite performance levels in most cases equaling or exceeding criterion levels.

Looking at the topmost contingency crosstabulations in Tables 4 and 5, it appears that the relationships between the lowest level of understanding on the prereequisite measures and levels on the criterion measures are subtly different although ordinally the same as the previously discussed relationships. It is also apparent that the number of subjects having criterion score levels exceeding their prerequisite score levels is greater (seven for Place Value and eight for Clock Reading) when the lowest prerequisite score is used. This is because the "worst cases" from the single prerequisite comparisons are used in this comparison.

By contrast, looking at the cells below the diagonal in these two contingency tables, it is evident that the frequencies for lowest prerequisite levels of understanding exceeding criterion levels of understanding (eleven for Place Value and nineteen for Clock Reading) are lower than any of the comparable frequencies for the single prerequisite comparisons with the same criterion. Combining these two observations, and totaling the frequencies on the diagonals for confirmaton, it is evident that the lowest prerequisite score predictor measure has the highest frequency of equal levels of performance on prerequisite and criterion measures when the relationship involves the lowest performance level across all four prerequisite predictors for the same criterion measure. Thus the hypothesis that the lowest level of understanding on any of the four prerequisite measures would be a good predictor of children's levels of success in understanding the Place Value and Clock Reading principles appears to be supported by the data, in that these contingency tables contain far more entries on the diagonal than for any of the four prerequisite measures considered separately.

TESTS OF HYPOTHESES

The primary and secondary hypotheses stated that the cells above the diagonal in the Place Value and Clock Reading contingency tables just discussed would all contain zeroes, but this is the case only for the One Equal to Many measure. Although it would seem that one could test whether the observed frequencies differed significantly from zero, no such test could be identified. Therefore it was decided to determine whether the sums of the frequencies in the three cells above the diagonals differ significantly from sums of the same cell frequencies expected by chance if no relationship exists between prerequisite and crtierion measures. A Binomial Test was employed for this purpose.

Before a Binomial Test could be executed, it was necessary to determine the expected proportion of higher scores that would occur if there was no relationship between the prerequisite tasks and the criterion tasks. A two-step computation was necessary. First, the expected frequencies, assuming no relationship, for each of the three cells above the diagonal were obtained by multiplying the two marginal totals for each cell and dividing by the total number of subjects (N = 48). Then these three expected cell frequencies were summed and divided by N (48) again to obtain the proportion of examinee's scores expected for the criterion measure higher than the prerequisite measure assuming no relationship between criterion and prerequisite measures. These proportions were computed for all contingency crosstabulations in Tables 4 and 5. The expected and observed proportions are reported in Table 6. The expected proportions are shown in the first and fourth columns and the corresponding observed proportions are shown in the second and fifth columns of Table 6. The observed frequencies displayed in the cells above the diagonal in each crosstabulation shown in Tables 4 and 5 were summed and divided by the total N (48) to convert the summed frequencies to the proportions needed.

The Binomial Test was performed using the normal curve approximation to the Binomial Distribution for $N > 20$ and using the corresponding expected and obtained proportions shown in Table 6. The resulting

Table 6

Binomial Test of the Difference Between Expected and Observed Proportions of Subjects Whose Performance Level on the Criterion Exceeds their Performance Level on Each of the Prerequisite Measures.

Prerequisite Measure	Criterion Measure: Place Value			Clock Reading		
	Expected Proportions	Observed Proportions	ω_p	Expected Proportions	Observed Proportions	ω_p
Addition Subtraction	.129	.063	-3.63[a]	.099	.083	-2.40[a]
Counting By Groups	.129	.042	-4.43[a]	.092	.083	-2.22[a]
One Equal to Many	.043	0.0	---[a,b]	.028	0.0	---[a,b]
Plus or Minus N	.122	.063	-3.43[a]	.095	.083	-2.30[a]
Lowest Prerequisite Performance	.223	.146	-3.61[a]	.174	.167	-3.07[a]

[a]Significant to the .01 probability level ($\alpha_{.01,1\text{-tailed}} = \pm 1.288$).

[b]Test statistic is not calculable because the observed score proportion equals zero.

statistics are displayed in the third and sixth columns of Table 6. This Binomial Test on the proportion of criterion measure scores higher than the prerequisite measure scores tests the null hypothesis that the proportion of children scoring a higher performance level on the criterion measure than their level on the prerequisite measure will equal the proportion expected by chance if there is no relationship. The alternative hypothesis against which the null is tested is that the criterion measure scores will not be higher than the predictor scores for the same subject. Since the alternative hypothesis is directional, a one-tailed test of the Binomial Test statistics was performed for α = .05.

It was found that the null hypotheses for all predictor-criterion relationships tested were rejected with the test statistics all being significant at the .01 probability level, as shown in Table 6. Thus the primary hypothesis that children's performance levels for the Place Value measure would not exceed their performance levels on any of the prerequisite measures is supported. Likewise, the secondary hypothesis that children's scored levels of understanding for the Clock Reading measure would not exceed their scored levels on any of the prerequisite measures is also supported.

DISCUSSION OF RESULTS

The hypothesized relationship between prerequisite principles and the Place Value and Clock Reading principles was supported by the research findings with refernce to each of the four prerequisite principles. The hypothesized relationship was supported, also, for the lowest performance observed across all four measures. Therefore, there is support for the assertion that the Addition-Subtraction, One Equal to Many, Counting By Groups, and Plus or Minus N principles are all prerequisites to a functional understanding of the Place Value and Clock Reading principles.

While the organization of the data in Tables 4 and 5 provides for some comparison fo the relative importance of each of the four prerequisite principles to the understanding of the criterion principles, there is another way to look at these predictor-criterion relationships. Table 7 indicates the number of subjects for whom performance on each of the four prerequisite measures was the lowest level of performance attained on the four measures. Figures are shown for each prerequisite measure giving the number of subjects for whom each prerequisite was: the unique lowest score, one of two equal lowest scores, one of three equal lowest scores, and one of four equal lowest scores. The number in parentheses indicates the number of subjects whose lowest score was a "false predictor" meaning that performance on the prerequisite was lower than the Place value criterion performance level in contradiction to the primary hypothesis. That is, the seven instances where the lowest prerequisite performance level was actually lower than the performance level on the Place Value measure are indicated in the parentheses. Likewise, the numbers in brackets indicate the "false predictors" for the Clock Reading criterion measure.

Table 7

Number of Prerequisite Principle Scores Equal to the Lowest Score Measured on All Four Prerequisties (N = 48)

Uniqueness of Predictor Scores	Prerequisite Principles: Addition Subtraction Concept	Counting By Groups Concept	One Equal to Many Concept	Plus or Minus N Concept
Unique Lowest	2 (2)[a] [1][b]	5 (1)	1	10 (3) [3]
One of Two Lowest	9 (1) [2]	8 (1) [3]	1	2 [1]
One of Three Lowest	5	6	3	4
Four Prerequisites All Equal	14	14	14	14
Total	30	33	19	30

[a]Entries in parentheses indicate the number of subjects for whom performance on the crtierion measure of Place Value exceeded performance on the indicated prerequisite measure in contradiction to the primary hypothesis.

[b]Entries in brackets indicate the number of subjects for whom performance on the criterion measure of Clock Reading exceeded performance on the indicated prerequisite measure in contradiction to the secondary hypothesis.

This procedure for analyzing the results shows clearly that eliminating the One Equal to Many measure would result in only one more subject violating the hypothesized prerequisite-criterion relationship. That is, in this one case, the subject achieved competence level scores on all other prerequisite principle measures but achieved only a transitional level of performance on the One Equal to Many and Place Value measures, and incidentally a no understanding level on the Clock Reading measure. This, of course, does not suggest that children need not master the One Equal to Many principle before mastering Place Value understanding, but only that, for this sample of second and third graders it tended to be mastered earlier or at the same time as the other three prerquisite principles. This is corroborated by the data in Table 4 showing that 37 out of the 48 subjects achieved the competence level on the One Equal to Many measure.

It is also interesting to note that the Plus or Minus N measure and the Counting By Groups measure accounted for all the other uniquely lowest prerequisite performance levels. However, a sizable number of subjects produced "one-of-their-two-lowest" prerequisite performances on the Addition-Subtraction measure suggesting that understanding of this principle is also possibly important to success with Place Value understanding, but not essential as a predictor since the Counting By Groups and Plus or Minus N predictors appear sufficient. Unfortunately, these two best prerequisite measures also account for 5 out of the 7 lowest scores that were lower than the same subject's performance on the Place Value measure, thus appearing to give some false predictions of lack of readiness. Nonetheless, the Counting by Groups measure alone gave the lowest prerequisite score (eliminating the two false lowest instances) for 31 out of the 41 instances where performance was predicted while the Plus or Minus N measure provided at least 7 more unique predictions consistent with the primary hypothesis.

It is important to consider the instances of "false predictions." These seven "false predictions" for the Place Value measure and eight "false predictions" for the Clock Reading measure were brought to attention earlier in the contingency table data. Focusing on the lowest prerequisite performance levels relative to the Place Value performance, the sum of the cell frequencies above the diagonal was seven. According to the primary hypothesis these frequencies should have been zero. The eight "false predictions" for the Clock Reading measure were shown similarly in Table 5.

Considering the strength of the support for the primary and secondary hypotheses, a likely explanation for these "false prediction" instances is that they result from measurement error. In fact, careful analysis of administration procedures and examiner comments suggests three different sources of measurement error. Four of the "false prediction" instances can be explained as a "warm-up" effect in that the anomalously low scores were produced on the first instrument administered to these four subjects. It could be that these subjects took some time to adjust to the novel research situation and that nervousness about the unfamiliar situation affected their performance negatively. Two of the "warm-up" instances involved the Addition-Subtraction instrument, one instance involved the Counting By Groups instrument, and one the Plus or Minus N instrument. So, it would seem that the "warm-up" effect was a generalized phenomenon, rather than specific to one instrument, the commonality being that they were the first tasks administered to these four subjects.

Since the Addition-Subtraction instrument provided one instance not ascribable to a "warm-up" effect and the Plus or Minus N instrument provided four such instances, other reliability factors must have affected the administration of these instruments. While examiner comments suggest that a particularly active and distractable subject was a factor in the other Addition-Subtraction anomaly, the Plus or Minus N instrument clearly posed a problem discussed earlier in Chapter 3. Examiner comments indicated that the subjects who produced the anomalously low scores on the Plus or Minus N instruments, all had considerable trouble in understanding the instructions at the beginning of the task. The unfamiliarity of this task was probably also a factor contributing to this problem, since this principle is not explicitly taught in most elementary mathematics curriculums. For all of these reasons, it would appear that measurement error could be a contributing factor for all the "false prediction" instances recorded in this research study.

Another problematical aspect of the results, not discussed earlier, is the especially low overall performance levels recorded on the Clock Reading measure. Only three subjects achieved the competence level on this measure, while thirteen subjects did not even achieve the transitional level of understanding on this measure. Furthermore, ninteen subjects scored lower on the Clock Reading measure than their lowest performance level on the four pererequisite measures. These statistics could be an indicator that some other prerequisite principle, not investigated in this study is also necessary to successful performance on the Clock Reading measure. This possibility should be considered further in the next chapter.

SUMMARY

The statistical tests of the results obtained in this study show that the null hypotheses with respect to both criterion variables were rejected at the .01 confidence level, thus lending support to the hypotheses that children would not score higher levels of understanding on the criterion measures of Place Value and Clock Reading understanding than their lowest score on the four hypothesized prerequisite principles. After analyzing administration order and examiner observations, it was found that a possible explanation for the few cases where performance was inconsistent with prediction is measurement error. The high incidence of scores indicating lower levels of understanding on the Clock Reading measure than the lowest level of understanding on the prerequisite measures suggests that other possibilities should be considered in assessing readiness for learning the Clock Reading principle. In addition, it was found that the Counting By Groups and Plus or Minus N measures were the best predictors of levels of success on both the Place Value and Clock Reading criterion measures.

Chapter 6

INTERPRETATION AND CONCLUSIONS

The purpose of this study was to identify basic numerical principles prerequisite to successful acquisition of a functional understanding of the place value principle. It was hypothesized that mastery of analog clock reading is dependent upon the same set of prerequisite principles. Therefore, a study was designed in which children's levels of understanding were measured on four hypothesized prerequisite principles and the two criterion principles of Place Value and Clock Reading.

The primary hypothesis was that if the four proposed prerequisite principles were truly prerequisites for a functional understanding of place value, then children would not score higher on the Place Value criterion measure than their lowest score on any of the prerequisite measures. The prediction for the secondary hypothesis was that if the four prerequisite principles selected were necessary to a functional understanding of the selected Clock Reading principle, then children would not score higher on the Clock Reading criterion principle than their lowest level scored on any of the four prerequisite measures.

The null hypotheses for both the primary hypothesis and secondary hypothesis were rejected at the .01 probability level. It was concluded that the four foundation principles tested are prerequisites to successful learning of the place value principle. However, some further consideration of the research design and its limitations is in order.

Since the research design was a correlational study, the results showed only that performance levels on the criterion measures were generally less than or equal to performance on the prerequisite measures as predicted. Also, the level of understanding of the Place Value principle tended to correspond with the lowest level of understanding demonstrated on any one of the four prerequisite principles. While this supports the predicted association, it does not establish a cause-and-effect relationship. Nevertheless, the established precedence of prerequisite principle acquisition coupled with the previous research which supports this relationship lends additional support to the likelihood that a cause-and-effect relationship does exist.

Additional support is also provided by the many studies cited which have established that instruction improving performance on readiness factors also tends to improve performance on the dependent measure, as well. If a follow-up study were conducted in which remedial instruction on the lowest scoring foundation principles was used as a treatment and the treatment was shown to improve performance on a measure of place value understanding, this remediation study result would indicate that the relationship herein identified is indeed a cause-and-effect relationship. Therefore, a remediation study would appear to be a very desirable follow-up to the results of this study.

Another limitation of this study, derives from the fact that all data for the pilot studies and the final study were obtained from interviews with children attending only two different county school systems within the same state. It is entirely possible that the relative timing for acquisition of understanding of the foundation principles is influenced by the order of presentation and relative amounts of teaching time allocated. These curriculum emphasis factors will vary from one school system and from one state to another although the widespread use of only a few different mathematics textbook series as well as nationally standardized achievement tests to monitor pupils' progress does provide for some uniformity of curriculum emphasis.

Since the One Equal to Many principle provided only one unique lowest score on the four prerequisite principles, deleting this measure would appear to be immaterial to the results of this study. However,at least four factors argue against deletion without further study. One factor is the limited sample tested, and a second is the curriculum emphasis factor just discussed. Another factor was pointed out by one of the experts providing face validity judgements of the instruments as discussed in Chapter 3. That is, that the One Equal to Many instrument tested only a limited generalization of the principle since it used only standard money equivalences to test for understanding of the principle. It is possible that a procedure testing for a more global generalization of the principle would not indicate such an early acquisition for most subjects. A fourth factor to be considered is the extensive support for inclusion of this principle in the set of prerequisites provided by results of research studies and recommendations of mathematics educators discussed in Chapter 2. For all of these reasons, further research would seem prudent before dismissing the One Equal to Many principle from the set of probable prerequisites to Place Value and Clock Reading understanding.

CONSIDERATIONS PERTAINING TO CLOCK READING

Consideration of observations made by examiners after many administrations of the Clock Reading instrument suggest at least two possibilities to explain the unexpectedly low levels of understanding obtained on the Clock Reading measure. The first observation was that many children exhibited quite noticeable difficulty with the process of integrating all the "rules" for the different elements of the clock reading process. These subjects seemed to be able to use and even verbalize rules for most, if not all, elements of the process but had difficulty applying them all simultaneously. This difficulty with the synthesis of all of the elements is most likely to be a developmental aspect of the learning process rather than a missing prerequisite, per se. Piagetian learning theory (see Chapter 2) suggests that the synthesis process requires some time and perhaps additional experience before competence with the synthesized principle is acquired.

This time delay limitation could account for the unexpectedly low levels of understanding obtained on the Clock Reading measure. However, a second possible explanation was derived from examiners' observations. A number of subjects in each pilot study and in the final study demonstrated a type of confusion associated with a variety of measurement tasks in which an equal interval scale must be read. Elementary pupils often exhibit the problem with number line activities or linear measurement activities where a ruler is used. Confusion lies with whether to count intervals or the division lines, including whether to count the zero point, and with whether to include the intervals or lines where the numerals are located or just those between the numerals. Counting the zero point yields a count consistently one higher than the correct count, while skipping the intervals or division lines where the numerals are located yields only four minutes for every full five minute block included in the count. Some children combined both types of errors in their performances as well as combining them with counting-by-fives strategies to compound the confusion.

It would appear, then, that understanding of how to use an interval scale properly in measurement tasks might be an untested additional prerequisite to successful learning of the Clock Reading principle. Whether it is truly a separate mathematics principle or an application of ordination-cardination synthesis (the Plus or Minus N principle) is debatable. Whether to count the zero point was a noticeable problem for some subjects on the Plus or Minus N tasks. Thus a new instrument could be added or a modification to substitute some interval scale application tasks for part of the present N More or Less instrument could be employed. Either way, some investigation of this interval scale reading problem seems appropriate to any further study of prerequisites to the Clock Reading principle.

The second difference in interpretation of the Clock Reading and Place Value results has to do with the issue of curriculum emphasis. While emphasis on place value understanding is pervasive in any second or third grade mathematics curriculum, emphasis on clock reading is considerably more variable, especially in amount of time devoted to instruction in the classroom. Furthermore, the pressure to acquire understanding of the principles in order to cope successfully with societal demands in the children's environments is at least as variable. Place value understanding is still a pervasive need encompassing handling of money, and use of calculators among others. In contrast, with the popularization of digital time pieces, many children manage to avoid having to cope with the task of reading an analog clock almost completely. Thus, curriculum and societal influences may be other factors contributing to the unexpectedly low scores on the Clock Reading measure, as well.

SUMMARY OF CONCLUSIONS

While the results statistically support the two hypotheses of this study, several reservations to the most straightforward conclusions must be offered. Although subjects included in this study tended to exhibit no higher level of understanding on the Place Value and Clock Reading measures than the lowest level of understanding demonstrated on the four prerequisite principle measures, a conclusion that these prerequisite principles are therefore necessary to a functional understanding of place value and clock reading is subject to several limitations or reservations.

The fact that the research design was a correlation study makes cause-and-effect conclusions somewhat risky although the support provided by other hierarchical studies on place value coupled with the strong directional relationships exhibited by the present results tend to compensate somewhat for the limitations inherent in the research design. Nevertheless this design limitation coupled with the small size (N = 48) and limited representativeness of the subject sample tested dictate that both remediation and replication studies would be desirable to confirm the probable cause-and-effect relationship implied by these results.

The high incidence of responses indicating lower levels of understanding on the Clock Reading measure than the lowest level scored on the prerequisite measures merits further investigation. This pattern of responses could have been produced because after learning, time is required for the process of integration and synthesis of new understandings before application is possible. However the possible importance of understanding the use of equal interval measurement scales before attempting understanding of analog clock reading should be investigated.

Nevertheless, the probable prerequisite status of the four principles identified in this study could be considered in classroom planning whether or not further research studies are attempted. Since previous research and curriculum designers tend to support the prerequisite status strongly implied by the results of this study, consideration of these probable prerequisites in diagnosis and remediation planning for instruction concerning place value and clock reading understanding would appear to be prudent at this time. Certainly, for children experiencing difficulty in mastering the Place Value and Clock Reading principles, assuring that they have mastered the four prerequisite principles identified in this study as a course of remediation is not likely to cause any harm. Furthermore, the findings of this study suggest that such a course of remediation might indeed be very effective for some children.

Chapter 7

SUMMARY

The motivation for this research effort was to find some remedy to the frustration encountered by many teachers and pupils when understanding of the place value aspect of our numeration system becomes important to successful learning in the elementary school mathematics curriculum. It was hypothesized that children who continue to fail to grasp the essential principle of place value after several attempts at reteaching lack of understanding of some prerequisite principle or principles. It was felt that identification of such prerequisties would make it possible for teachers to identify the missing principle(s) and then proceed successfully to teach the place value principle, so essential to continuing success in the mathematics curriculum. Four plausible prerequisite principles were identified and instruments were developed to measure levels of understanding for each principle.

In all, six interview protocol instruments were developed. The four prerequisite principle instruments are the Addition-Subtraction, Counting By Groups, One Equal to Many, and N More or Less instruments. In addition, instruments were developed to measure levels of understanding for the two criterion principles, the Place Value and the Clock Reading instruments, since both of these principles were hypothesized to be dependent on the same four prerequisites. All six instruments were individual interview protocols giving verbatim instructions and questions for ten task items and an explanation item for each instrument. All items included the use of manipulative or representational materials to aid in communication between examiner and subject. Four successive pilot studies were conducted to develop, revise and validate these protocols before the final data were obtained.

Scoring of responses was based upon the three levels of understanding favored by Piaget (1952) to measure the stages in development of the understanding of a logico-mathematical operation. Item scores of 0-3 correct responses were assigned the no understanding level (N), item scores of 4-7 were assigned the transitional level of understanding (T), and item scores of 8-10 including a satisfactory explanation were assigned the competence level (C). This system of designating three levels of understanding was used throughout the analysis of the results.

The tested hypotheses were that children would not achieve a higher level of understanding on the measure of place value understanding (and likewise for clock reading understanding) than their lowest level of understanding achieved on the four proposed prerequisite measures.

Forty-eight children, twenty who had just completed second grade and twenty-eight who had just finished third grade, were tested on all six instruments during the Summer of 1982. Both the Place Value and Clock Reading hypotheses were tested statistically using the Binomial Test of Proportions. For both hypotheses, the proportions of subjects demonstrating higher levels of understanding on the place value and clock reading measures than the lowest level of understanding exhibited on the foundation principle measures were tested against the corresponding expected chance proportions assuming no relationship between the criterion and foundation principles. The null hypotheses for both the Place Value and Clock Reading hypotheses were rejected at the .01

probability level. Thus the four proposed prerequisite principles were supported to be necessary to a functional understanding of place value.

Further analysis of the results and examiner observations led to the following conclusions. The same statistical results would be obtained for the reduced set of three foundation principles (Addition-Subtraction, Counting By Groups, and N More or Less) if the One Equal to Many principle was deleted from the set of foundation principles. However, findings from research reviewed coupled with limitations affecting this reserach study argue for further consideration before discarding the One Equal to Many principle as a prerequisite to place value and clock reading understanding.

Next, examiner observations suggested two possibilities for additional prerequisites to a functional understanding of the Clock Reading principle. One possibility was that synthesis of the understandings of the four principles requires considerable time and perhaps additional experience or practice before a higher level of understanding on the Clock Reading principle could be achieved. Furthermore a problem with the process of integrating all the elements of clock reading to apply them simultaneously was also observed so that the synthesis or integration process itself may be a missing prerequisite. Another possibility suggested was that an additional prerequisite principle encompassing the understanding of interval scale reading might indeed be necessary to the understanding of clock reading.

Finally, it was concluded that follow-up research studies of two types appear desirable. One type would be remediation studies to test the probable cause-and-effect relationship between the foundation principles and the criterion principles. The second type would be replication studies to determine if the results can be replicated in other school systems with differing curriculum emphases. Implications for instructional planning were also discussed.

LITERATURE CITED

Almy, Millie, and Associates
1970. Logical Thinking in Second Grade. New York: Teachers College Press.

Almy, Millie, E. Chittenden, and P. Miller
1966. Young Children's Thinking. New York: Teacher College Press.

Ashlock, Robert B.
1972. Error Patterns in Computation. Columbus, Ohio: Charles E. Merrill Books, Inc.

Baker, Beverly Elizabeth
1977. "The Effects of Sequencing Levels of Representation on the Learning and Retention of Place Value Skills." Dissertation Abstracts (University of South Florida, 1977) 38:7076.

Barr, David Campbell
1976. "A Comparison of Three Methods of Introducing Two-Digit Numeration and the Effects of these Methods on Certain Groups of Children." Dissertation Abstracts (University of Illinois at Urbana-Champaign, 1975) 36:5799.

Bednarz, Nadine and Bernadette Janvier
1982. "The Understanding of Numeration in Primary School." Educational Studies in Mathematics, 13(1):33-57.

Bidwell, James K.
1969. "Learning Structures for Arithmetic." The Arithmetic Teacher, 16:263-268.

Biskin, Donald, Wayne Ellis, Drusilla Gotow, and Dianna Richardson
1975. "An Empirical Test of the Necessity of the Justification Criteria for Assessing the Initial Occurrence of Conservation." Paper presented at the Annual Meeting of the American Educational Research Association, Washington, D.C.

Brownell, William A.
1970. "Arithmetical Readiness as a Practical Classroom Concept." Current Research in Elementary School Mathematics, eds. R. B. Ashlock and W. L. Herman, Jr., New York: The Macmillan Company.

Brownell, William A.
1967. Arithmetical Abstractions: The Movement toward Conceptual Maturity under Differing Systems of Instruction. California: University of California Press.

Callahan, Leroy G. and Vincent J. Glennon
1975. Elementary School Mathematics: A Guide to Current Research, 4th ed. Washington, D.C.: Association for Supervision and Curriculum Development.

Chissom, Brad S., Delores G. Collins, and Jerry R. Thomas
1974. "Relationships Among Perceptual-Motor Measures and their Correlations with Academic Readiness for Preschool Children." Perceptual and Motor Skills, 39:467-473.

Copeland, Richard W.
1974. How Children Learn Mathematics: Teaching Implications of Piaget's Research, 2nd ed. New York: Macmillan Publishing Co., Inc.

Cox, L. S.
1975. "Systematic Errors in the Four Vertical Algorithms in Normal and Handicapped Populations." Journal for Research in Mathematics Education, 6:202-220.

Cunningham, Joseph G. and Richard D. Odom
1978. "The Role of Perceptual Salience in the Development of Analysis and Synthesis Processes." Child Development, 49:815-823.

Daugherity, Walter Chisholm
1978. "The Ordering of Topics in Teaching Mathematics." Dissertation Abstracts (Harvard University, 1977) 39:3431.

De Flandre, Charles, Jr.
1974. "The Development of a Unit on Place-Value Numeration Systems, Grades Two, Three, and Four." Dissertation Abstracts (Temple University, 1974) 35:6434.

Dienes, Zoltan P.
1969. Mathematics in the Primary School. London: MacMillan and Co. Ltd.

Dodwell, P.C.
1968. "Relation Between the Understanding of the Logic of Classes and of Cardinal Number in Children." Logical Thinking in Children, eds. Irving E. Sigel and Frank H. Hooper. New York: Holt, Rinehart and Winston, Inc.

Dodwell, P.C.
1960. "Children's Understanding of Number and Related Concepts." Canadian Journal of Psychology, 14:191-205.

Duncan, Ruth Kruger
1980. "The Effects of Two Instructional Treatments on Increasing Computational Skill in Subtraction with Regrouping of Grade Three Students." Dissertation Abstracts (Indiana State University, 1980) 41:3876.

Edge, Douglas Richard Montgomery
1980. "The Effects of Using Multiple Embodiments of Selected Multi-Digit Numeration Concepts on the Understanding and Transferability of Those Concepts." Dissertation Abstracts (University of Maryland, 1980) 41:1990.

Elkind, David
1974. Children and Adolescents: Interpretive Essays on Jean Piaget, 2nd ed. New York: Oxford University Press.

Ellis, Leslie Clyde
1972. "A Diagnostic Study of Whole Number Computations of Certain Elementary Students." Dissertation Abstracts International (The Louisiana State University and Agricultural and Mechanical College, 1972) 33:2234.

Feigenbaum, Kenneth D.
1963. "Task Complexity and I.Q. as Variables in Piaget's Problem of Conservation." Child Development, 34:423-432.

Flournoy, Frances, Dorothy Brandt, and Johnnie McGregor
1963. "Pupil Understanding of the Numeration System." The Arithmetic Teacher, 10:88-92.

Frege, Gottlob
1884. The Foundations of Arithmetic: a Logico Mathematical Enquiry into the Concept of Number. Eng. trans. by J. L. Austin, 2nd rev. ed. Evanston: Northwestern University Press, 1953 ed.

Gagné, Robert M.
1975. Essentials of Learning for Instruction. Hinsdale, Illinois: The Dryden Press.

Gagné, Robert M.
1970. "Learning and Proficiency in Mathematics." Current Research in Elementary School Mathematics, eds. R. B. Ashlock and W. L. Herman, Jr., New York: The Macmillan Company.

Gagné, Robert M.
1965. The Conditions of Learning. New York: Holt, Rinehart and Winston.

Ginsburg, Herbert
1977. Children's Arithmetic: The Learning Process. New York: D. Van Nostrand Company.

Ginsburg, Herbert and Sylvia Opper
1969. Piaget's Theory of Intellectual Development. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Grossman, Anne S.
1983 "Decimal Notation: An Important Research Finding." Arithmetic Teacher, 30(9):32-33.

Grouws, Douglas A.
1974. "Solution Methods Used in Solving Addition and Subtraction Open Sentences." Arithmetic Teacher, 21:255-261.

Grouws, Douglas A.
1972. "Open Sentences: Some Instructional Considerations from Research." Arithmetic Teacher, 19:595-599.

Gruen, Gerald E.
1966. "Note on Conservation: Methodological and Definitional Considerations." Child Development, 37:977-983.

Hamrick, Anna Katherine Barr
1976. "An Investigation of Oral Language Factors in Readiness for the Written Symbolization of Addition and Subtraction." Disseration Abstracts (University of Georgia, 1976) 37:4931.

Hiebert, James, Thomas P. Carpenter and James M. Moser
1982. "Cognitive Development and Children's Solutions to Verbal Arithmetic Problems." Journal for Research in Mathematics Education, 13:83-98.

Hinkle, Dennis E., W. Wiersma and S. G Jurs
1979. Applied Statistics for the Behavioral Sciences. Boston: Houghton Mifflin Company.

Hirstein, James Jay
1978. "Childrens Counting in Addition, Subtraction, and Numeration Contexts." Dissertation Abstracts International (University of Georgia, 1978) 39:7203.

Howlett, K. D.
1973. "A Study of the Relationship Between Piagetian Class Inclusion Tasks and the Ability of First Grade Children to do Missing Added Computation and Verbal Problems." Unpublished doctoral dissertation, State University of New York at Buffalo.

Inhelder, Bärbel and Jean Piaget
1969. The Early Growth of Logic in the Child. New York: W. W. Norton and Company, Inc.

Kennedy, Leonard
1975. Guiding Children to Mathematical Discovery, 2nd ed. Belmont, California: Wadsworth Publishing Company, Inc.

Keppel, Geoffrey
1982. Design and Analysis--A Researcher's Handbook, 2nd ed. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Kingma, J. and W. Koops
1983. "Piagetian Tasks, Traditional Intelligence and Achievement Tests." British Journal of Educational Psychology, 53(3):278-290.

Larson, S. C., R. M. Parker, and B. Trenholme
1978. "The Effects of Syntactic Complexity upon Arithmetic Performance." Learning Disability Quarterly, 1(4):80-85.

Le Blanc, John F.
1968. "The Performances of First Grade Children in Four Levels of Conservation of Numerousness and Three I.Q. Groups When Solving Arithmetic Subtraction Problems." Dissertation Abstracts (University of Wisconsin, 1968) 29:67.

Lindvall, C. Mauritz and Cheryl Gibbons Ibarra
1980. "Incorrect Procedures Used by Primary Grade Pupils in Solving Open Addition and Subtraction Sentences." Journal for Research in Mathematics Education, 11:50-62.

Lombard, Thomas J. and Robert G. Riedel
1978. "An Analysis of the Factor Structure of the WISC-R and the Effect of Color on the Coding Subtest." Psychology in the Schools, 15(2)176-179.

Lovell, Kenneth
1971. The Growth of Understanding in Mathematics: Kindergarten through Grade Three. New York: Holt, Rinehart, and Winston, Inc.

Markman, E. M., and J. Seibert
1976. "Classes and Collections: Internal Organization and Resulting Holistic Properties." Cognitive Psychology, 8:561-577.

Matlow, Robert Harvey
1976. "The Effects of Token Reinforcement and a Task Analyzed Program on the Acquisition of the Place Value Concept for Learning Handicapped Children with Low Mathematical Abilities." Dissertation Abstracts (University of Southern California, 1976) 37:7076.

Matthews, Julia
1981. "An Investigation into Subtraction." Educational Studies in Mathematics, 12:327-338.

May, Lola June
1974. Teaching Mathematics in the Elementary School, 2nd ed. New York: The Free Press, A Division of Macmillan Publishing Co., Inc.

Moser, James M.
1971. "Grouping Objects as a Major Idea at the Primary Level." The Arithmetic Teacher, 18(5):301-305.

Murray, Dennis E.
1973. "Mathematical Readiness of Six-Year-Olds and the Placement of First Grade Topics." Dissertation Abstracts (The Pennsylvania State University, 1973) 35:174.

Nicholas, Susan
1977. "Number Concept for School Beginners. Is It Possible?" The Arithmetic Teacher, 24(4):275-282.

Overholt, E. D.
1965. "A Piagetian Conservation Concept." The Arithmetic Teacher, 12:317-325.

Paschal, Billy J.
1968. "Readiness for Mathematics Learning." The Arithmetic Teacher, 15:5-6.

Payne, J. N. and E. C. Rathmell
1975. "Number and Numeration." Mathematics Learning in Early Childhood, ed. Joseph N. Payne. Virginia: The National Council of Teachers of Mathematics, Inc.

Piaget, Jean
1952. The Child's Conception of Number. London: Humanities Press, Inc.

Reisman, Fredricka K.
1977. Diagnostic Teaching of Elementary School Mathematics. Chicago: Rand McNally College Publishing Company.

Resnick, Lauren B., Margaret C. Wang, and Jerome Kaplan
1973. "Task Analysis in Curriculum Design: A Hierarchically Sequenced Introductory Mathematics Curriculum." Journal of Applied Behavior Analysis, 6(4):679-710.

Richmond, P. G.
1970. An Introduction to Piaget. New York: Basic Books, Inc.

Roberts, G. H.
1968. "The Failure Strategies of Third Grade Arithmetic Pupils." The Arithmetic Teacher, 15:442-446.

Robinson, Inez C.
1968. "The Acquisition of Quantitative Concepts in Children." Dissertation Abstracts (University of Southern California) 28:3038.

Schminke, C. W., Norbert Maertens, and William Arnold
1973. Teaching the Child Mathematics. Hinsdale, Illinois: The Dryden Press, Inc.

Simpson, Robert E.
1971. "The Effects of Cognitive Styles and Chronological Age in Achieving Conservation Concepts of Substance, Weight, and Volume." Dissertation Abstracts (University of Alabama, 1970) 31:3353.

Smith, Robert F.
1973. "Diagnosis of Pupil Performance on Place-Value Tasks." The Arithmetic Teacher, 20(5)403-408.

Steffe, Leslie P.
1967. "The Performance of First Grade Children in Four Levels of Conservation of Numerousness and Three I.Q. Groups When Solving Arithmetic Addition Problems." Dissertation Abstracts (University of Wisconsin, 1966) 28:885-86.

Suydam, Marilyn and Fred Weaver
1975. "Research on Learning Mathematics." Mathematics Learning in Early Childhood: Thirty-Seventh Yearbook ed. Joseph N. Payne. Virginia: The National Council of Teachers of Mathematics.

Thompson, Patrick Wilfrid
1982. "A Theoretical Framework for Understanding Young Children's Concepts of Whole Number Numeration." Dissertation Abstracts International (University of Georgia, 1982) 43:1868.

Underhill, Robert G.
1977. Teaching Word Problems to First Graders." The Arithmetic Teacher, 24:54-56.

Underhill, Robert G.
1972. Teaching Elementary School Mathematics. Columbus, Ohio: Charles E. Merrill Publishing Company.

Underwood, Benton J.
1957. Psychological Research. New York: Appleton-Century-Crofts, Inc.

Warner, Beverly Grace
1979. "Teaching Place-Value Understanding with Multi-Base Manipulatives in the Elementary School." Dissertation Abstracts (University of Arkansas, 1978) 43:1330.

Weaver, J. F.
1973. "The Symmetric Property of the Equality Relation and Young Children's Ability to Solve Open Addition and Subtraction Sentences." Journal for Research in Mathematics Education, 4:45-56.

Weaver, J. F.
1972. "The Ability of First-, Second-, and Third-Grade Pupils to Identify Open Addition and Subtraction Sentences for Which No Solution Exists within the Set of Whole Numbers." School Science and Mathematics, 72:679-91.

Weaver, J. F.
1971. "Some Factors Associated with Pupils' Performance Levels on Simple Open Addition and Subtraction Sentences." The Arithmetic Teacher, 18:513.

Wheatley, Grayson H., Jr.
1968. "Conservation, Counting, and Cardination as Factors in Mathematics Achievement among First-Grade Students." Dissertation Abstracts (University of Delaware, 1967) 29:1481-82.

Ziegenbalg, Sherry Bauer
1981. "Manipulation Versus Non-Manipulative Approaches to Teaching Renaming in Addition and Subtraction in Second Grade: An Experimental Study." Dissertation Abstracts (Rutgers University The State University of New Jersey, 1981) 42:1524.

Appendix A

THE INSTRUMENTS

INSTRUMENT #1

THE ADDITION-SUBTRACTION DIAGNOSTIC TEST

THE PRINCIPLE TO BE TESTED:
THE WHOLE SET AND ITS SUBSETS ARE THE COMPONENTS OF THE ADDITION AND SUBTRACTION PROCESSES

The Diagnostic Test

SECTION 1: ADDITION AND SUBTRACTION STORY PROBLEMS

Description of the Procedure

Six problems will be posed, in all. The first one is a warm-up problem which receives no score. The remaining five problems include difference, missing addend and missing minuend problems. Each of the five in this group will contain verbal "cue words" such as "3 or more are added..." or "he had 6 in all," to aid in the verbal interpretation of the problem presented.

For each problem, the examiner will read the problem to the subject and repeat all of the problem, as needed until the subject makes a response. Then the examiner should give the subject an opportunity to confirm or change his response in every case, whether the subject's first response is right or wrong.

For each problem, the subject will have concrete or pictorial items which can be used to represent the quantities referred to in the story problems. However, the examiner only makes these available for use and must never demonstrate or make suggestions as to how these materials might be used.

The Materials to be Used

1. A set of 20 small stacking blocks of uniform size should be used. Unifix Cubes, Multi-link Blocks, Lego bricks or any similarly sized stacking blocks suitable for table use will suffice.

2. A set of 20 playing cards with bird pictures on them is needed. Wild-life Federation Cards are used. These are sold as a children's card game. A regular deck of cards with bird pictures on the back or index cards with bird pictures pasted or sketched would suffice.

3. A set of 20 playing cards with picturs of various wild animals on them is needed. Cards from a National Wildlife Federation card game were used. Substitutes such as those described in the previous paragraph would be equally satisfactory as long as the cards have some sort of animal pictures on them.

Explanation to the Subject

"I am going to ask you some addition problems, like 3 + 5 = 8, and some subtraction problems, like 8 - 4 = 4, but you will not need to use any paper or pencil. I will read you a story and you will have some materials, which can be used to help you remember the numbers in the stories if you want to use them. I will be happy to repeat the stories as many times as you need because sometimes you can't remember the beginning by the time we get to the end. When you are ready you can tell me what you think the answer should be. Be sure to think about each story carefully, because I want to find out just how well you can figure these out when you are doing your best."

For each problem,

1. Give the subject the complete undivided collection of materials referred to in the story, but do not give any cues about their manipulation; (See Figures 3-5.)

2. Repeat the problem, slowly, exactly as written;

3. Pause, and then repeat the problem, if needed, but only verbatim. Repeat, as often as needed.

4. When the subject has responded, ask him to "listen very carefully," as you "read the story one more time to check (his) answer." and

5. Record the subject's final response opposite the question code number on his answer sheet.

A-S Warm-up Problem

"If your friend made a tower of 8 of these blocks, and then you add 5 blocks to his tower; afterwards, how many blocks tall would the tower be?"

A-S Trans. 1

"If you had a tower that was 12 blocks tall, and your friend had a tower 3 blocks tall, how many blocks taller than your friend's tower would your tower be?"

A-S Trans. 2

"Pretend that you made a tower 5 blocks tall, and then your friend put some more blocks on your tower. If the tower was 11 blocks tall after your friend added the blocks, how many blocks did your friend put on the tower?"

A-S Trans. 3

"Imagine that several birds were settled in a small tree and then a cat came along and 6 of the birds on the lower branches flew away. If 5 birds were still left in the tree after the cat scared 6 away, how many birds were in the tree before the cat came?"

A-S Trans. 4

"Pretend that some animals were running from a forest fire. Soon 6 animals joined them as the fire spread. Then there were 14 animals in all running from the fire. So, how many animals were running from the fire at the beginning, before 6 more came?"

A-S Trans. 5

"Suppose that 13 animals are drinking water around a forest pond. Then, at one end of the pond, a crow screams a warning that something dangerous is coming and the animals near the crow run into the forest. Afterwards, there are 4 animals still drinking from the pond, so how many of the 13 animals ran to hide when the crow screamed his warning?"

SECTION 2: CLASS INCLUSION QUESTION

Description of the Procedure

Three class inclusion questions will be posed, in which the subject is presented with two concrete subsets of an inclusive class, such as spoons and forks of the class silverware, and the subject will be asked to compare the size of the inclusive set to the larger of the subsets. The first questions will be a warm-up question followed by the two scored Addition-Subtraction test questions. As with the story problems, these questions are to be read exactly as written and may be reread, verbatim, as needed, but the wording must not be changed in any way.

If the warm-up question is answered incorrectly, the subject will be asked to repeat the question and the question will be reread if he cannot repeat it, in order to ensure that the subject is answering the question asked, and not a different question which he expected to be asked. However, this procedure will only be followed for the warm-up question. The first answer given will be recorded for the two scored Addition-Subtraction test questions.

Materials to be Used

1. One very clear but sturdy plastic bag about the size of a sandwich bag should be filled with at least 20 paper clips and 3 nails about 1 1/2 to 2 inches long. The bag should be sealed with a piece of scotch tape at the open end.

2. Three spoons and eleven forks are needed. Plastic picnic silverware would be quite satisfactory. (A silver-colored plastic will get you less arguments from smart little third graders about the "silverware" label.)

3. Two very clear but sturdy plastic bags are needed, about the size of a sandwish bag or larger, if longer pretzels are used. Put two cookies in one bag and seal it with scotch tape. Put at least 20 pretzel sticks in the other bag and seal it, also.

Explanation to the Subject-Warm-up Question

1. Take out the sealed plastic bag with nails and paper clips in it and place it flat in front of the subject. Spread out the paper clips and nails so they seem to fill the bag.

2. Say: "Now I'm going to ask you some tricky questions about several sets of things I'm going to show you. Be sure to listen very carefully to the questions because they may not be just the way you would expect....Now, here's the first question."

Warm-up Question

"In this plastic bag, I have a few nails and many paper clips, but the nails are made of metal and the paper clips are made of metal so all the things in the bag are metal things, aren't they?" (Pause)

"Now listen very carefully to the question...Are there more paper clips or more metal things?"

3. If the subject answers "more metal things," go directly to Part C. But, for the subject who answers "more paper clips," ask the subject--

"Okay, now I'd like you to tell me what the question was, that you just answered." (Pause)

If no response, say--"Try to describe the question I asked."

4. If the subject indicates that he was comparing the amounts of paper clips and metal things, go on to Part C directly. Otherwise, say--

"I'll reread the question slowly once more. Please listen very carefully because I don't think you heard all of the question correctly." (Reread the warm-up question--pause for the answer--reread once more, only if asked to, and after answered, proceed to Part C.

Presentation of the Problem-Part C

For each problem:

1. Present the appropriate materials in two separate sets, one kind of item grouped to the subject's left, and the other kind of item grouped to the subject's right, with at least 6 inches of space between the two groups.

2. Read the question slowly and clearly, pause for the answer rereading the question if necessary, and record the subject's response opposite the question code number on his answer sheet.

A-S Comp. 1

"Okay, now I have put out a few spoons and a whole lot of forks, havent't I?...But, both the spoons and the forks are kinds of silverware, aren't they? (Pause) Now, listen carefully to the question...Are there more forks or more pieces of silverware?"

A-S Comp. 2

"Fine, now here's the stuff for the last question. In this bag I have quite a few pretzel sticks and in this bag I have a couple of cookies--but both the pretzels and the cookies are food aren't they?...So, here's the tricky question...Are there more pieces of food, or more pretzels?"

SECTION 3: PICTURE PARTS AND WHOLES QUESTIONS

<u>Description of the Procedure</u>

Three sets of composite pictures will be presented to the subject, with each set including 5 different pictorial representations of the addition or subtraction process, within which only 1 of the 5 representations is correct. For example, the top equation in Figure 2 is a correct representation of the addition .process while the bottom equation is an unacceptable representation of the subtraction process since a part not found in the remainder was not subtracted from the whole, so that the remainder is less than it should be.

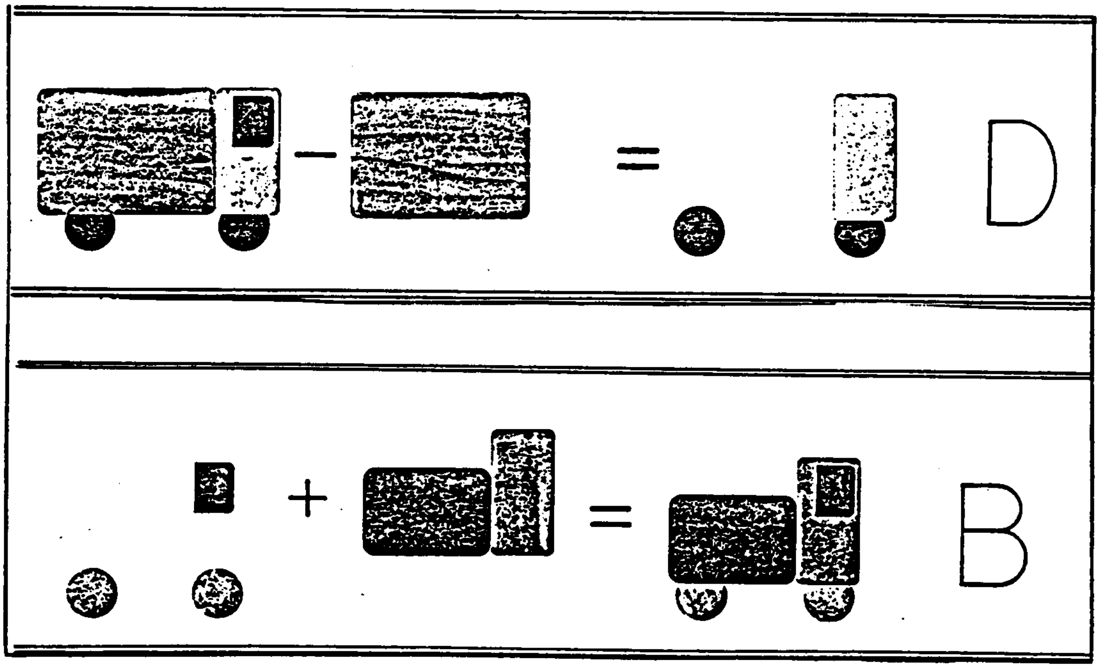

Figure 2. Examples of Picture Part and Whole Equations

Thus, when the subject is asked to select a correct symbolic representation of addition, from one set of 5 illustrations, the subject is essentially being asked if he, or she, knows that the symbolic representation of addition correctly includes the representation of two mutually exclusive subsets separated by a plus symbol with the combined total of the two subsets indicated to the far right of the horizontal sequence separated from the subsets by an equal symbol. Likewise, for the subtraction question, the subject is being tested both for the correct left to right sequence of representation (Whole - Part A = Part B), and for the inclusiveness of the Whole and the mutual exclusiveness of the two parts.

Pilot tests of this type of question have indicated that no warm-up question is needed as the 5 different illustrations presented initially seem to be enough examples of this unique representation form to permit the subject to give a reliable response to the first question posed. Therefore, just three problem sets are presented, one at a time, the question is posed, each time including an admonition to consider all possibilities carefully, and then each response is recorded.

If the subject answers at least 3 out of 5 Competence Level Addition-Subtraction questions correctly, he, or she, is asked to justify the last correct answer by explaining why it was selected. Two elements are required for an adequate justification, first a description of the generalized addition process as Part A + Part B = Whole (or the inverse for subtraction), and second, the stipulation that each piece of the whole picture is included in one and only one of the two partial pictures, with no duplications, extras or omissions. If the subject does not give an adequate explanation to the first request, then the subject is asked to explain <u>why</u> the first illustration, choice A, for the set being discussed, was not selected as a good example. On all three pilot tests of these instruments, it was observed that some subjects were more likely to explain all the essentials of a problem solving principle when confronted with an unsatisfactory example, for which they could explain the deficiencies. (See Figures 3-5.)

The Materials to be Used

Illustrations for the three problem sets to be used are included at the end of this section. Each set is constructed on 11" x 14" white posterboard and then laminated with clear contact paper for durability. The lines dividing the five examples and the identifying letters for each of the five were drawn in green ink while the parts of the object depicted were constructed using Avery self-adhesive imprinted labels (sizes 1 1/4" x 1 3/4", 7/8" x 1 1/4", 5/8" x 1 1/4", 1/2" x 3/4" and 1/2" diam.). Each label was colored with a selected artists' marking pen color so as to give both visual contrast and color balance among the parts.

Explanation to the Subject

"I'm going to show you some addition and subtraction examples that are a bit unusual." (Display the first set of 5 illustrations.) "As you can see, these show adding pieces of pictures instead of numbers, but I think you'll see that the addition process works about the same way that you're used to, even if these are adding different things. Of course, it might be that not all of these examples are really showing how to add exactly the right way so I want you to look over each set of examples very carefully, 1 by 1, and see if you can find any examples that you feel are really showing an addition sentence that is exactly right with no mistakes or errors. You have to look them over really carefully--not to get tricked!"

Presentation of the Problem

For each of the three problems,

(1) place the set of illustrations on the table in front of the subject,

(2) repeat the question verbatim, and

(3) record the response by the identifying letter, or letters, if the subject feels more than one is correct.

A-S Comp. 3

"On this first set, they're all addition sentences, so I want you to look each one over really carefully 1 by 1 and tell me why each one is a good or bad addition sentence...How about the next one?"

A-S Comp. 4

"Now, this set has all subtraction sentences, that's take away, so I want you to look them each over especially carefully 1 by 1 and tell me why each is showing the subtraction process correctly or incorrectly...And, the next one?"

A-S Comp. 5

"Okay, this one is the last one, and these examples are showing addition, again. Be sure to look them over quite carefully before you pick any. Do you see any addition sentences that are doing it exactly as they should, with no mistakes?...(if necessary) "Which ones?"

A-S Comp. Explanation

If subject answered 4 out of 5 A-S Comp. questions correctly, show him the last set of picture parts and wholes which was answered correctly and say, "Now I want to go right down through each of these examples and I want you to tell me just why each one is a good or bad example of the addition (or subtraction) process." Start with example A and record the subject's reason for accepting or rejecting each as an example of the addition (or subtraction) process.

Record the answers verbatim, describing gestures in parentheses if essential to the explanation. Ask the subject for a repeat if necessary to get it all down.

Examiner ______________________

Addition-Subtraction Score Sheet

Subject's Number ____________ Date and Time ____________

Story Problems		Response
Warm-up	# of blocks tall	______
Trans. 1	# of blocks taller	______
Trans. 2	# of blocks put on	______
Trans. 3	# of birds before	______
Trans. 4	# of animals at beginning	______
Trans. 5	# of animals that run	______

Class Inclusion	Response (Circle)	
Warm-up	paper clips	metal things
Comp. 1	more forks	more silverware
Comp. 2	pieces of food	pretzels

Picture Parts and Wholes		Response (Circle Choice)				
Comp. 3	House	A	B	C	D	E
Comp. 4	Van	A	B	C	D	E
Comp. 5	Dumptruck	A	B	C	D	E

A-S Comp. Explanation

A: ______________________

B: ______________________

C: ______________________

D: ______________________

E: ______________________

Correct Incorrect

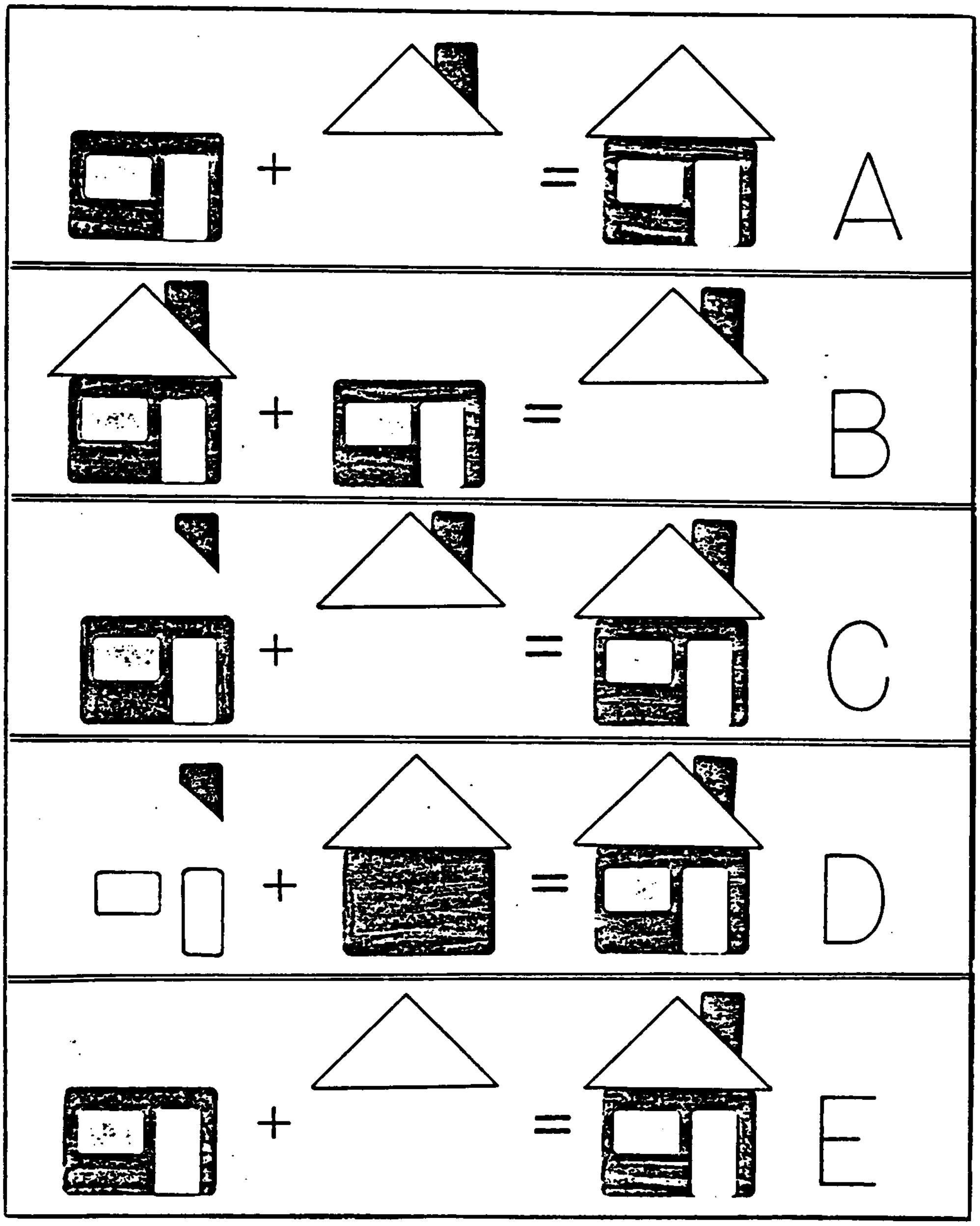

Figure 3. A-S Comp. 3 Set of Equations

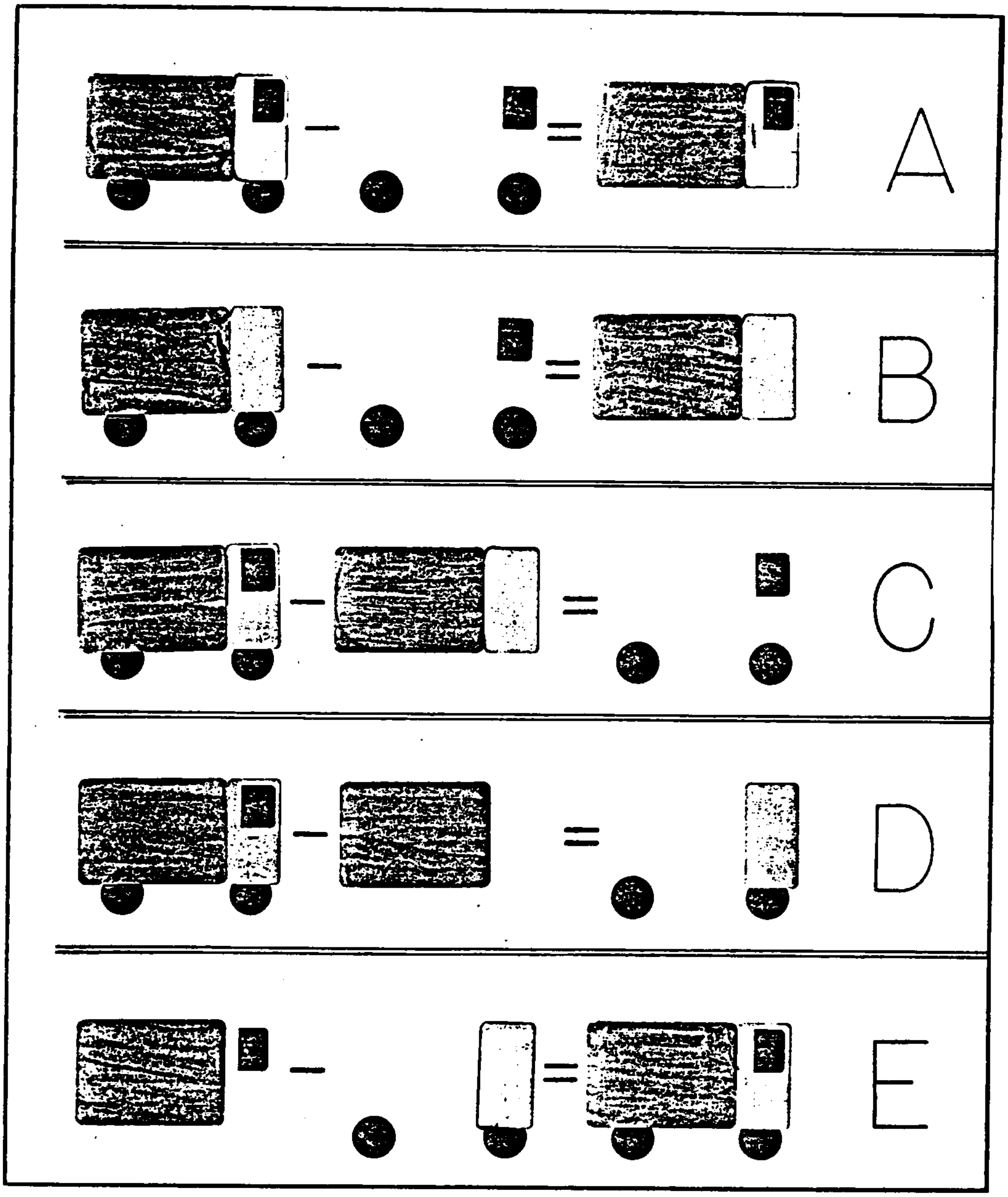

Figure 4. A-S Comp. 4 Set of Equations

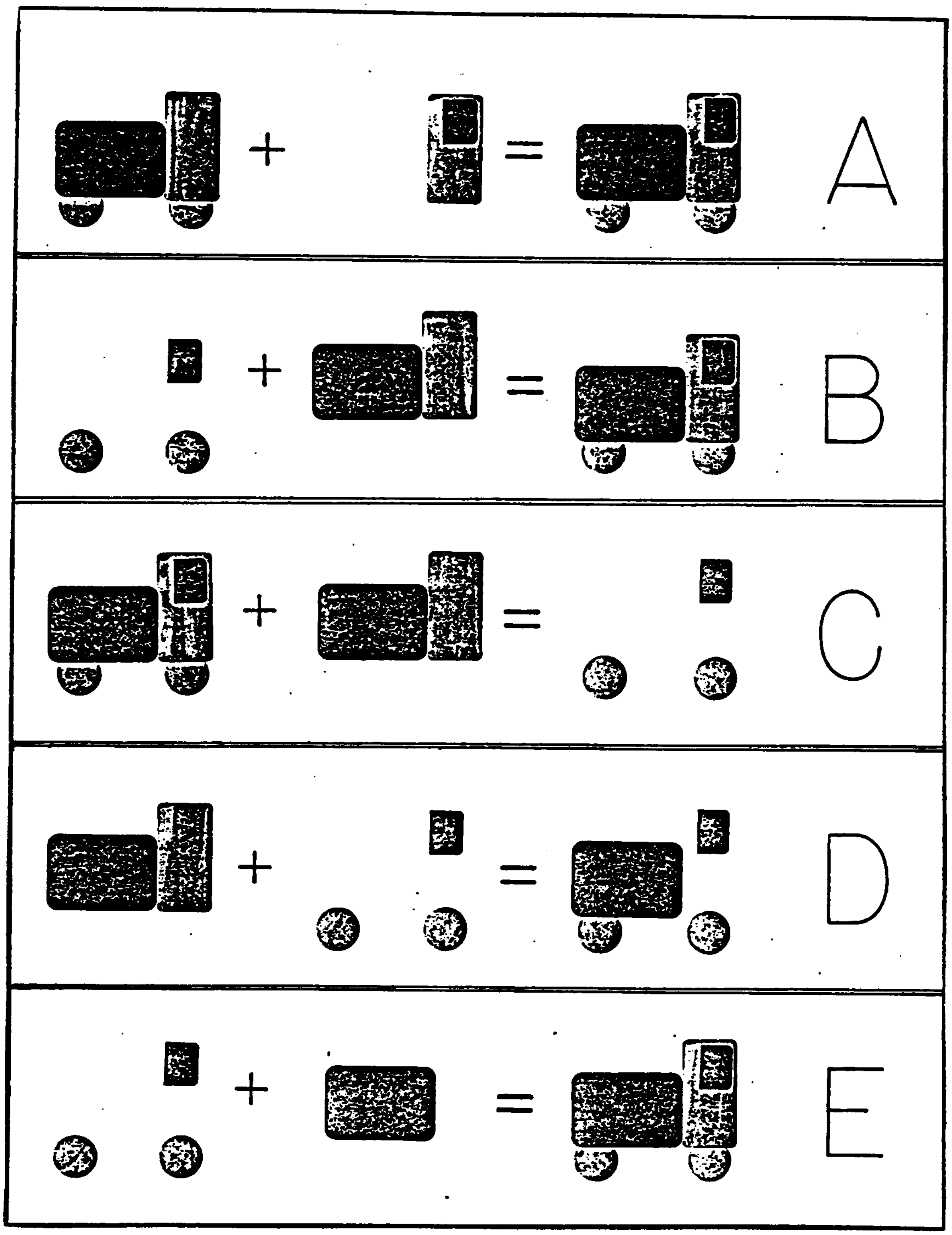

Figure 5. A-S Comp. 5 Set of Equations

INSTRUMENT #2

THE COUNTING BY GROUPS DIAGNOSTIC TEST

THE PRINCIPLE TO BE TESTED:
COUNTING BY GROUPS IS A FASTER, NO MORE ERROR PRONE COUNTING PROCEDURE THAN COUNTING ONE-BY-ONE

The Diagnostic Test

Description of the Procedure

There are two sections to this diagnostic test. The transitional understanding measure is comprised of five tasks in which the subject is instructed to count out a specified number of objects--with two trials counting by twos, one trial counting by tens, and two trials counting by fives. The criterion for success on each trial is simply to obtain the correct count using the specified grouping procedure. Since this is to determine if a transitional level of competence has been reached, the subject will be allowed one prompted recount on his first undetected miscount in the five trials. The subject will be warned, at that time, to be careful since an accurate count is important.

The competence level section of this diagnostic test is designed to determine if the subject will, of his own volition, employ the skill of counting objects by naturally presenting groupings rather than counting them one by one when pressed to count with as much speed and accuracy as possible. If the subject should count the number of groupings and then multiply by the number in each group, this is accepted as meeting criterion, also. One example of objects presented in pairs, one example grouped in fours (where counting by twos or fours will be acceptable), two examples grouped by fives, and one example grouped by tens (where counting by fives or tens will be acceptable) are included. In all the examples, the groupings are ones which always occur with the same number together such as in pairs of shoes, fingers on the hand, and wheels on a car, and are groupings commonly within the experience of a primary school child.

For the usual competence explanation measure, if the subject counts at least 4 of the examples by groups, either directly or using multiplication, he will be asked to justify his choice of this method at the conclusion of the five trials. The subject must state positively that he believes the counting by groups is both faster than and as likely to be accurate as counting one-by-one for the situation given.

Because the competence level tasks are designed to determine if the subject will choose to employ the more efficient counting-by-groups method without prompting, of his own volition, the competence level tasks are presented before the transitional level tasks. Clearly, the requests to count by twos, tens and fives on the transitional tasks would constitute a broad hint as to how to approach the competence level tasks and would eliminate the possibility of getting an unprompted, volitional performance measure. Furthermore, the concern for the competence level difficulty affecting performance on subsequent trials is not a factor here, since the child is free to count the objects in whatever way he chooses. Thus, the transitional level places more demanding constraints on the task, as far as explicit instructions are concerned. This is one of just two instruments for which this reverse order of presentation was found to be necessary and advisable.

As with all the diagnostic tests a subject must succeed on at least three of the five trials and eight out of the total ten trials or he will have failed the tested level of competence. This three out of five criterion applies both to the procedure of counting by groups and the accuracy of the count. Also, as in all the tests, self-correction is permitted, with the last response being recorded; thus, the examiner must refrain from giving facial or other clues as to the acceptability of all responses.

In order to reduce the total testing time, if the subject demonstrates competence in counting by twos, fives, and tens on the volitional section of the test, by counting accurately and without marked hesitation, each time counting by groups is employed, corresponding tasks on the transitional level will be scored as successes without repeating them.

Materials to be Used

For the first section, a watch with a second hand or a stop watch is needed. The sets to be counted are attached to poster board sheets about 10 inches by 12 inches in size. Spacings will vary, but in all cases, a natural group will be within close enough proximity to permit a visual grouping as they are affixed.

Card one----5 paper dolls with all fingers visible are affixed to the card. (See Figure 6.)

Card two----Cutouts of 7 gloves with all fingers showing in sizes from 2 1/2" to 3" long. (See Figure 7.)

Card three--5 cardboard cutouts of traced hands, each with fingernails done in a different color of fingernail polish. (Actual cosmetic counter samples of polished nail colors glued on a hand shape is very effective, but any coloring technique will do.) The hands can be overlapped as long as the visual groupings of 5 fingernails is facilitated. (See Figure 8.)

Card four---6 small plastic vehicles, each with exactly four wheels and no extras such as a spare.

Card five---9 pairs of Barbie-doll shoes, each pair clearly distinguishable by color and shape. (See Figure 9.)

For the second section, one set of 20 loose objects such as buttons, lego bricks, crayons, etc., is needed and one group of 45 objects attached, in some way, in groups of 5, such as lego bricks stuck together or cards of 5 buttons stapled or sewn on.

Explanation to the Subject

Give the instructions verbatim as follows: "In this game, all you have to do is count things, but the trick is to count them as quickly as you can without making any mistakes. I'll be timing you and writing down the time you take to count each set, but I have to add time for each one you're too low or too high on the count, so you need to count carefully as well as quickly to get the best score...We want to find out what ways of counting seem to work best, so please count out loud so I can see just how you are doing it, to be careful and fast, at the same time."

Presentation of the Task

Note 1. Remind the subject to count out loud, immediately withdrawing the card if he forgets, and then starting him over with the time.

Note 2. When the subject responds with the total for each question on the first section, immediately record on the subject's score sheet opposite the question-code-number the following:

a. the count total he told you,

b. the time in seconds elapsed since he began to count (or start and stop times), and

c. check either the Grp-Cnt box or the One-by-One box. The Grp-Cnt box is checked if he counted at least two-thirds of the objects by groups or by the multiplication method. Otherwise the One-by-One box is checked.

Grp-Cnt Comp. 1

Say: "On this first card, we want to know how many fingers there are in all, including all the fingers on all the dolls,...counting out loud, of course."

Present Card One and Start Timing. (You may reassure the subject that there are 5 fingers on every doll's hand if he asks.) If the subject asks, at any time during the presentation of the five Grp-Cnt tasks, if he is permitted to count by groups or by fives, tens, twos, or fours, repeat that the research purpose is to find the best ways to count things to be both fast and accurate and that the subject should get the total count in whatever way he feels is best for that purpose. Use the same explanation if the subject asks about using multiplication. Record and continue.

Grp-Cnt Comp. 2

Say: "This is the next card--it has gloves on it. I want you to count how many fingers there are in all, counting all the fingers and thumbs on all the gloves. Be sure to count aloud."

Present Card Two and Start Timing.

Record and continue.

Grp-Cnt Comp. 3

Say: "On this card, we want to know the total number of fingernails on the card, including all the fingernails on all the hands. Don't forget to count aloud."

Present Card Three and Start Timing.

Record and continue.

Grp-Cnt Comp. 4

Say: "On this next card, I want you to count aloud how many wheels there are in all--including all the wheels on all of the vehicles."

Present Card Four and Start Timing.

Record and continue.

Grp-Cnt Comp. 5

Say: "Okay, for this last card, I want you to count how many shoes there are in all on the card...Don't forget to count out loud."

Present Card Five and Start Timing.

Record and continue.

Grp-Cnt Comp. Explanation

If the subject counted at least three out of the five trials by groups, rather than one-by-one, refer to the last items counted by groups. Say: "When you counted these (item), you didn't count them one-by-one. You counted groups of (#) together. Weren't you afraid you would miscount or count too slowly counting by groups of (#)?"

If the subject does not justify his procedure in terms of both speed and reliability, ask the appropriate follow-up question(s).

(1) "Do you think counting one-by-one is faster or counting by groups is faster?...Why?

(2) "When things are bunched in natural groupings, as these are, do you think counting one-by-one is better for you, to count correctly without miscounting or is counting by groups better for you to avoid miscounting?...Why?"

Note 3. Do not readminister the counting by twos tasks if the subject performed this task successfully on the competence level tasks. Simply score these tasks as successes. Do likewise for the counting by fives and tens tasks.

Note 4. Place the appropriate set of objects in front of the subject in a pile, read the question directions, then time the subject and record the number actually counted out, and check the grp-cnt box if the counting procedure was as directed.

Grp-Cnt Trans. 1

Present Set One (20 loose buttons) and say: "Now we're going to do things a little differently, I'm going to tell you how to count these and how many to count out from the pile. You just need to do it carefully to be sure you get the right number. I won't time these, but be sure to keep counting out loud so I can see how you're doing it."

"First, I want you to count by two's (like 2, 4, 6, 8,...) and count out 14 buttons from the pile counting by two's."

Allow one recount if his first is a miscount, and record.

Grp-Cnt Trans. 2

Say: "This time, I want you to count out 18 buttons counting by two's."

Allow one recount if his first is a miscount and record.

Grp-Cnt Trans. 3

Put away set One and place Set Two (cards of buttons) in front of the subject.

Say: "Okay, three more to go! First, count out 20 buttons, counting by five's."

Grp-Cnt Trans. 4

Say: "Now the next to last one! Let's see you count out 30 buttons, counting by 5's again."

Grp-Cnt Trans. 5

Say: "Now for the last one, I want you to count out 40 buttons, counting by <u>tens</u> this time."

Allow one recount for the first miscount and record.

Examiner ______________________

Counting by Groups Score Sheet

Subject's Number ______________ Date ______________

Sets-of-Items-on-Cards Count

	Count	Time	Response: By Groups	Response: One-by-One
Comp. 1 (5's or 10's)				
Comp. 2 (5's)				
Comp. 3 (5's or 10's)				
Comp. 4 (2's or 4's)				
Comp. 5 (2's)				

Grp. Cnt. Comp. Explanation

____

____

____

____

Count-Out-By-Groups	Correct Response: Actual # Counted	Incorrect Counted by Groups: Yes	Incorrect Counted by Groups: No
Trans. 1 (14 by 2's)	______________		
Trans. 2 (18 by 2's)	______________		
Trans. 3 (20 by 5's)	______________		
Trans. 4 (30 by 5's)	______________		
Trans. 5 (40 by 10's)	______________		

Figure 6. Grp-Cnt Comp. 1 Task Card

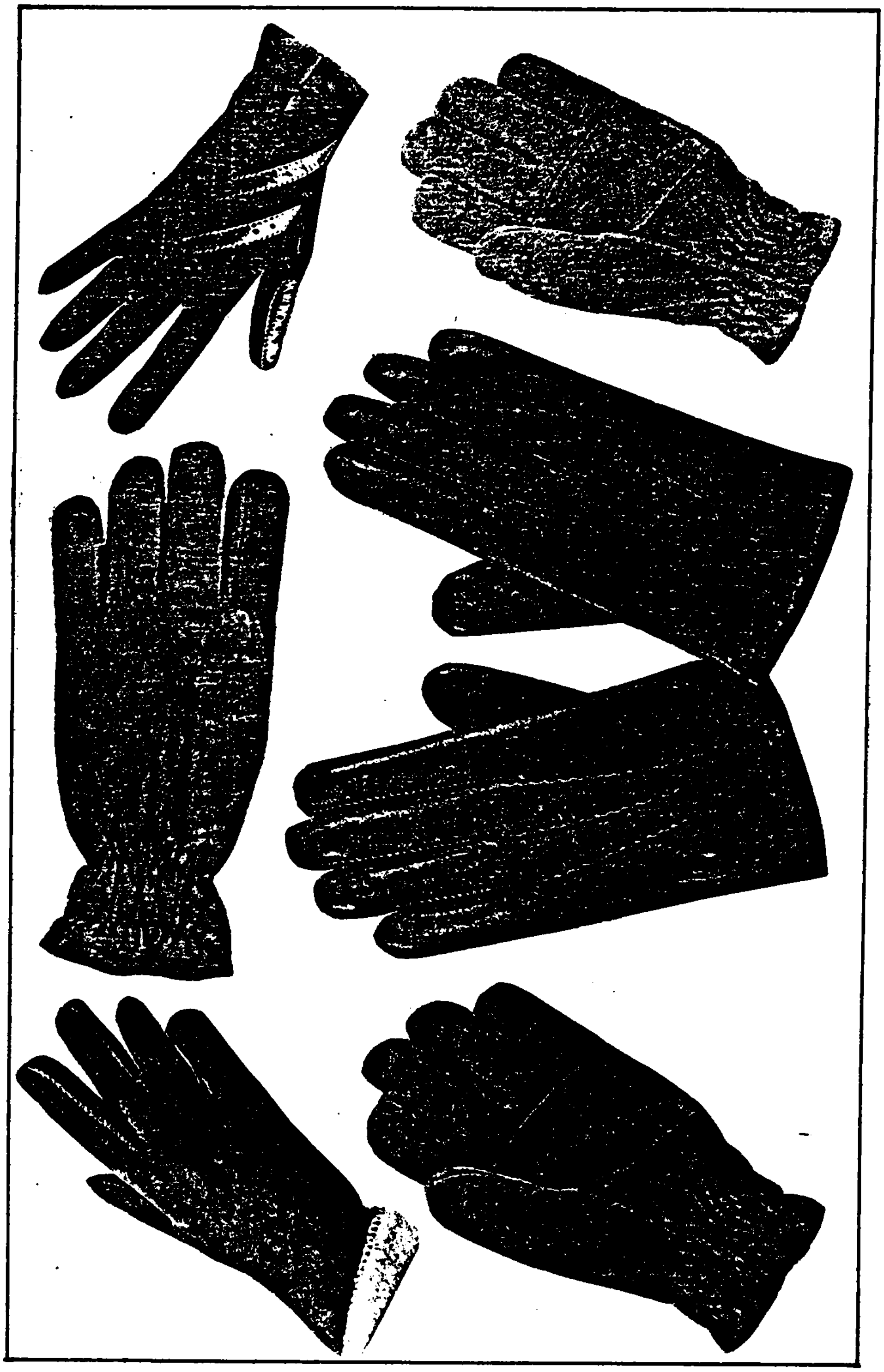

Figure 7. Grp-Cnt Comp. 2 Task Card

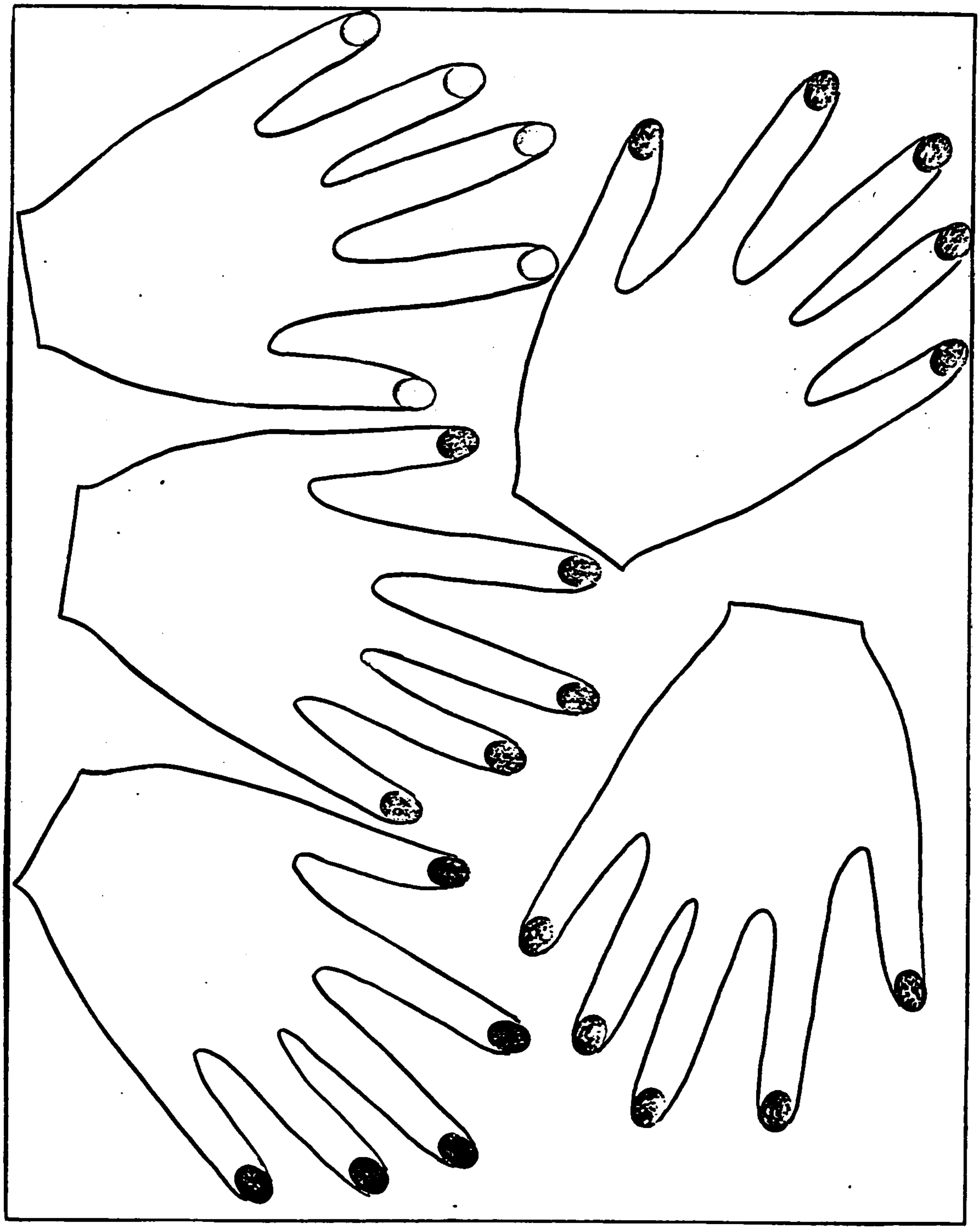

Figure 8. Grp-Cnt Comp. 3 Task Card

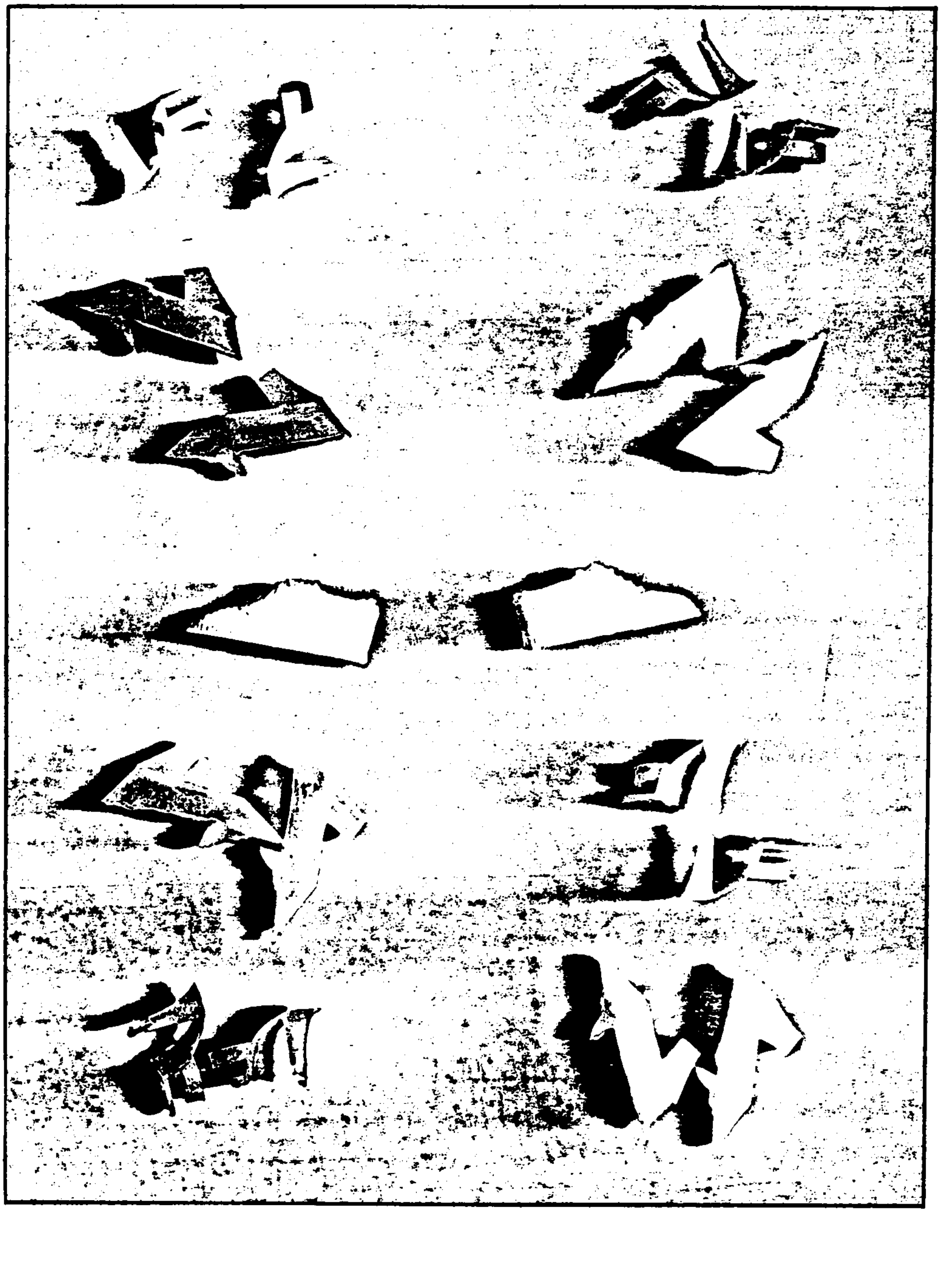

Figure 9. Grp-Cnt Comp. 5 Task Card

INSTRUMENT #3

THE ONE EQUAL TO MANY DIAGNOSTIC TESTS

THE PRINCIPLE TO BE TESTED: ONE OF A DESIGNATED ITEM MAY BE CONSIDERED TO BE EQUIVALENT TO MANY OF ANOTHER ITEM

The Diagnostic Test

Description of Procedure

This diagnostic procedure will use money exclusively to test the level of development of the concept of different denominations with different values. Money was chosen because previous pilot testing indicated that the greatest number of second and third graders had developed a high level of understanding of denomination equivalences with respect to money as compared with measurement of other parameters such as time, length, liquid and dry volumes of weight. While all of these types of measurement are at least introduced by the end of second grade, there seemed to be more consistent opportunities for continued experience with money measurement than with others in the groups tested.

The transitional level of the test will consist of five trials each consecutively using a higher total amount of money than previous ones. In each trial, the subject will be asked to determine the value of the cash presented. After each response, the subject is asked how many pennies would be needed to have the same amount of money, but with all pennies and no other kinds of coins or bills. This is to determine if the actual comparative values of the coin and bill denominations are really understood or if the counting of the cash is just a learned rote response. Both the conventional counting of the coin total and the penny equivalent in each of the trials must be correct to earn a transitional level success score on each task in this section.

The other group of questions is designed to measure the applied level for real competence. Thus, in this group of five trials, the subject will be given a collection of coins and bills and then will be asked to make up a specified amount in two different ways. For example, 22¢ could be made up with two dimes and two pennies, or with one dime and twelve pennies or with one dime, two nickels and two pennies, etc. In this way, the subject will be required to use his understandings of equivalence relationships among the different money denominations in ways other than the most conventional. In the last three trials, the coins available will be carefully limited so that not all of the many possible ways of making up the amount can be used. This, of course, makes further demands on the subject's ability to apply his understandings of equivalences among different money denominations. As in all the diagnostic tests, at least eight correct trials out of the ten attempted will be necessary for a success rating on the competence level.

For both parts of the test, commercially available facsimile coins and bills will be used, since real cash tended to disappear during the pilot testing.

The competence explanation will be elicited by giving the subject a five dollar bill to put with his remaining coins on the last trial (there should be 20¢ remaining), and asking him he can pay $2.55 again. If the subject responds "yes", he will be asked to do so. If he responds "no" eventually, he will be asked to explain "why not" and asked if he has more or less than $2.55. A correct score will be given if the subject indicates that he needs to exchange the five dollar bill for an equal amount in smaller denominations, whether this is indicated positively—as the solution to the dilemma—or negatively—as the reason why he cannot pay the amount. As long as the subject indicates that:

1) he has more than $2.55 and

2) he does not have the amount in the proper denominations to take out two dollars and 55 cents, he will receive the "correct" score.

Because these trials are relatively time consuming, and because results on the pilot tests indicated that children did reliably obtain equal or higher scores on the transitional level than the competence level, the competence level is given before the transitional level, with an automatic equal score given for the transitional level if 4 or 5 correct answers are given at the competence level. If a score of 3 or less is obtained at the competence level, then the transitional level tasks are administered to determine the subject's transitional level of performance.

Materials to be Used

Facsimile coins and bills are used to make up all the amounts specified. Care should be made to obtain clearly recognizable facsimiles closely resembling real money in size, color and markings. Many educational supply companies have inexpensive fake coins and bills which would serve the purpose.

For the transitional section, separate packets or divisions of a box with the amounts already made up will expedite the testing procedure. The amounts of each trial are as follows:

Trial #	Pennies	Nickels	Dimes	Quarters	Dollars	Total
1	1	1	1			$ .16
2	2	2	2			$ .32
3	5	3	3	0	1	$1.50
4	2	3	4	1	1	$1.82
5	3	0	1	3	2	$2.88

For the competence section, the following collections of coins should be ready in separate packets or sections of a box:

Trial #	Pennies	Nickels	Dimes	Quarters	Dollars
1	20	5	3	0	0
2	15	5	3	3	0
3	9	3	7	4	0
4	13	2	4	3	1
5	0	1	5	7	3

For the Competence Explanation Question, you will need one five dollar bill.

Explanation to the Subject

Give the explanation verbatim, as follows:

"In this activity, you're going to be counting up different amounts of money. Since we're using fake money rather than the real stuff, if you're not sure what kind of coins or bills they are supposed to be just ask me. I think they look pretty much like they're supposed to, but you may not be sure of the amount for some of them, so just ask me, to be sure, if you want to." (They should only be identified by the examiner as pennies, nickels, or quarters--not for their numerical value.) I'll give you a handful of money and you have to pick out the money you need, to pay me a certain amount. You just need to be careful to give me just the right amount."

One-Many Comp. 1

For the first trial in the competence section,

1) present the coins specified for Trial 1 and explain: "I want you to pretend that is your money and you are paying me for something. I just want you to give me 24 cents from these coins. Pretend that you are buying something, and that you must pay me exactly 24 cents for it, because I don't have any change."

2) record the number of each denomination used, and the total amount the subject selected;

3) then say: "But now, you still have quite a lot of money left. Can you pay me 24 cents again?...Try and see if you have enough to pay me 24 cents again;" and

4) record the number of each denomination used and the total amount selected. (Record zeroes if the subject claims 24 cents cannot be paid from the money still left.)

One-Many Comp. 2

Repeat the same procedures 1) through 4) using the specified money packet for trial 2, but this time, ask the subject to pay you 52 cents each time. Record his responses in the One-Many Comp. 2 spaces.

One-Many Comp. 3

Repeat the same procedures 1) through 4) using the trial 3 money packet and asking the subject to pay you 87 cents and recording the responses in the appropriate spaces so that the subject pays you 87 cents twice.

One-Many Comp. 4

Do the same as for trial 3, except use the trial 4 money packet and ask the subject to pay you 1 dollar and 13 cents each time. He should pay it twice, in all. If the subject uses all the pennies on the first payment, he may be allowed to redo the first payment after he has indicated that he has a problem with the second payment.

One-Many Comp. 5

Repeat the same procedures as for trials 3 and 4, but use the trial 5 money packet. Ask the subject to pay you 2 dollars and 55 cents each time. If the subject has only quarters left after the first payment and recognizes that this is a problem, return the first payment and let him try it another way, anticipating the second payment.

One-Many Comp. Explanation

If the subject has responded correctly to four out of five previous trials, do the following:

After putting away the trial 5 money packet, ask while giving the subject the five dollar bill, "If I give you this five dollar bill, can you pay me two dollars and 55 cents again?"

If he says "yes," ask him to pay you that amount.

When he says he cannot pay you the exact amount, ask him to explain why he cannot. Also, if he does not volunteer the information in his explanation, ask him if he has "more than $2.55 or less than $2.55," and "What could you do with the $5 bill so that you could pay me exactly $2.55?" Record the subject's explanations carefully.

If the subject did not score at least 4 out of 5 on the competence section, either because of miscounts or because he quit due to the difficulty of the tasks, proceed with the transitional level questions below. However, if the subject has scored at least 4 out of 5 at the competence level; give him an equal score on the transitional level and discontinue the test at this point.

One-Many Trans. 1 through 5

Present each of the coin collections for Trials 1-5 successively, ask the subject how much money he has in each collection, and record the subject's responses under appropriate columns successively, on the subject's response sheet.

For each trial, remind the subject to count the money carefully (as you give him the money) because you want to know how well he can count money when he's doing his best.

For each trial, after the subject has told you how much money he has, ask him, "How many pennies would you need, to have that same amount of money, but all in pennies?" Record these responses under the pennies column of Trans. 1-5 on the recording sheet.

Examiner ______________________

One Equal to Many Score Sheet

Subject's Number ______________ Date and Time ______________

Counting Out the Requested Amount — **Responses**

	am't requested	. P . N . D . Q . $.	total value
Comp. 1	$ .24	. . . . . .	______
Comp. 2	$ .52	. . . . . .	______
Comp. 3	$ .87	. . . . . .	______
Comp. 4	$1.13	. . . . . .	______
Comp. 5	$2.55	. . . . . .	______

One-Many Comp. Explanation

____

____

____

____

Correct Incorrect

Value of the Cash Presented — **Responses**

		one-many value	pennies am't
Trans. 1	. 1 . 1 . 1 . . .	______	______
Trans. 2	. 2 . 2 . 2 . . .	______	______
Trans. 3	. 5 . 3 . 3 . 0 . 1 .	______	______
Trans. 4	. 2 . 3 . 4 . 1 . 1 .	______	______
Trans. 5	. 3 . 0 . 1 . 3 . 2 .	______	______

INSTRUMENT #4

THE N-MORE OR LESS DIAGNOSTIC TEST

THE PRINCIPLE TO BE TESTED:

N MORE THAN ANY NUMBER OF ITEMS IS GIVEN BY THE N^{TH} NUMBER ABOVE THE ITEM NUMBER IN THE COUNTING SEQUENCE

The Diagnostic Test

Description of the Procedure

First, the subject will be asked to order the color-grouped dot strips from 1 to 27, so that the one with the least number of dots is at one end, then the one with the next to smallest number of dots,...up to the one with the greatest number of dots. (The dot strips are shown in Figure 10 at the end of this instrument.) This is done so that the subject will know that all the numbers from 1 to 27 are represented and that the strips are all properly ordered in counting order.

Next a set of small counters (such as 1 cm. cubes) are given to the subject and he is asked to place them on the dot strip which has the same number of dots as he has counters, placing one counter on each dot to make a self check. Then, the subject is given one more counter and asked to predict which dot strip all the counters will just match, now. The subject is asked to check his prediction by placing all the counters on the strip he chose and to correct the placement if his prediction is wrong. This is to ensure that the subject understands the prediction objective clearly.

The next step is to set up the actual testing situation. The counters are removed and then posterboard strip is laid over the upper part of all the dot strips so that only the first four or five dots on each strip are visible, rendering strips 6-27 indistinguishable except by position. Then the examiner quickly whisks off several strips on both ends of the 1-27 sequence, simultaneously, so that the subject will not know just how many strips have been removed. At this point, the subject should <u>not</u> be sure of how many dots are on any one particular strip remaining on the table.

With this set-up, counters are placed on the exposed dots of one strip near the center of the sequence. Then the subject is asked to pretend that the strip with counters has counters <u>under</u> the posterboard, also, so that there are just as many counters as there are dots on that strip, just as they were before when <u>he</u> placed the counters on their matching strip.

From this point, the actual trials begin. The subject is given more counters or some counters are removed (the extras are placed on the posterboard above the exposed counters). Then the subject is asked to predict which strip all the counters would just match—<u>including</u> the counters we are pretending are under the posterboard. The N More-Less transitional trials will be two more, two less, three more, two more, and three less, in that order. If the subject is very hesitant about the prediction, he is asked if he thinks it is impossible to know what strip will match the set with 1 less counter (or whatever the present task). If the subject is quite sure that there is no way to predict reliably without knowing the actual number of dots on the strip with the counters, the test is discontinued.

In addition, if the subject is obviously counting all the counters and is assuming the dots equal them (even though not seen), and is evidently making decisions based on finding the strip with dots equal to the resultant number of counters by counting in sequence from the first or last strip, the response should be scored as incorrect, even if the subject does happen to move the correct number up or down. This scoring procedure is necessary because a subject is apt to get the last several responses correct by applying the unacceptable procedure consistently, simply because we assume the starting point correct at the beginning of each move, and this illegal counting system will produce correct moves after initial errors if it is applied consistently without changing initial assumptions.

The test can be discontinued after the first five trials if the subject is consistently counting all the counters for each move in order to make a decision. This procedure is quite obvious to spot, simply because it is so time consuming.

Otherwise the N-More-Less Competency trials are presented. The same procedure is continued with 5 less, 6 more, 7 less, 8 more, and 7 less. If the subject responds on at least three out of the last five trials, he then is asked to explain his "trick" for figuring out which dot strip the counters will match.

Maaterials to be Used

1. The set of dot strips which are pictured in Figure 10 are made from posterboard strips, approximately 7/8" x 11" each, and AVERY brand stick on dots in two colors. The dots are applied in color groups of five as pictured. It is important that they be equally spaced on the strips so length of dot groups can be used for a visual check of proper ordering both by the subject and the examiner. The strips used had white and orange dots on sky blue posterboard. The strips were laminated for durability after the dots were applied.

2. Twenty counters which will fit on the individual dots on the dot strips (such as Cuisinaire one rods or one centimeter cubed plastic one gram cubes sold by many educational supply companies or electrical connector caps, etc.)

3. A piece of posterboard about 18" long and 8" wide.

Explanation to the Subject

Give the child the twenty-five dot strips and say:

"I want you to put these dot strips in a row, in order, so that the strip with the smallest number of dots is over here (point to the child's left) and then the strip with the next to smallest number of dots comes next, and then the next smallest,...(sliding finger to subject's right)...and ending the row over here with the strip with the largest number of dots."

If the subject is confused and does not understand the directions, you may help the subject get started by asking him to first find the strip with the smallest number of dots. When this is done, tell the subject to look over the remaining strips and find the one with the smallest number of dots left. Have him place that one, and look for the next smallest, etc. After three to five strips have been placed, the subject should be able to continue on his own. If the subject still has not caught on after the tenth strip, discontinue the test.

If the subject misplaces any of the strips, point to the misplaced one(s) and ask him to check and see if it is really in the right place. Continue until all strips are placed correctly.

When all twenty-five strips are correctly ordered, give the subject a set of nine of the counters (picking up 3 + 3 + 3 without any visible counting) and say:

"I would like you to count these carefully and tell me how many you have."

If the subject cannot count them correctly, even with a second try (to allow for carelessness), do not continue this task.

If counted correctly, then, say:

"Okay, good, now I want you to find the strip with that same number of dots, 9 dots, and point to it." (Pause until subject points.)

"Okay, now I want you to check and see if you counted correctly by putting one of your counters on each dot of that strip. If you are right, the counters will just fit on the strip with no dots left over and no counters left over."

*If necessary, demonstrate for the subject placing the counters on the selected strip.

**If the subject picks the wrong strip, move the counters to the correct strip saying, "let's try them on this strip and see if they just fit."

Now say:

"Okay, now we're getting to the tricky part of the game. We have our counters on the 9 strip now don't we? So, you have 9 counters right now."

"From now on, I don't want you to count all the counters or the dots on a strip any more. You may count the counters I give you or take back from you, but not the ones that are already on the strip."

"First, I'm going to give you one more counter." (Place the new one next to the 9 strip.) "Now, you have too many counters to fit on that strip because you have one more. Now, without counting your counters or any dots, which strip do you think all the counters will fit on just exactly right?" (Pause until he decides, but if he takes very long, caution him not to count, but make his best guess.) When the strip has been selected, say:

"Okay, let's check and see. Place your counters on the dots on that strip and see if you have exactly one counter for each dot."

If the subject misplaces the counters so that they don't fit exactly on the strip selected, ask the subject to select another strip and place the counters on it to check his selection.

When the subject has placed the 9-plus-one-more counters correctly, say, "Very good, but now our game is going to get a little more complicated."

(Perform the described procedures as you talk. To remove the strips at both ends of the sequence, place your outspread hands over the end four or five strips at both ends of the sequence simultaneously and sweep several strips into your hands with your thumbs immediately and place the removed strips out of sight quickly while still talking.) Say:

"I'm going to remove these counters now, so I can put this cardboard down over the top part of all the strips..." (See Figure 11 for correct placement.) "And, then we'll take some of the strips off each end so our row won't be quite so long...Now, I'm going to put counters on the dots that aren't covered up on this strip here." (Place the counters on a strip about at the middle of the sequence still remaining and add a few more counters to the ten you had on the strips before covering them up.) "I want you to pretend that there are counters on this strip underneath the cardboard, too, so there are just the right number of counters to fit on all the dots on this strip, just as we did before we covered up part of each strip."

Say, "Of course, we don't know just now many counters we have now, counting the ones we're pretending are under the cardboard, too, but that's okay!...That's the tricky part of the game now."

<u>Presentation of the Task</u>

N-More-Less Trans. 1

Place two extra counters below the strip with the counters and say, "Now I'm giving two more counters. Pretend that you now have two more counters than 11 the counters that would fit on this trip exactly. Including these two extra counters, and the counters we can see, and all the pretend counters that would be on this strip, on which strip do you think they will all fit, now, so there are just enough for each dot?" If the subject seems hesitant or confused, urge him to "give it a try" as he did before the strips were covered and see how it goes. Do not discontinue the test until at least two moves have been tried with the strips covered. (Pilot studies demonstrated that some subjects just have to plunge in to understand the directions clearly.) (Pause until the subject selects a strip.) Then say, "Okay, let's move the real counters to that strip—we can put the extras on top of the cardboard, about where they would fit on dots. We'll still pretend we have just enough to cover all the dots."

While the subject moves the counters to the newly selected strip, record his response as plus two, if correct, or as plus X if he moved the counters X strips up or as minus X if he moved the counter x strips down, in number. (X is of course the actual number of strips to the right or left of the starting point for each step.) The response is, of course, recorded on the subject's score sheet, opposite N-More-Less Trans. 1.

For Trans. 2 through Trans. 5, follow the exact same procedure as for Trans. 1, adding or removing counters as indicated below. If the subject seems to be making blind guesses or appears very insecure after at least two trials have been completed, ask the subject, "Do you think that it is impossible to be sure of which strip will be the right one for (2 more) counters, without knowing just how many counters we have, in all, including the pretend ones?"

If the subject is not sure it is impossible, continue the trials. If the subject is quite positive that he cannot predict the strips with any assurance of being right, discontinue the trials for both sections of the test and check the Judged Impossible box on the score sheet at this point in the trials.

N-More-Less Trans. 2

Take two counters away from the total on the strip and cardboard coverup.

N-More-Less Trans. 3

Add three counters to the total below the strip with the counters.

N-More-Less Trans. 4

Add two more, as before.

N-More-Less Trans. 5

Take three counters away, as before

N-More-Less Comp. 1

Take away five counters.

N-More-Less Comp. 2

Add six counters.

N-More-Less Comp. 3

Take away seven counters.

N-More-Less Comp. 4

Add eight counters.

N-More-Less Comp. 5

Take away seven counters.

N-More-Less Comp. Explanation

If the subject predicted correctly on at least three out of five of the competency trials, remove the "cover-strip" from on top of the ordered dot strips and place the correct number of counters (one on each dot) on the strip with 13 dots. Say, "Okay, now we're all set up again, so we can see what we're doing. Now I want you to explain to me--if I give you 8 more counters, how will you figure out where they fit exactly, without counting the dots on the strips?" Rephrase the question for 7 less counters if the explanation is not entirely complete or doesn't include the removing of counters as well as adding them.

The explanation is scored as correct if the subject indicates that he doesn't need to know the total number of counters, but only how many <u>more</u> or <u>less</u> than <u>before</u> and gives the rule of N steps up for N more and N steps down for N less. An implied indication that the total count is unnecessary is sufficient. For example, "You <u>only</u> need to know how many more..." would be acceptable.

<u>Try to get</u> a <u>generalized</u> explanation independent of the number of counters being added or removed, but if not forthcoming, explanations

for several different specific additions and removals will suffice to imply a general rule.

Examiner ________________________

N-More-Less Score Sheet

Subject's Number __________ Date and Time ______________

N-More-Less	Response	Judged Impossible
Trans. 1	____________	
Trans. 2	____________	
Trans. 3	____________	
Trans. 4	____________	
Trans. 5	____________	

N-More-Less	Response	Judged Impossible
Comp. 1	____________	
Comp. 2	____________	
Comp. 3	____________	
Comp. 4	____________	
Comp. 5	____________	

N-More-Less Explanation

____

____

____

____

____

____

Correct Incorrect

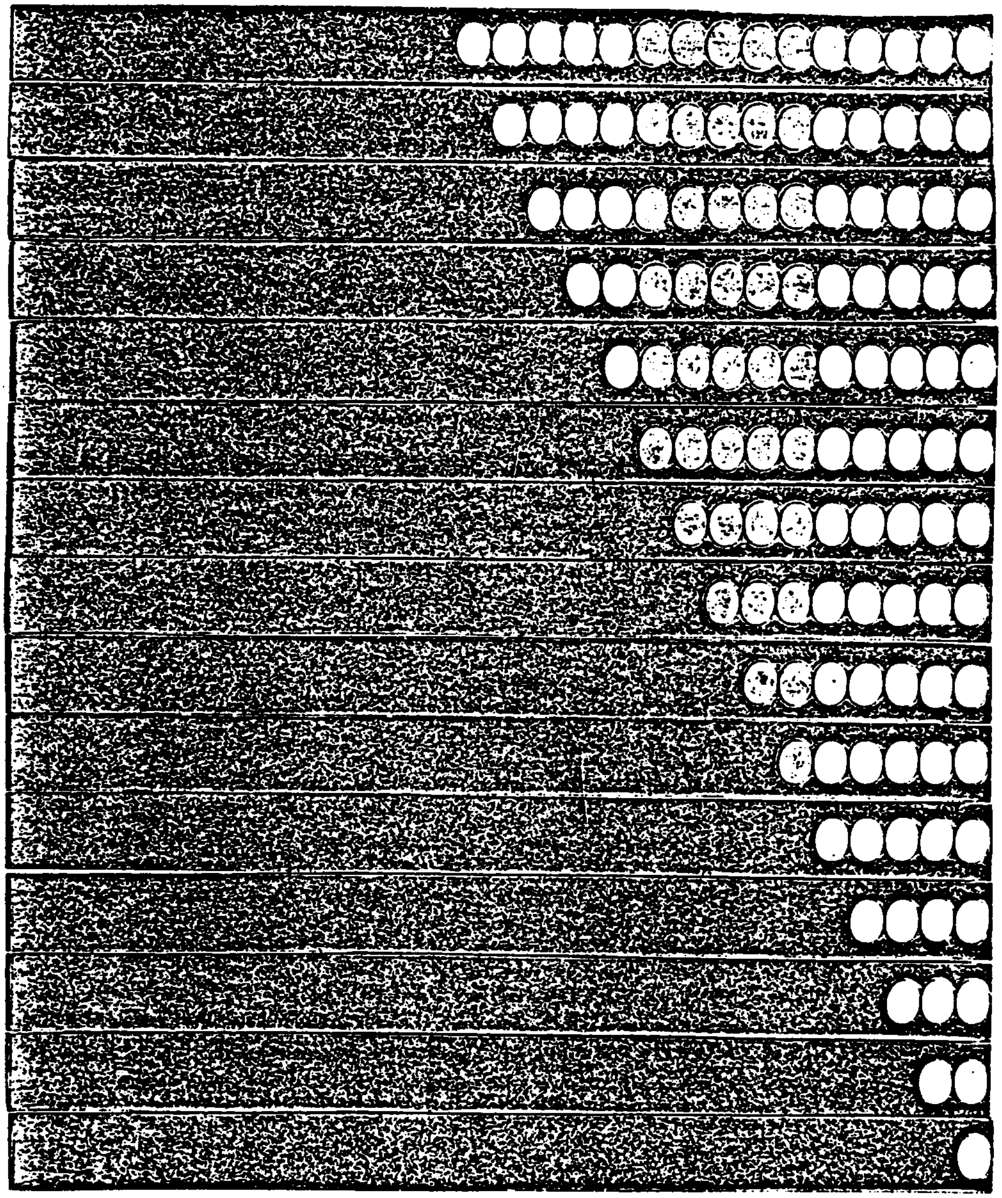

Figure 10. Fifteen of the Dot Strips Arranged in Order

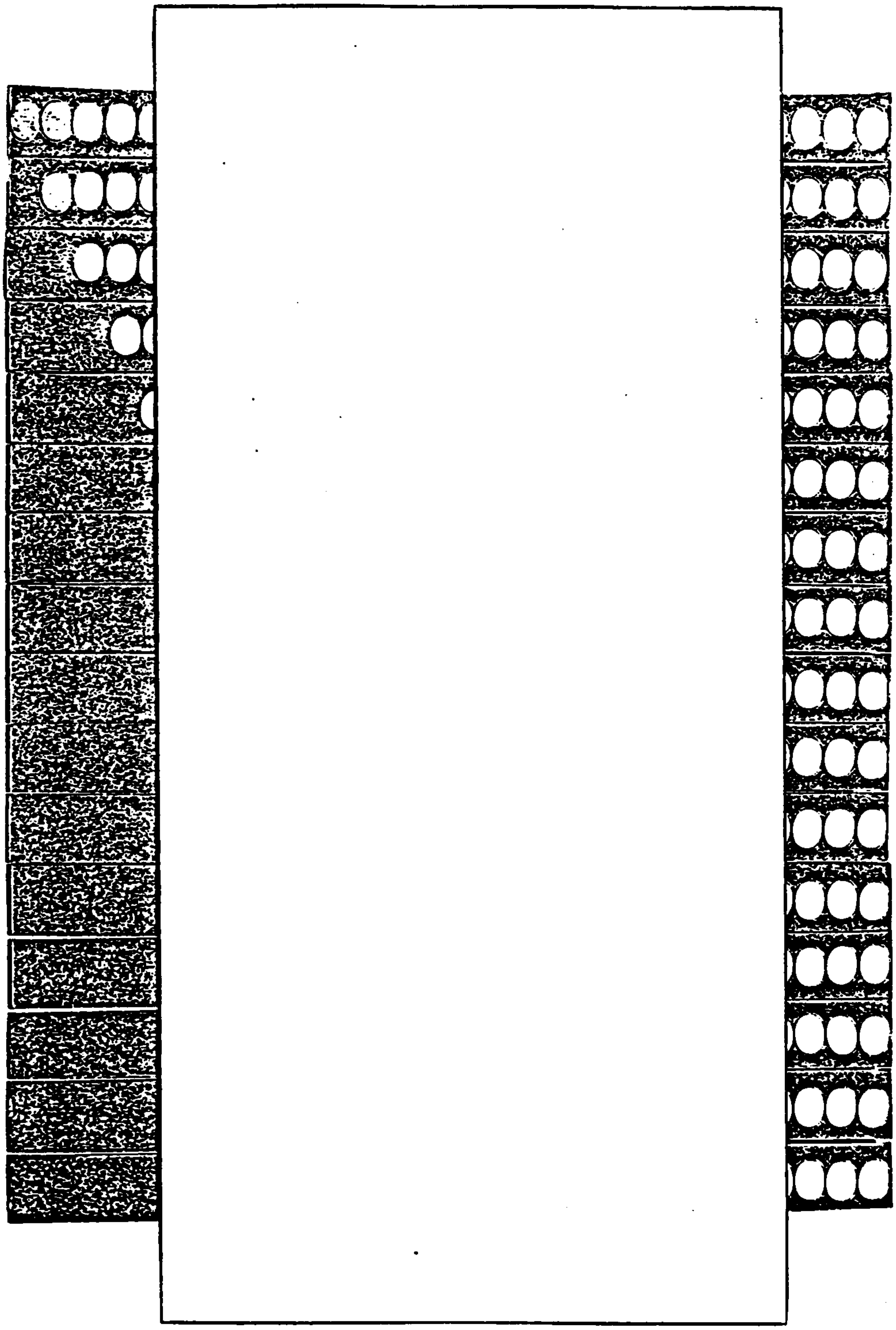

Figure 11. Ordered Dot Strips as They Appear after Being Covered

INSTRUMENT #5

THE PLACE VALUE DIAGNOSTIC TEST

THE PRINCIPLE TO BE TESTED:

THE NUMERALS 0, 1, 2,...,9 DO NOT ALWAYS CARRY THE SAME VALUE IN OUR NUMBER SYSTEM. THEIR POSITION IN A NUMBER DETERMINES THEIR VALUE. EXAMPLE, IN THE NUMBER 999, THE RIGHT-HAND 9 HAS VALUE OF 9, THE MIDDLE 9 HAS A VALUE OF 90 (OR 9 GROUPS OF TEN), AND THE LEFT-HAND 9 HAS A VALUE OF 900 (OR 9 GROUPS OF ONE HUNDRED)

The Diagnostic Test

Description of the Procedure

For the first part of this test, measuring the transitional level of initial understanding, two types of questions will be posed. The first three questions involve placing numerals in their sequential positions in a ten by twenty grid of boxes on a magnetic number board. The board is presented with 20 numerals between 20 and 219 already placed in their correct positions as shown in Figure 12. The sequence of numerals is started at 20 instead of 0 in order to make rote responses related to classroom tasks less automatic. Since most classroom number grid tasks use the numerals from 1 to 100, rather than from 0 to 99, etc., the arrangement on the grid in this task is apt to look somewhat different, to a rote responder, in more than one way.

The first three trials at the transitional level, then, require the child to place the numerals 67, 116, and 202 on the magnetic number board. The numerals are presented to the subject <u>without</u> being named verbally by the examiner, so the subject has no verbal cue to aid in their placement.

For the second part of the testing measuring the transitional level of initial understanding, two types of questions will be posed. First, a card is shown with the number 24 in the left upper corner. To the right of the 24, two rows of 10 and one row of 4 heart shaped red cinnamon candies are shown. The circles are intended to suggest, but not define, coins. See Figure 13 for a facsimile of the card. The subject will have all the elements of the card pointed out to him, and will then be asked why the digits 2 and 4 are used together to symbolize 24 items when, in fact, one combines two coins and four coins, one gets only six coins--not twenty-four coins. The subject's explanation is patiently solicited and recorded.

The second type of question, and the last trial for the transition level section of the test, is a request to solve an adition problem, with regrouping required only in the ones position. In pilot tests, it was found that children succeeded with this type of regrouping problem with a much higher frequency than with a straightforward regrouping problem involving the subtraction process. Theoretically this can plausibly be explained by the fact that in an addition problem, the regrouping aspect of the process is necessitated as a <u>result</u> of the simple addition process; while in a subtraction problem, the regrouping process must be employed at the <u>outset</u> of the process in order to make simple subtraction possible. Furthermore, the children must modify their perception of the previously learned precept that "you can't subtract a larger quantity from a smaller quantity" to accommodate the possibility of regrouping, where higher denominations are available. The addition problem is presented on Card One.Five, shown in the

Materials section which follows:

As in all the diagnostic tests, four correct responses out of ten trials are needed to obtain a success rating for the transitional level of understanding of the place value principle.

The second half of this diagnostic test has two different kinds of trials, also. The first two trials are similar in design to the fourth trial of the transitional section, but are more abstract in nature, since money is used to model the place value aspect of numerals. Two more cards, with coin representations of 24 and 124, are presented to the subject. He is asked if each would be a good device for helping a child understand what the 1 means, what the 2 means, and what the 4 means when these numerals are written together to represent one quantity. He is asked to explain why he thinks (or does not think) this example would be a good representation. If the subject thinks the representation is <u>not</u> satisfactory, he is asked if he would suggest some changes to make it a better representation.

These two cards, shown in Figures 14 and 15, are constructed as follows:

<u>card two</u> has the number 24 at the top with two dimes lined up vertically under the 2 and 4 dimes lined up vertically under the 4; and

<u>card three</u> has a 124 at the top, with one dime under the 1, two nickels under the 2 and four pennies under the 4.

In each of these trials, a success rating is awarded if the subject correctly identifies the ones, tens and hundreds values of the appropriate numerals, suggesting appropriate substitutions to correct the coin examples, as needed. Since the line of questioning is intentionally (to avoid rote responses devoid of understanding) quite unlike the typical place value testing questions used in the classroom, both cards will be kept on the table until both have been presented and responded to. The subject will be allowed to add to either of his explanations until both have been explained (or not explained) to the subject's satisfaction. In the pilot testing, this was found to be necessary since the coin examples seemed to help quite a few subjects figure out what was being sought, after having been thrown quite "off balance" by the initial line of questioning. Accepting additions and/or modifications to explanations throughout the duration of this part of the task administration should improve, rather than impair, the reliability and validity of the score obtained since the measurement goal is to determine whether or not the subject is capable of explaining the place value aspect of numerals as used in our number system and the timing of the responses is essentially irreleveant (within the constraints of the testing situation). In essence, if a subject is able to explain, refusing to accept a "late" but valid explanation would result in a false negative score and would therefore impair the validity and reliability of the instrument. On the other hand, obtaining a false positive score by accepting "late" explanations is quite unlikely because of the novelty of the examples displayed, and because the examiner protocol provides for no verbal clues during the procedure.

The last three trials test the subject's abilities at solving vertical addition and subtraction problems which require regrouping: complicated by internal zeros and successive regrouping. The third and fourth trials require the subject to solve one addition and one subtraction problem, while the last trial presents a subtraction problem already solved (incorrectly) and requires the subject to check the solution to see if it is correct. In this last trial, an explanation is required. The subject is asked to explain why the solution is incorrect, if the subject does find it incorrect. If the subject claims it is correct, he is asked to demonstrate the procedure which must have been used to solve the problem in the hope that he may discover the error. In the event that the subject does not discover the error in this last trial, but has answered the previous four questions correctly, an explanation will be sought for the procedure used in solving the problem for Trial 4.

For this section of the test, an explanation will be scored as successful if it explains the regrouping process in the problem including firstly, the fact that an adjustment in the quantity represented for a particular denomination was necessary, secondly, a reason why the adjustment was necessary, and thirdly, an assertion that the total value of the quantity represented was not altered by the adjustments made. The reason for the adjustment in subtraction is, of course, not having enough elements of a particular denomination to subtract from, while the reason for regrouping in addition would be because more than 9 elements cannot be represented in any one denomination.

Materials to be Used

1. Three 5 x 8 index cards set up as depicted in Figures 13, 14, and 15 and then laminated (or covered with clear contact paper). The 1 inch high numerals can be used. The hearts were stenciled and colored

with a marking pen. The coins were real ones positioned with double faced scotch tape before laminating.

2. A magnetic number board lined with a 10 x 20 grid as shown in Figure 12 is needed with numeral cards for the numerals 20, 22, 26, 29, 33, 38, 41, 44, 47, 74, 80, 89, 92, 137, 163, 166, 171, 178, 185, 67, 116 and 202. A 17" x 20" stove and counter mat with a steel surface works well. A felt marking pen with permanent ink can be used to make the grid lines. The numeral cards can be cut out of posterboard, stenciled with the numerals, and then laminated. Strips of stick-on plastic magnets can be bought in many sewing notions departments and cut into small pieces with utility scissors to attach to each numeral card. The numeral cards must be sized to fit the grid blocks, about 1 1/2" x 1" for the 17" x 20" steel mat described above. (A much more economical, but less durable, number board can be made on a piece of poster board covered with clear contact and then written on directly with an overhead projector pen, which wipes off easily.)

3. Four 5 x 8 cards on which the problems below are depicted before the cards are laminated (or covered with clear contact). Stenciled numerals or the plastic stick-on can be used. One problem is shown on each card.

55	52	206	201
+26	−27	+97	−196
			105
Card One.Five	Card Four	Card Five	Card Six

4. An overhead projector pen such as the Vis-a-Vis, Crusader or Staedtler is needed if the subject is to write on laminated cards which will be wiped off. These pens wipe off easily and cleanly with a damp sponge.

Explanation of the Subject

Say, "In this activity, I'm going to be asking you to do a lot of explaining. You know, in the beginning, I told you that we were doing all these activities to help teachers find better ways of explaining math ideas to kids who are having trouble with their math. So, in this activity, I'm going to be writing down just how you would explain some math things that are kind of confusing to lots of kids in second and third grade. I know explaining is often harder than just doing something so take your time and think about how you would explain these to a second grader."

Presentation of the Problem

Present the number board with the twenty given numbers already placed correctly and position it facing the subject. Explain, "This number board has boxes for two hundred numbers, but we started with 20 up here in the left hand corner, and then you count...(point at the correct boxes) 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32,...and so on, so we can put numbers up a little above 200 on the board. Of course, each numeral has to go in just the right box so they'll all count in the right order if we fill them all in."

"Now I'm going to give you some numerals, and I want you to put them up there on the board in the right box, so they'd all be in the right counting order if we filled them all in." "Here's the first one." Give the subject the numeral card with 67 on it. When the numeral card has been placed to the subject's satisfaction, record its placement and proceed to the next one.

Plc. Val. Trans. 2

Present the numeral card with 116 on it for the subject to place on the board. When finished, record the response and proceed to the next one.

Plc. Val. Trans. 3

Give the subject the numeral card with 202 on it and have him/her place it on the number board in the correct box, record the response and then proceed to the next task.

Plc. Val. Trans. 4

Present the card, facing the subject, which has the numerals 24 with 24 hearts and 6 coins pictured on it. Say, "You know, when we write the numeral 2 by itself, we mean we have two things--like these 2 coins here (pointing)--and when we write the numeral 4 by itself, we are talking about 4 things--like these 4 coins here (pointing)--; but when we write the numerals 2 and 4 together, we are thinking about 24 things, like the 24 pieces of cinnamon candy shown over here (pointing)--not 2 things and 4 things like the coins!" Q--"But, just what does the numeral 2 mean and what does the numeral 4 mean when we put them together to make us think of 24 things instead of 6 things?...(Pause)..."Why do you think we use a 2 and a 4 together to mean 24 things? (Be patient for at least one full minute, then repeat the question from Q again, if necesary to elicit a response.) Record Response Verbatim.

Continue with the remaining questions when the subject seems to be finished explaining, but keep each question card visible over at the side of the table in case the subject decides to go back and explain further before the Plc. Val. questions are all completed. Record any additional explanation opposite the appropriate question when it is given.

Plc. Val. Trans. 5

Present card one.five facing the subject and say, "This is an addition problem like you've probably done in class before. I'd like you to show me how you would solve the problem, by doing it right on this card." Give the subject the overhead projector pen, and record the response when the problem is completed.

Plc. Val. Comp. 1

Present the card with the number 24, two dimes under the 2 and four dimes under the 4. This is card two. Say, "This card has coins on it to show what the numerals 2 and 4 mean when we write them together to show 24. Do you think this is a good way to help a second grader understand what the 2 means and what the 4 means in 24...?" "Why do you think so?" Record Responses. Then ask, "Do you think different coins on the card would be better for explaining the 2 and 4 in 24?"..."
Why do you think so?" Record Responses.

Plc. Val. Comp. 2

Present card three facing the subject and say, "On this card, we used different coins for each numeral. We have one dime under the 1, two nickels under the 2, and four pennies under the 4. Do you think this could be a good example to help explain how much the 1 counts, how much the 2 counts and how much the 4 counts in one hundred and twenty-four?"..."Why do you think so?"...Record Responses.

Plc. Val. Comp. 3

Present card four facing the subject and give him the special pen provided for the test, saying, "On this card, I want to see how you would figure out the answer to this subtraction problem, just as we did before with the addition problem. You can use the pen again to show me how you would do it."

Plc. Val. Comp. 4

Present card five facing the subject and say, "This is another addition problem like the one before. I'd like you to show me how you would solve the problem, right on the card again.

Plc. Val. Comp. 5

Present card six to the subject and say, "You don't need to do this problem, because another third grader already did it for me. What I want you to do is check his work and see if it's done correctly." Pause, until the subject gives a response. Record the Response.

If the subject says the problem is not correct ask him to explain why and Record the Responses.

If the subject says the problem is correct ask him to explain how he thinks the third grader did the problem and Record his Response.

Continue to pursue the "why and how" questions until the subject has given his best explanation, but not to the point of making the subject uncomfortable if he has no better explanation. The explanation needs to include the point that regrouping was necessary and why it was necessary plus an explanation of how it would be done correctly so as not to change the actual value of the total number represented.

The subject's justification must explain that one hundred was exchanged for ten tens and one ten for ten ones. "Crossing out a one and putting a ten here, then crossing out the ten and making it a nine because you need a ten here" does not justify the equivalent values of one hundred and ten tens, etc. The place values of one, ten, and one hundred must be included appropriately in the explanation. Otherwise, it is not complete or satisfactory.

If the subject has answered four out of the five Comp. questions correctly, but did not include all the essential elements in the solicited explanation, bring out the subject's solution to the Comp. 4 problem and ask him/her to explain the procedure used to solve the problem and why it was done that way.

If the subject's explanation does not include a complete explanation as outlined before, pursue his explanation further with such questions--"Why did you cross out the 2 and make it a 1?...or..."Won't you get the wrong answer, changing that number?...or "How did you subtract 7 from 6?" etc. Questions, of course, must be tailored to the procedure demonstrated, and should not give undue clues about the justification needed. If the subject obviously is not capable of rendering the appropriate explanation, do not pursue the questioning further. However, do continue the questioning sufficiently to elicit a satisfactory explanation if it is to be forthcoming. BE CAREFUL not to "put words in the subject's mouth" or to "drag out" an explanation through prompting with clues. Record Exact Configurations of the solved Problem Explanation.

Examiner ______________________

Place Value Score Sheet

Subject's Number ____________ Date and Time ______________

Place Value

Trans. 1 (numeral 67) placed in square ____________

Trans. 2 (numeral 116) placed in square ____________

Trans. 3 (numeral 202) placed in square ____________

Trans. 4 (card one) Explanation ____________

Incorrect Correct

Trans. 5 (card one.five)

$$\begin{array}{r} 55 \\ \underline{+26} \end{array}$$

Comp. 1 (card two) Explanation ____________

Incorrect Correct

Examiner ______________________

Place Value Score Sheet (continued)

Subject's Number ______________

Comp. 2 (card three) Explanation ______________________

____

____

____

Incorrect Correct

Comp. 3 (card four)
$$\begin{array}{r} 52 \\ -27 \\ \hline \end{array}$$

Comp. 4 (card five)
$$\begin{array}{r} 206 \\ +97 \\ \hline \end{array}$$

Comp. 5 (card six)
$$\begin{array}{r} 201 \\ -196 \\ \hline 105 \end{array}$$
if changed, record
$$\begin{array}{r} 201 \\ -196 \\ \hline \end{array}$$

Explanation ______________________________________

____

____

____

Comp. 4 Explanation (if needed) ______________________

____

____

____

Incorrect Correct

20		22				26			29
			33					38	
	41			44			47		
				74					
80					85				89
		92							
							137		
			163			166			
	171							178	
					185				

Figure 12. Magnetic Number Board with Initial Placement of Numbers

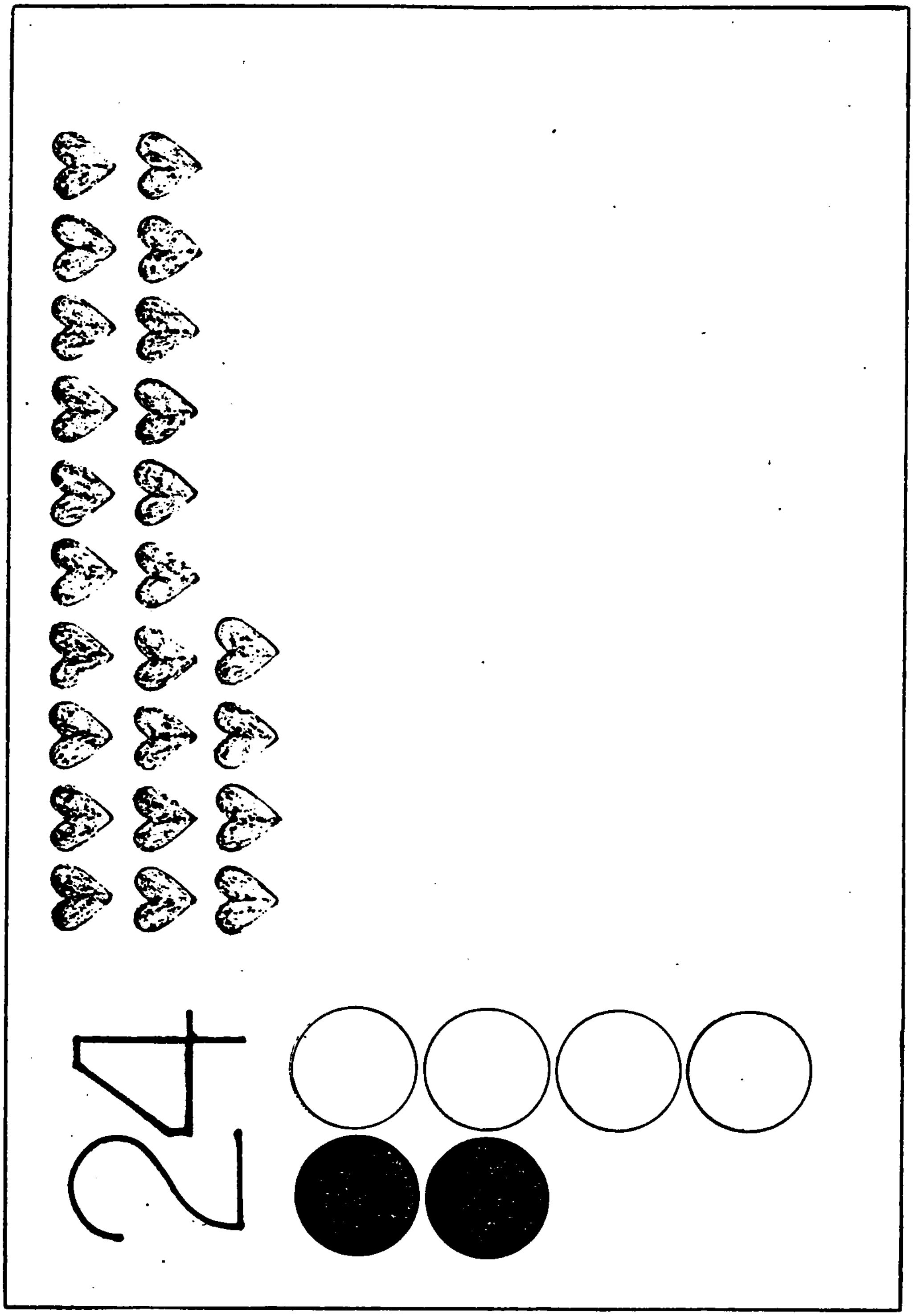

Figure 13. Card One for the Trans. 5 Task

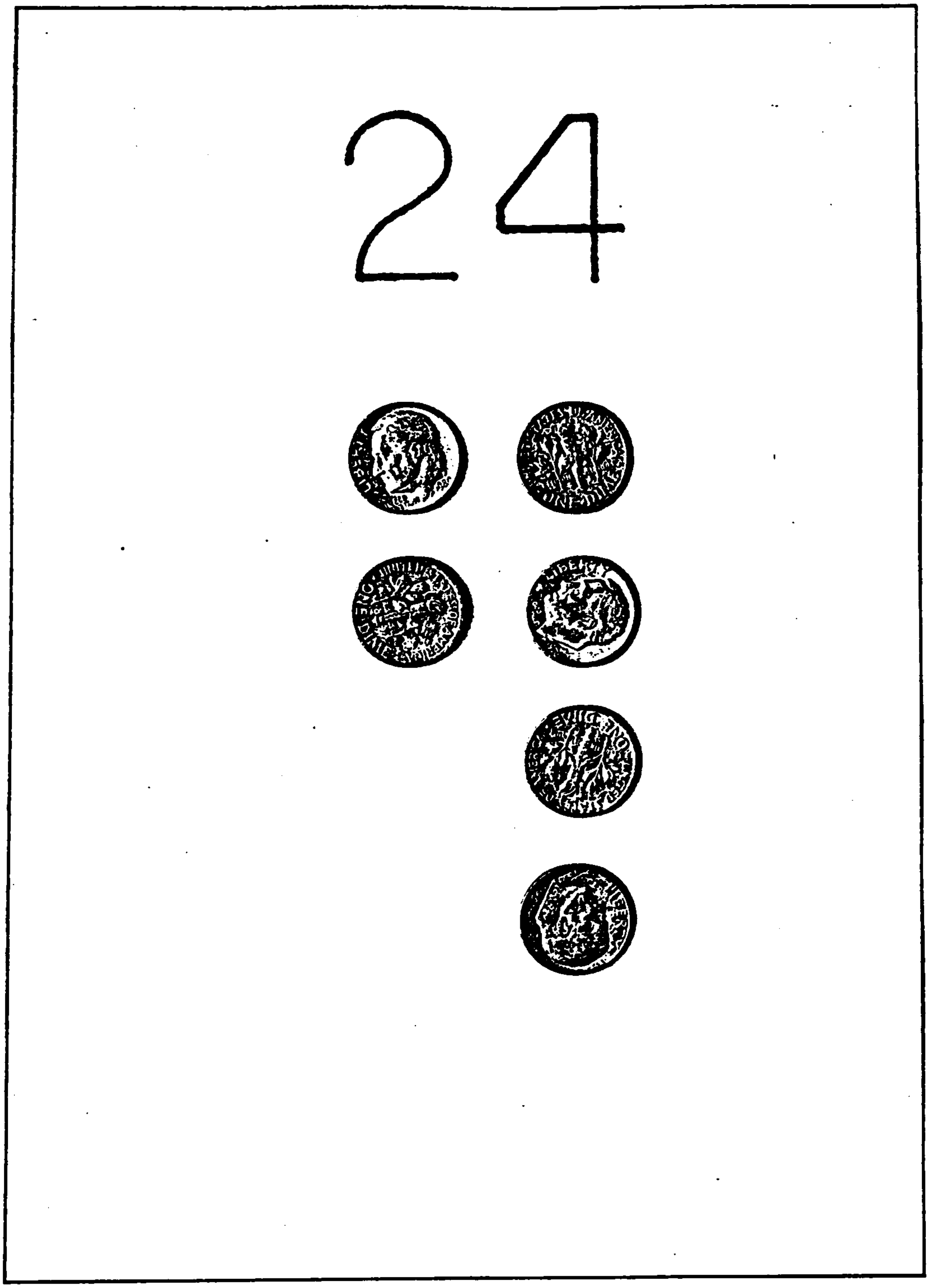

Figure 14. Card Two for the Comp. 1 Trial

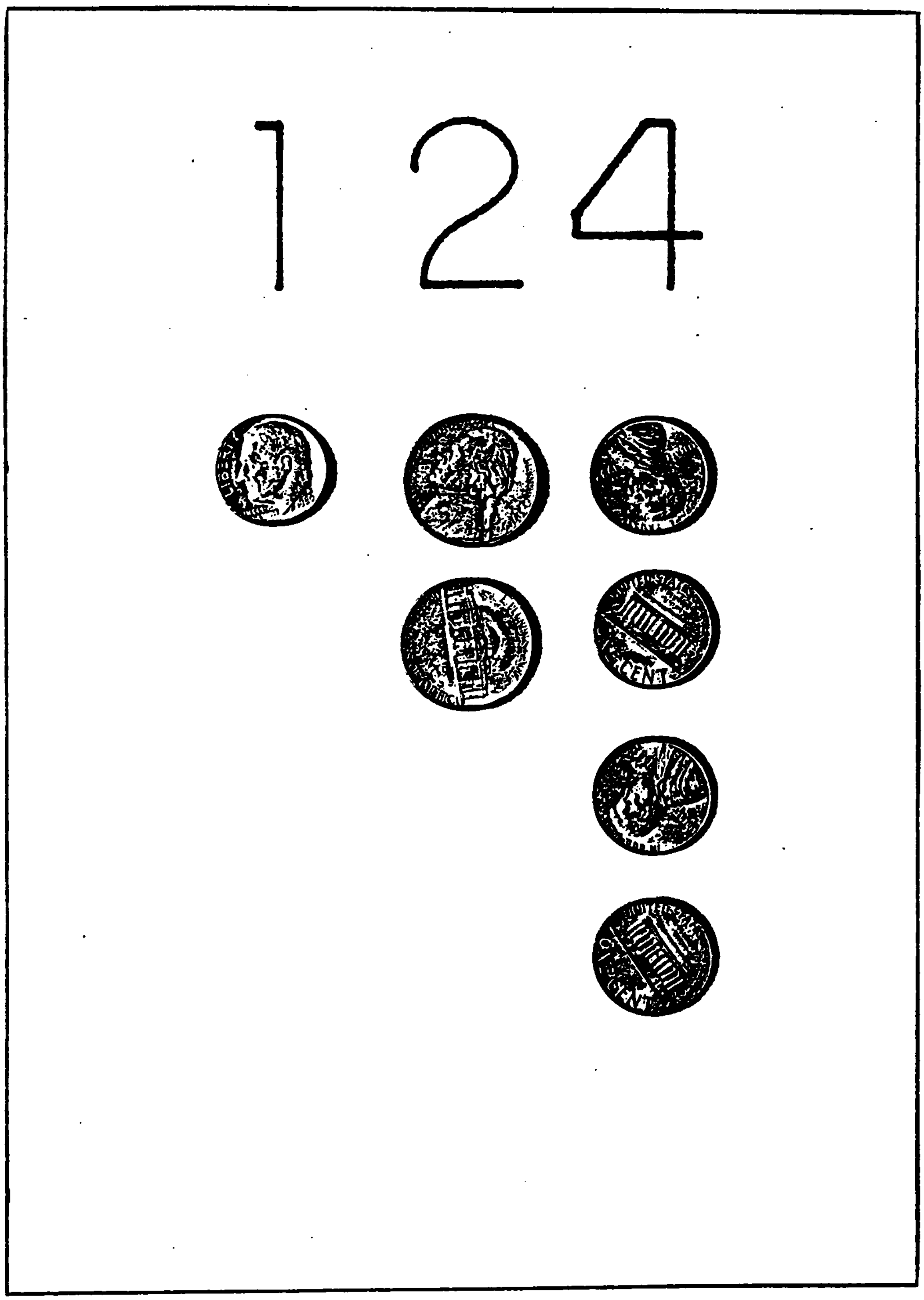

Figure 15. Card Three for the Comp. 2 Trial

$$\begin{array}{r} 55 \\ +\,26 \\ \hline \end{array} \qquad \begin{array}{r} 52 \\ -\,27 \\ \hline \end{array}$$

$$\begin{array}{r} 206 \\ +\;\;97 \\ \hline \end{array} \qquad \begin{array}{r} 201 \\ -\,196 \\ \hline 105 \end{array}$$

Figure 16. Problems for the Trans. 5, Comp. 3, Comp. 4 and Comp. 5 Tasks

INSTRUMENT #16

THE CLOCK READING DIAGNOSTIC TEST

THE PRINCIPLE TO BE TESTED;

THE HANDS TELL WHAT IS BEING COUNTED AND THE VALUES FOR EACH NUMERAL ON THE CLOCK. FOR THE SHORT-HAND, HOURS ARE BEING COUNTED AND EACH NUMERAL COUNTS ONE HOUR. FOR THE LONG-HAND, MINUTES ARE BEING COUNTED AND EACH NUMERAL COUNTS FIVE MINUTES WITH THE SHORT LINES BETWEEN THE NUMERALS MARKING THE ONE MINUTE INTERVALS

The Diagnostic Test

Description of the Procedure

The large Judy Clock™ from Silver Burdett is used to set up times for the subject to read off. For the transitional level, the subject will be required to read three times where the minutes are an exact multiple of five and to set up two such times himself. Either the minutes <u>from</u> the previous hour or the minutes <u>before</u> the next hour are acceptable.

For the complete competence level, the subject will be asked to read off three different times to the exact minute where the minutes are not exact multiples of 5 and then to reset the clock to a given number of hours and minutes <u>before</u> and then <u>after</u> a given setting for the last two trials. For time past the half hour, the subject will be asked to give it in minutes after the previous hour <u>and</u> in minutes before the next hour. The subject will also be asked to explain his system for reading the exact time on the clock.

The explanation of the procedure for determining the time indicated on the clock is considered acceptable if:

1) the short hand is identified as the indicator for hours,

2) the interval between any two numerals on the clock is given a value of one hour when using the short hand indicator (which may be explained by assigning 1 hour, 2 hour, and 3 hour values to the numerals on the clock),

3) the long hand is identified as the indicator for minutes,

4) the interval between any two numerals on the clock is given a value of 5 minutes when using the long hand indicator (may explain numeral values as 5 min., 10 min., 15 min., etc.),

5) the number of minutes <u>after</u> the hour is determined by properly counting clockwise from 12 (as the zero point), and

6) the number of minutes <u>before</u> the hour is determined by properly counting counterclockwise from the 12 (as the 0 point).

The usual criterion of four correct responses out of five trials plus an acceptable explanation will determine a success score.

<u>Materials to be Needed:</u>

1. A clock with movable hand and with clearly placed arabic numerals with five equal intervals delineated with four short lines between each of the twelve hour numbers. The hands should be clearly distinguishable in length. The Silver Burdett Judy Clock™ is an excellent one. Other educational supply models can be used or a poster board model can be made. It should be at least nine inches in diameter.

2. A list of the test times to be set for each trial should be at hand when testing. The times are written out in English words so that they cannot be read off too easily by the subject who prefers to read your list instead of the clock.

<u>Clk Rdg Trials</u>

Trial 1 - three thirty

Trial 2 - four fifteen

Trial 3 - four forty-five

Trial 4 - five twenty-five

Trial 5 - seven forty

Trial 6 - seven thirty-eight

Trial 7 - ten eighteen

Trial 8 - eleven forty-four

Trial 9 - <u>from</u> twelve-eight <u>to</u> forty-one minutes <u>later</u>

Trial 10 - <u>from</u> one twenty-two <u>to</u> twenty-five minutes <u>earlier</u>

<u>Explanation to the Subject</u>

Say, "We've noticed that reading clocks is really hard for many second and third graders, so we would like to see how you go about reading some times off the clock. Afterwards, I'll ask you to explain your system for reading the clock times. Be sure to read each time carefully so I can see how your system works."

<u>Presentation of the Trials</u>

Place the clock on the table facing the subject and say, "I'll set the times and you tell me what time it is. Okay?!"

Clk Rdg. Trials 1 through 5

Set each time as shown in the Materials List and record the subject's response for the appropriate trial number on his score sheet. If the subject should give you an inexact time, such as "a little after 3 o'clock," ask the sujbect if he "can tell you the time more exactly."

Clk Trials 6-8

Continue as before setting the trial time given and asking the subject to give you the time "to the exact minute" if his reading is an approximation and not just the wrong time.

However, for each trial, if the subject gives the time after the previous hour, ask, "How would you say that as the number of minutes before the next hour? That's time 'til the hour." Record Responses.

Clk Rdg Trials 9 and 10

For each of these trials, set the clock at the time given and then ask the subject to set the clock at the Trial prescribed number of minutes before or after the time given. Then ask the subject to read the time he set. Record the setting and Record his reading of this setting.

Clk Rdg Comp. Explanation

Reset the clock at eleven forty-four (Comp. 3 Trial), and say to the subject, "I'd like you to explain to me just how you figure out what the time before the hour is, explaining how you figure the hours and how you figure the minutes." Record Response. Then ask, "How do you figure the time out if I want the time after the hour? I need to know how you determine the hour and the minutes after the hour."

Record Response.

Scoring of Responses

Trials 1 and 2 are warm-up items and are not scored. Trials 3, 4, and 5 receive one point each for correct responses. Trials 6, 7, and 8 receive up to two points, one for each of the first two correct readings out of the six readings. Trials 6, 7, and 8 also receive up to three additional points, one for each time setting for which both readings are correct. Trials 9 and 10 receive one point for each correct trial.

Examiner ______________________________

Clock Reading Score Sheet

Subject's Number ______________ Date and Time ____________________

Clk Rdg	Reading Given
Trial 1	______________
Trial 2	______________
Trial 3	______________
Trial 4	______________
Trial 5	______________

Clk Rdg	Time Before Hour	Time After Hour
Trial 6	______________	______________
Trial 7	______________	______________
Trial 8	______________	______________

Clk Rdg	Time Set	Reading Given
Trial 9	______________	______________
Trial 10	______________	______________

Clk Rdg Comp. Explanation

Before Hour: ____

____

____

After Hour: ____

____

____

Correct Incorrect

Appendix B

Pilot Study Forms

NAME ______________________ NUMBER 2309

# RIGHT :	Response by Question Type / Concept Totals	INTERMEDIATE LVL 0	1	2	3	4	5	COMPETENCE LEVEL 6	7	8	9	10	E
ADDITION	Story Problems	X	X	X	X	0	X						
	Class Inclusion	W						X	X				
SUBTRACT	Picture Parts & Whole									X	0	X	
	Explanation												X
	ADD-SUBT TOTAL	0	1	2	3	4	5	6	7	8	9	10	E
COUNTING	On Request: By 2s		X	X									
	By 5s & 10s				X	X	X						
BY	Volitional: By 2s & 4s							0	X				
	By 5s & 10s									X	X	X	
GROUPS	Explanation												X
	CNT BY GROUPS TOTAL	0	1	2	3	4	5	6	7	8	9	10	E
ONE	Cash pnny,nckl,dime		X	X									
	Given p,n,d,qrtr,dllr				X	X	0						
EQUAL TO	Given Am't Made 2 Ways							X	X	X	X	0	
	Explanation												X
MANY	ONE = MANY TOTAL	0	1	2	3	4	5	6	7	8	9	10	E
N-MORE	From +/-1 to +/-3		X	X	X	X	X						
&	From +/-5 to +/-8							X	X	X	X	X	
N-LESS	Explanation												X
	N MORE & LESS TOTAL	0	1	2	3	4	5	6	7	8	9	10	E
PLACE	On # Board (67,116,202)		X	X	X								
	Addition with Regrping					X							
VALUE	Recognition of 10s & 1s						0						
	Model of 100s, 10s, 1s							X	X				
	Add & Subt with Regrpng									X	0	X	
	Explanation												X
	PLACE VALUE TOTAL	0	1	2	3	4	5	6	7	8	9	10	E
CLOCK	Multiple of 5 Minutes		X	X	X								
	To Exact Min. - 1 way					X	X						
READING	To Exact Min. - 2 ways							X	0	0			
	+ or - Exact Minutes										0	0	
	Explanation												0
	CLOCK READING TOTAL	0	1	2	3	4	5	6	7	8	9	10	E

Figure 17. Profile Sheet for Reporting Subject Responses (Followed by Explanation for Interpretation)

Interpretation of the Profile Sheet

INTERPRETATION OF THE PROFILE SHEET

Each of the six concepts are listed by title down the left-hand side of the profile page:

ADDITION-SUBTRACTION: the generalizations that the addition process is PARTa + PARTb = WHOLE while the subtraction process is WHOLE - PARTc = PARTd. (Typical unsatisfactory rules tend to "count all the pieces" for addition or "take away the smaller number from the larger" for subtraction. These latter rules don't work for problems such as N + 3 = 7 and N - 6 = 3 .)

COUNTING BY GROUPS: Quantities can be counted more quickly and with less chance for error by counting them in naturally occuring or arranged groups rather than one-by-one counting. (e.g. pairs of shoes, wheels on cars, fingers on gloves)

ONE EQUAL TO MANY: it is possible for one thing to have a quantitative value equivalent to many of something else. (e.g. 1 dime = 10 pennies)

N-MORE & N-LESS: N more than any number is represented by the N^{th} number above it in the counting sequence. (e.g. 6 more than 10 is shown by the 6^{th} number above 10 in the counting sequence.)

PLACE VALUE: all digits in a number are not counting just 1 item; the position of each digit in the number determines the size of the groups being counted. (e.g. the middle digit in 124 is counting 2 groups of 10 items, etc.)

CLOCK READING: there is a complex set of rules using the numerals and the hands to determine time before the hour and after the hour as shown on the clock.

If a circle is blacked in or checked on your child's profile sheet, then your child answered the corresponding question correctly. The W on Class Inclusion represents a warm-up question which does not add to the score, but is used to familiarize the child with that type of question. In general, the number of circles in each row indicates the number of questions of that type, and those blacked in were answered correctly.

For each concept, the first 5 questions are the Intermediate level questions and the second 5 questions are the Competence level questions which are logically more difficult. Thus, for Add-Subt, the 5 Intermediate level questions are all story problems. In all the other concept sets except Place Value, the first 2 types of problems (listed in the first 2 rows of their box) are for the Intermediate level. For the Place Value concept, 3 different problem types are included in the lower level questions.

The last row in each Concept box summarizes the child's performance for that concept by indicating the total number of questions answered correctly. Thus the profile:

CNT BY GRPS TOTAL 0 1 2 3 4 5 6 7 8 9 10 E

would indicate that 8 out of 10 questions were answered correctly and the explanation was satisfactory, also.

NAME ______________________ NUMBER __________ GRADE __________

CONCEPT / DESCRIPTION	LITTLE OR NO CORRECT PERFORMANCE	SOME CORRECT PERFORMANCE	DEPENDABLE CORRECT PERFORMANCE AND APPLIES WHEN APPROPRIATE
ADDITION - SUBTRACTION PROCESS: Performs the addition and subtraction processes correctly, distinguishes correctly between the two processes in a problem solving situation, and notices when the addition or subtraction procedure has been performed incorrectly.			
COUNTING BY GROUPS: Counts correctly by 2's, 5's, & 10's, up to 50 and uses this ability for speed and accuracy in suitable situations.			
ONE EQUAL TO MANY: Counts money (pennies, nickles, dimes quarters, & dollars) correctly, and can make up a required amount whenever possible using whatever coins are provided.			
N MORE AND N LESS Uses the From-To Method of adding and subtracting as in 6 + 3 is (from 6) 7,8, 9' so 12 + 3 is 3 more than 12 and 17 - 6 is 6 less than 17, etc.. Uses the generalization that N more than a number is N up in the counting sequence, and N less, is N down in the number sequence.			
PLACE VALUE Demonstrates understanding of the value of digits in the 100's, 10's & 1's places, locating numerals correctly in the counting sequence to 205, and in doing addition and subtraction with regrouping correctly. Also can explain why we use regrouping in a problem.			
CLOCK READING Reads time correctly to the nearest 5 minutes, reads time correctly to the exact minute, and identifies the correct time for N minutes more or less than the time indicated on the clock.			

Figure 18. Evaluation Form for Teacher Estimate of Subject's Levels of Understanding of the Six Principles Investigated.

Names and Spring 1982 Positions of Recognized

Authorities for the Face Validity Study

Dr. Donald S. Biskin,	Director of Research and Evaluation State Department of Mental Health, VA
Dr. Gerald Brazier,	Professor of Mathematics Education Pan American University
Dr. Larry Brooks,[a]	Supervisor of Science and Mathematics Montgomery County, VA Public Schools
Dr. Edgar Edwards,[b]	Mathematics Supervisor State Department of Education, VA
Mrs. Chris Gilley,	Coordinator of Special Education Montgomery County, VA Public Schools
Mrs. Evelyn Graham,[a]	Supervisor of Mathematics Chesapeake, VA Public Schools
Dr. Carole Greenes,	Professor of Education Boston University
Dr. Vergie Kieth,	Math-Science Supervisor Pulaski County, VA Public Schools
Dr. Raymond C. Spaulding,	Professor of Mathematics Radford University
Dr. Robert G. Underhill,[a]	Professor of Education Kansas State University

[a]Reviewed the Addition-Subtraction, Counting By Groups, and One Equal to Many protocols.

[b]Reviewed the N More or Less, Place Value, and Clock Reading protocols.

The two page vita has been removed from the scanned document. Page 1 of 2

The two page vita has been removed from the scanned document. Page 2 of 2