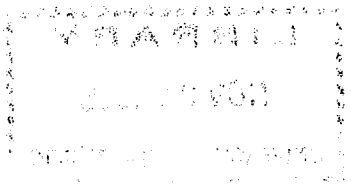
1965

The Relationship Between A Kindergarten Screening Test And Other Measures

Mary Page Mcguire
University of the Pacific



THE RELATIONSHIP BETWEEN A KINDERGARTEN SCREENING TEST AND OTHER MEASURES

A Dissertation

Presented to

the Faculty of the Graduate School

University of the Pacific

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Education

---

by

Mary Page McGuire

June 1965

This dissertation, written and submitted by

Mrs. Mary Page McGuire ,

is approved for recommendation to the Graduate Council, University of the Pacific.

Department Chairman or Dean:

J. Marc Jantzen

Dissertation Committee:

Jay H. Reeve, Chairman

David K. Brunen

Edward Bolen

Juanita G. Curtis

Dated May, 1985

TABLE OF CONTENTS

CHAPTER PAGE

LIST OF TABLES

TABLE PAGE

FIGURE PAGE

CHAPTER I

THE PROBLEM AND DEFINITION OF TERMS

Concern has been expressed in the past by kindergarten teachers, elementary principals, and curriculum consultants about kindergarten promotion. Their specific concern has been with the promotion of borderline cases. It is relatively easy to identify those children who are so very immature that it is quite obvious they need to repeat the kindergarten experience and those children who are clearly ready for a first grade program. However, every kindergarten class will contain some children about whom the teacher is doubtful as to whether or not they would be able to succeed in a first grade program.

At the present time the kindergarten teacher must rely solely on her own judgment, refer the child to a school psychologist, or administer group reading readiness tests.

Kindergarten teachers are reluctant to base passing or failing of questionable cases on subjective judgment alone and so refer to the school psychologist, when available. The psychological referral is time-consuming, limited in the number of children who can be seen, and costly to the school

district. The school psychologist must not only administer and interpret tests but also should observe the child several times in different situations. These factors alone will limit the number of children who can be seen.

Group reading readiness tests are available but do not appear to meet the particular needs in borderline cases. Such tests are too time-consuming for use with the one to five children about whose progress the teacher is particularly uncertain.

How, then, can we improve the kindergarten teacher's evaluation of borderline promotional cases?

An experimental instrument is needed to explore the possibility of providing the kindergarten teacher with an objective screening device which could be used in conjunction with teacher judgment for the purpose of decision-making. The writer proposed to develop such an instrument. This instrument would give information about visual-motor coordination, vocabulary and basic general concepts and could be easily administered. Such a test could provide kindergarten teachers with relatively objective data for evaluation of the slower child's readiness for first grade, and might also provide quantifiable information about the maturation level

of any child.

I. THE PROBLEM

Statement of the problem. Are there significant relationships among the experimental instrument, the Lee-Clark Reading Readiness Test, kindergarten teacher judgment, and first grade reading group placement?

Purpose of the study. There are three areas of maturation which appear to be critical in predicting first grade success: visual-motor coordination, vocabulary, and basic general concepts. This study is concerned with the development of an experimental instrument which will measure these three areas and which can be administered by a kindergarten teacher in one brief, individual session. This instrument would provide an additional means for the evaluation of pupil progress and maturation at the end of the kindergarten period.

The purpose of this study is one of exploratory validation. Concurrent validity was established by showing the relationship between the experimental instrument and the Lee-Clark Reading Readiness Test. The relationships among

the experimental instrument, the Lee-Clark Reading Readiness Test, kindergarten teacher judgment, and first grade post-readiness reading group placement determined the predictive validity. The statistics offered in this study are for descriptive rather than definitive purposes.

Delimitation of the study. The population included in this study was limited to kindergarten and first grade pupils in three Modesto elementary schools.

The data used were obtained by the investigator during the school years 1963-64 and 1964-65. The test instruments administered were the Lee-Clark Reading Readiness Test and the investigator's experimental instrument, called the Kindergarten Screening Test. Kindergarten teachers' evaluations of expected success in a first grade program were obtained in May, 1964, without prior knowledge of test scores. Actual reading group placement was obtained from first grade teachers during the months of October and November, 1964.

Hypotheses. When expressed in terms of the null hypothesis, this investigation is designed to test the hypotheses that:

1. There will be no significant relationship between pupil scores on the Kindergarten Screening Test and the Lee-Clark Reading Readiness Test.

2. There will be no significant relationship between pupil classification on the Kindergarten Screening Test and kindergarten teacher judgment.

3. There will be no significant relationship between pupil classification on the Kindergarten Screening Test and actual first grade reading group placement.

II. DEFINITION OF TERMS

Borderline cases. This term refers to those two or three children found in any regular kindergarten class about whom there is doubt as to possible first grade success. Children who obviously would not be successful are not considered as borderline.

Visual-motor coordination. Within the frame of reference of the study, this term refers to the ability of the child to look at a design and reproduce it.

Basic general concepts. This term is used to include those subject matter items and informational items of a general nature which are expected to be within the experiences of the average five- and six-year-old. These desired common

learnings may be the result of experiences within or outside of the kindergarten program but are expected by the first grade teacher to have been mastered and so to have become a part of the child's readiness for first grade.

Individualized reading. This term describes a reading program where a large number and variety of textbooks are used and children choose what they read. Book selections are made by the child, and instruction is conducted on a one to one basis. Grouping for reading experiences may be done on a short term basis.

III. ORGANIZATION OF THE CHAPTERS

Chapter II deals with pertinent literature in the areas of psychometric techniques, child growth and development, social forces, first grade curriculum materials, and other tests available.

The Pilot Project is discussed in Chapter III. Included in this chapter are the development of the experimental instrument, selection of the sample and revisions, and results of the administration of the instrument.

Chapter IV contains the experimental design of the

study which includes sample selection, instrument administration, data collection,and plans for data treatment.

Chapter V presents the data obtained in the experimental study and their interpretations.

The conclusions and recommendations indicated by this study constitute Chapter VI.

CHAPTER II

REVIEW OF THE LITERATURE

A survey of the literature pertinent to this study led quite naturally into specific topics within certain areas. The child growth and development point of view was explored generally, but specific emphasis was placed on research on the early school years. The emergent social conditions which might affect a child's beginning school experience were explored. Recent shifts or trends in curriculum which would influence the project were studied. Statistical treatment for the investigation was determined from authoritative texts and articles discussing certain techniques.

In reviewing literature for this exploratory study data relating to primary sources were lacking. Therefore the content of this chapter will be more conceptual than might be the case if more specific studies were available. An effort will be made to construct a frame of reference in the following discussion.

CHILD GROWTH AND DEVELOPMENT

The child growth and development point of view implies

that children grow and develop physically, emotionally, socially, and intellectually, each at his own rate of speed. At times this development is rapid and seemingly in spurts. At other times it is slow, appearing to come almost to a standstill or to reach a plateau. Each stage in a child's development brings with it specific problems, and each developmental step the child takes in infancy and early childhood leads to another. The child moves from creeping and crawling to standing and walking. He moves from infant babbling to talking and from the simple to the complex task.

Gesell, an early pioneer in the recognition of growth and development, and his staff at the Yale Clinic made a biographic-developmental study of the patterning of behavior from infancy through age ten. Gesell's area of investigation was designed to determine whether or not "a given behavior had an assignable status in a gradient of growth as indicated by converging evidence of the total data for all the children of all the ages."[1] Gesell's studies established the concept that every child has an individual growth pattern which is

---

[1]Arnold Gesell and Frances Ilg, The Child From Five to Ten (New York: Harper and Brothers, 1946), p. 3.

unique for him. Thus the developmental concept as related to changes in motor behavior was brought into focus. Gesell's longitudinal studies constituted a framework on which others might build.

To the concept of developmental growth Ruth Strang added the premise that a child's behavior is produced by forces both within and outside of himself. She conceived of each phase in a child's development as bringing with it special problems which needed to be solved. Strang believed that it is easier for the parent and more gratifying to the child, to help a child through each phase than to try to undo wrong moves or detrimental ways of coping with life's situations. Thus, although a child might grow and develop at his own rate of speed, resultant behavior could be controlled.[2]

Havighurst was concerned with the developmental task concept. He conceived of development as learning and maturing as the result of inner and outer pressure. Havighurst identified developmental tasks as being three in nature.

---

[2]Ruth Strang, An Introduction to Child Study (New York: The MacMillan Company, 1951).

> Some tasks arise mainly from physical maturation, such as learning to walk, learning to behave acceptably to the opposite sex in adolescence, and (for women) adjusting to the menopause in middle life. Other tasks, arising primarily from the cultural pressure of society, are learning to read, and learning to participate as a socially responsible citizen in society.
>
> There is a third source of developmental tasks--namely, the personal values and aspirations of the individual, which are a part of his personality, or self.[3]

Havighurst sums up the importance of developmental tasks to the educator with the statement, "When the body is ripe, and society requires, and the self is ready to achieve a certain task, the teachable moment has come."[4]

Kephart recognized the human organism as living in the most complex of all environments and the complex environment in which the child lives as leading to the most extreme demands for learning. He saw more demanded in the way of learning for the human organism than for any other living organism. Kephart described the situation as one wherein not only is behavior at a high level of achievement demanded of the organism but also one in which the organism has not

---

[3]Robert J. Havighurst, Human Development and Education (New York: Longmans, Green and Company, 1953), p. 4.

[4]Ibid., p. 5.

developed sufficiently for such adaptation. Such demands are therefore dual. Behavior is demanded on the one hand, while a complex learning activity is demanded as well. This underlying learning activity is so complex that no other species could accomplish it, let alone develop efficiency in it. Such learning is most difficult; it strains the capacity of the normal child.[5]

The complex learning activity described by Kephart is one which is expected in the classroom as routine. Curricular theory holds that the developmental level of each child must be recognized before anticipation of performance or learning. Behavior must be accepted as a function both of the characteristics of the person and of the characteristics of the situation.

During and immediately following World War II the educator began to recognize more clearly the implications of child growth and development. Learning was then perceived as more than an acquisition of factual materials. Huggett and Millard at Michigan State University represented this point

---

[5]Newell C. Kephart, The Slow Learner in the Classroom (Columbus, Ohio: Charles E. Merrill Books Incorporated, 1960), p. 6.

of view. It was asserted that acquisition of facts was not to be discarded, but rather emphasis should be placed upon those facts which would contribute to effective living in our civilization. Thus the acquisition of facts becomes a tool in the more fundamental goals of well-rounded physical development, health, mental growth, social adjustment, personality development, character formation, balanced living and emotional balance.[6]

Crow and Crow present the concept of child growth and development as a tracing of the patterns of development and adjustment experienced by the child as his potentialities are developed through experiences with physical, social, and cultural factors of influence. They treat the various aspects of physical, social, motor and emotional development as separate entities, as though the child could be segmented and analyzed as separate parts instead of an integrated whole. The dynamics of child development are sacrificed for operational precision and become presented to the educator as specifics, such as development of skill in written language

---

[6]Albert J. Huggett and Cecil V. Millard, Growth and Learning in the Elementary School (Boston: D. C. Heath and Company, 1946).

or the sequences of social development.[7]

Carrie Lou Goddard speaks for this century when she calls it the century of the child and sums up the problems of understanding child growth and development by stating that there are no magic words that can be used to help, no list of characteristics which can be learned, no one book that can be read to provide the necessary skills to help children to live the more abundant life. She feels that research and study have shown that there is no general body of knowledge applicable to every human being; that each person is unique, born into a particular environment, stimulated by specific events which happen to him only, and responding to the individuals who make up his society. To help youngsters to develop to their maximum potentialities is the problem which must be faced.[8]

SOCIAL FORCES

Outside of the school situation, major social forces

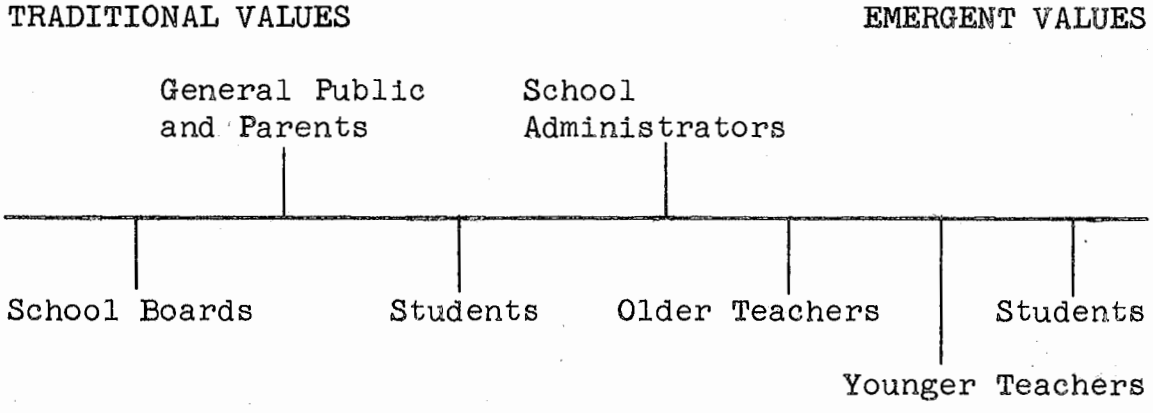
---

[7]Lester D. Crow and Alice Crow, Child Development and Adjustment (New York: The MacMillan Company, 1962), 491 pp.

[8]Carrie Lou Goddard, The Child and His Nurture (New York: Abingdon Press, 1962).

have been making their impact felt on the development of the child prior to his entering school. These forces have not only affected his actual physical and mental growth but also his anxiety level. His ability to participate successfully in a school situation has been determined at least partially by these social forces.

According to Spindler, American culture is undergoing a confused transformation. A major shift in value systems is occurring which he labels a shift from traditional to emergent value systems. This shift is quite evident as one studies the relationships between the school and the community and the placement of groups along a continuum. Spindler presents this placement of groups by use of the following diagram.



The difference in value systems illustrated by the

above groups would naturally lead to conflicts in ideas about education. These conflicts would of necessity directly affect the educational program as a whole and a parent's expectations concerning school success.[9]

The traditional values which may be clung to by some school boards, parents, and the general public would include a set of moral and spiritual values that are more frequently expressed in the form of maxims of early Americana, such as "honesty is the best policy" or "a stitch in time saves nine." Traditional values are thus expressed as a set of rules which can apply only as the rules are mutually understood and acceptable. In education a traditional value system stresses the need for a formalized educational system which emphasizes content and skills.

We move in a period of transition toward a new core of values which include a new interpretation of civil liberties, the belief that every individual is entitled to full personal development, a faith in human intelligence, and the right of

---

[9]George Spindler (ed.), Education and Culture: Anthropological Approaches (New York: Holt, Rinehart and Winston, 1963), pp. 136-39.

the individual to participate in the formulation of policies within which he will live and work. It becomes then the primary task of the school to study the principles comprising the core of the American culture to the same degree that the three R's or science are stressed. The school now finds itself in the position of having its feet rooted in traditional values and its head reaching toward the concept of education as being more than the mastery of information.

The shift in value systems appears to be an integral part of American civilization. The occurring events are described as "progress" and "advancement." These events demand mobility, adaptability, and the capacity for growth.[10]

Curriculum and the child's ability to achieve success in a school program are affected by the attitudes toward education which are held by his parents and peers. Not all groups in society are equally interested in schools or education. The American public school system is primarily designed to serve the huge middle class. Most children from the upper classes attend private schools, and so the

---

[10] Solon T. Kimball, "Those First School Years," The National Elementary Principal, XL (September, 1960), pp. 18-32.

influences of their parents on the schools are not particularly strong. Many lower-class parents present an uninterested attitude toward education, and so do not to any great degree affect the school program. So the public school finds itself in a dilemna. The school must be aware of middle-class values and standards of behavior but must also be aware of the values and behavior standards of lower-class families.[11]

Middle-class mobility may also be considered a factor affecting a child's ability level. The upwardly mobile middle-class male feels that he must be as well educated as possible and alert in order to stay ahead. He feels that his children must be prepared with the skills, knowledge, and manners that he believes necessary for successful competition in life. So it is that he believes success in school insures a successful career, and in turn he puts pressure on the youngster for a high level of achievement.[12]

---

[11] Robert H. Beck, Walter Cook, and Nolan C. Kearney, Curriculum in the Modern Elementary School (Englewood Cliffs, New Jersey: Prentice-Hall, 1960), p. 119.

[12] Ibid., pp. 123-24.

CURRICULUM

The kindergarten curriculum in the United States has emphasized various aspects of a child's growth and development at different periods of time. During the early part of the twentieth century, a great concern was evidenced in the physical development and welfare of the child. In the nineteen twenties and thirties the major emphasis was placed on the child's social development. The forties and fifties showed a definite concern with the emotional development of the child. The era of Sputnik has brought an increased interest in intellectual development. The cry for engineers and scientists has brought academic life into focus. This influence is felt on all curriculum levels.[13]

Many recent curriculum conferences indicate that changes in curriculum are not only numerous but rapid. Curricula must reflect the facets of modern life. Good schools are not obligated simply to reflect the world but to help the young person to live in it worthily, gracefully, and even,

---

[13]Josephine Foster and Neith Headley, Education in the Kindergarten (New York: American Book Company, 1959), p. 438.

hopefully, to beautify it.[14]

Apparently pioneering is needed in every subject taught in schools. The school can no longer give children a summary knowledge of any field. The challenge now is to analyze the characteristic structure of each field, the peculiarities of the discipline by which it was built, and the distinctive nature of each subject.[15]

The traditional subject matter fields offer new frontiers for exploration. Changes are taking place both in the scope of the subject in the elementary school and in its sequence. Changes in the teaching of mathematics and science appear to be the most startling to the layman, but changes are also occurring in the field of language arts.[16]

Some newly published mathematics textbooks present a curriculum based on a discovery approach to the patterns, relationships, and concepts embodied in the structure of

---

[14]A. Harry Passow (ed.), Curriculum Crossroads: A Report of a Curriculum Conference (New York: Bureau of Publications, Teachers College, Columbia University, 1962), pp. 76-86.

[15]Ibid., p. 9.

[16]Aubrey Haan, Elementary School Curriculum (Boston: Allyn and Bacon, 1961), pp. 190-272.

mathematics.[17] Others, such as the Greater Cleveland Mathematics Program,[18] concentrate on the problem-solving technique introduced by the S.M.S.G. program.[19]

In the new frontiers of the language arts field we find a re-emphasis on motivation for reading, an expansion in the understanding of the meaning of readiness, and a tendency to involve the family with readiness problems, reading disabilities, and reading incentive. We also find popularized the practice of individualized reading, altering the language arts program according to the school district it serves, extensive reading, and new rapid methods of teaching reading such as the experiment in Dallas with five-year-olds using colors for the various sounds in word identification.[20]

---

[17] Robert E. Eicholz et al., Elementary School Mathematics, Primer (Palo Alto, California: Addison-Wesley Publishing Company, 1963).

[18] Greater Cleveland Mathematics Program (Chicago: Science Research Associates, 1962).

[19] School Mathematics Study Group: A curriculum study sponsored by the National Science Foundation and conducted by a group of mathematicians, psychologists, and educators expressedly for the purpose of stimulating commercial publishers to produce new mathematics materials.

[20] Haan, loc. cit.

Reading plays a definite role in the kindergarten curriculum. Not only are experiences with phonics and with name and label reading an integral part of the program, but as Gertha Williams believes, this is a time for enlargement of personal experiences, an expansion of the child's environment. She advocates the giving of the child more real experiences so that when he is ready to widen his experience through reading about others' experiences, he will have a full and rich background upon which to build the new experiences gained through reading.[21]

Reading may be described as the process in which the individual brings meanings to a word in order to take meanings from it. Accumulated experiences thus bring deeper meanings into a word. Reading is interrelated with the other aspects of the language arts; it is a part of the sequential order of language development. A child first develops a language of understanding, then a language of communication, and finally the ability to read. A child's readiness for

---

[21]Gertha Williams, "The Kindergarten Teacher and Reading: Another Viewpoint," Childhood Education, October, 1963, pp. 77-78.

reading is a stage in his total language development sequence.[22]

PSYCHOMETRIC TECHNIQUES

The language content of a test is of prime concern when constructing test items. The language used must not only be precise but within the vocabulary level of the group to be tested.

Thorndike and Horn[23] both made studies on adult word usage in books and correspondence, but comparatively little has been done in research on the language patterns of children. In 1936 a federal grant was given to the University of Oklahoma for the study of the vocabulary of elementary school children. The research problem was to present the actual frequency of words, grade by grade, of many children, from all sections of the country. The words in the tabulation

---

[22] A Teachers Guide to Education in Early Childhood (Sacramento, California: State Department of Education, 1956), p. 319.

[23] Henry D. Rinsland (ed.), A Basic Vocabulary of Elementary School Children (New York: The MacMillan Company, 1945), p. 4.

were those used by children in their own writing. This extensive study included 353,874 words used in stories by first grade children.[24]

According to Charles T. Meyers, there appears to be a lag between practice and theory in the selection of items in test construction. Meyers asserts that this lag may be because the test constructor used validity as his main concern, and reliability is sought only as a means to the end of increased validity. Meyers was concerned with whether or not a selected set of average difficulty items would be more reliable and valid than a selected set of items half of which were considerably easier than the average. His experimentation did not show any difference in validity between peaked and u-shaped tests. The experiment did, however, support the theory that peaked tests tend to be more reliable than tests with other distributions of item difficulty.[25]

---

[24] Ibid., pp. 1-21.

[25] Charles T. Meyers, "The Relationship Between Item Difficulty and Test Validity and Reliability," Educational and Psychological Measurement, XXII (Autumn, 1962), pp. 565-71.

Guilford suggests one method to be used in the selection of items to be included on a test, the assessment of item-total correlation. This is a comparison of each item with the sum of the items. Those items which appear to be poorly correlated with the total are eliminated.[26] The biserial coefficient of correlation is designed for a situation such as this when one variable is reduced to two categories.[27]

Solomon agrees that most psychologists use the method described by Guilford. He, however, prefers the method of choosing a certain number of items, then correlating each item with the latent factor. The final choice of items would be based on the order of the items from the highest to the lowest. This procedure is one which is usually followed by psychologists, but since the latent factor is not always known, the psychologist must often resort to the use of a

---

[26]Joy P. Guilford, Psychometric Methods (New York: McGraw-Hill Book Company, Incorporated, 1954), pp. 417-43.

[27]Joy P. Guilford, Fundamental Statistics in Psychology and Education (New York: McGraw-Hill Book Company, Incorporated, 1956), p. 297.

manifest equivalent, the observed test score.[28]

Statisticians as a whole are in agreement concerning the definition, use, and types of reliability coefficients found in general use. English and English in their dictionary of psychological terms describe reliability as a generic term used to refer to several types of evidence. Reliability describes the extent to which a pupil would obtain similar results on a test when readministered, assuming that no additional learning or practice effects had occurred. Reliability, then, describes consistency of scores. There are primarily three types of reliability coefficients. The coefficient of internal consistency refers to an analysis of data obtained on a single trial of a test. The split-half method and analysis of variance are the most common examples of this method. The coefficient of equivalence refers to a correlation of scores from two equivalent test forms, administered at essentially the same time. The coefficient of stability refers to a correlation between a test and retest

---

[28]Herbert Solomon (ed.), Studies in Item Analysis and Prediction (Stanford, California: Stanford University Press, 1961), p. 9.

with some time period intervening.[29]

Another basic concept used by all test constructors is that of validity. English and English represent the statistician in their discussion of validity. They state grossly that a test is valid to the extent that it measures what it purports to measure. Content validity refers to how well the content of the test measures the subject matter or situation. Construct validity involves the psychological qualities a test purportedly measures. A common use of concurrent validity involves the correlation of the test score with other test scores. Predictive validity more often is considered to involve the relationship between test scores and actual criterion performance. Predictive validity refers to how well predictions made from test scores are related to data collected, often at a later time.[30]

Cronbach differs only slightly from English and English in his description of validity. He describes the criterion of predictive validity as the record of an outcome;

---

[29]Horace B. English and Ava C. English (ed.) A Comprehensive Dictionary of Psychological and Psychoanalytical Terms (New York: Longmans, Green and Company, 1958), p. 456.

[30]Ibid., pp. 574-76.

this record, or criterion, is compared to the prediction. Concurrent validity also employs empirical comparison. Two sources of information are obtained at very near the same time and compared. His interpretation of construct validity is so similar to that of English and English that very little further discussion is needed. Cronbach describes construct validation as an analysis of the meaning of test scores in terms of psychological concepts.[31]

Although test constructors concern themselves with both the reliability and validity of the measure, Terman, in his introduction to McNemar's discussion of the Revised Stanford-Binet, states that "the futile war of words regarding the validity of this or that intelligence test seems to have died down and validity is defined more and more in operational terms. A test tests what it tests, and the nature of the 'what' only becomes clear as the test is used and the results checked."[32]

---

[31]Lee Joseph Cronbach, Essentials of Psychological Testing (New York: Harper and Brothers, 1960), pp. 103-5.

[32]Quinn McNemar, The Revision of the Stanford-Binet Scale (New York: Houghton Mifflin Company, 1942), p. 13.

APPROPRIATE KINDERGARTEN TESTS

Casual observation might indicate a plethora of tests available. However, those appropriate for use with a kindergarten child are very limited in number. Some are appropriate for individual use only while others may be group administered. Some tests are unifactorial while others are multifactorial. Table I on the following page illustrates an evaluation of some of these tests according to use.

TABLE I

TESTS APPROPRIATE FOR USE WITH KINDERGARTEN PUPILS

Test	Function			
	Intelligence	Coordination	Reading Readiness	General Information
W.I.S.C.	X	X		X
Stanford-Binet	X	X		X
Van Alstyne Vocabulary	X			
C.T.M.M.	X	X		
Slosson Intelligence	X			X
Readiness for Formal Work		X	X	
Pintner-Cunningham	X			X
Ammons Quick Test	X			
Ammons Picture Vocabulary	X			
Easel Age Scale	X			
Goodenough	X	X		
Bender Motor Gestalt	X	X		
Kent E.G.Y. Scale				X
Lee-Clark Reading Readiness			X	

CHAPTER III

PILOT PROJECT

The pilot project consisted of three major steps:

1. Development of the experimental instrument.
2. Selection of the sample.
3. Revision of the instrument and results of administration.

I. THE EXPERIMENTAL INSTRUMENT

The three-part Kindergarten Screening Test was conceived, in general, as one which would require relatively few verbal responses, one which could be administered in a brief period of time, and one which could be used by a kindergarten teacher in decision-making about promotion. Development of such an instrument was undertaken by the investigator in this study.

Test description. Section 1, Visual-motor Coordination, consists of eight drawings which are arranged according to difficulty. Each drawing, or design, is presented to the child who is asked to reproduce it. The reproduction of the design appears to involve gross motor coordination,

eye-hand coordination, and form perception.

Section 2, Vocabulary, involves identifying pictured vocabulary words. It is composed of eight cards, four of which picture separate items arranged on the card and four which consist of complete pictures. This section is based on the premise that identification of words through recognition may be more efficiently measured by the child's pointing to the word rather than through verbal responses. The use of complete pictures requires the child to be able to discriminate details.

Section 3, Basic General Concepts, includes subject matter and informational items which measure the child's understanding of color, number, size, differences, similarities, and position in space.

II. SAMPLE

William Kessen, in discussing research designs in the study of developmental problems, aptly described the problem encountered by this study.

Kessen stated that, realistically speaking, a completely random sample could not be done; the cost in time and money would be prohibitive. The way out used by the

pollster, the use of a stratified sample, is not available to the child psychologist. These constraints lead to the almost universal use of unsampled or selected subjects in developmental research. The investigator uses children who are available when he needs them, often without representativeness of a larger population.[1]

The sample used by the investigator in the pilot study consisted of 227 beginning first grade pupils, including 118 males and 121 females ranging in age from C.A. 5-9 through C.A. 7-5. This sample included the total population of the native non-repeating first graders in three Modesto schools. Schools "A", "B," and "C" were selected because of their comparable size and because each school was representative of a distinct socio-economic level.

School "A" is an all-white school located in the section of Modesto where most homes range in price from a low of approximately $15,000. There is very little rental property and most families own their homes. Occupations of the fathers include many professional, managerial, and skilled

---

[1] Paul H. Mussen (ed.), Handbook of Research Methods in Child Development (New York: John Wiley & Sons, 1960), p. 40.

workers. School "B" is found in a section where most families rent and the wage-earners fall into a low income grouping. There is a large federal housing project within the School "B" attendance area. Several families in this school district receive Aid to Needy Children. The school's population includes Mexican and negro children as well as white. School "C" is representative of wage-earners in the middle-income bracket. The wage-earners are primarily clerical workers, salesmen, craftsmen, truck drivers, and other blue collar workers. Family homes range in price from $7,500 to $12,500. Many families own their own homes but rental property is also available in this area. The homes, whether rented or self-owned, are generally kept in a state of cleanliness and good repair. School "C" is primarily white.

In order to describe the representativeness of the total sample, a comparison was made between the occupations of the employed fathers of the pupils and the occupations of wage-earners as reported by the 1960 census for the city of Modesto.[2] The same classification system was used by the

---

[2] United States Census of Population, 1960, California, General Social and Economic Characteristics (Washington: Government Printing Office, 1960) p. 6-316.

investigator as that used by the Bureau of Census rather than the one used by the Dictionary of Occupational Titles. The latter system is much more detailed than that used by the Bureau of Census, and it calls for many fine distinctions which could not be made from census information.

The actual number of male persons gainfully employed in Modesto was reported by the 1960 census as 8,736. The actual number in each major category was translated into percentages. These percentages are shown in Table XX on page 102 in Appendix I.

The same process was used with the sample, and at a later date with the exploratory population, to determine the relation of the pilot sample and the kindergarten sample with that of the total Modesto population. Table II on the following page shows this relationship.

III. REVISIONS AND RESULTS

Revisions. The Kindergarten Screening Test in its original form was administered to 227 beginning first grade pupils. The original test included a total of 103 items, eight of which were visual-motor coordination items, 67 were vocabulary items and 27 were basic general concept items.

TABLE II

PERCENTAGE DISTRIBUTION OF OCCUPATIONS
OF MALE WAGE-EARNERS

Occupation	Modesto Total	First Grade Sample	Kindergarten Sample
Professional, Technical, etc.	14%	13%	12%
Farmers & Managers	1	0	1
Managers, Officials, Proprietors, etc.	18	7	11
Clerical	5	5	5
Sales Workers	12	11	10
Craftsmen, Foremen	15	22	16
Operatives	13	8	9
Private Household	0	0	0
Service	6	8	7
Farm Laborers	2	0	1
Other Laborers	7	13	14
Occupation Not Reported	7	13	14

The vocabulary items were checked against the basic vocabulary list formulated at the University of Oklahoma[3] in order to determine the frequency of usage in first grade child-produced stories. Nine words used in the Kindergarten Screening Test were found to be not a part of a first grade written vocabulary, but seven were retained in the final product as they were considered by the investigator as being a part of a first grader's oral vocabulary. The frequency classification of words according to the University of Oklahoma study is found in Table XXI on page 103 in Appendix I.

After the administration of the test, each correct item was tallied. Through examination it appeared that fifteen vocabulary items and two basic general concepts items should be eliminated. Each item, including those seventeen items which appeared to be non-discriminatory, was then tested by use of the formula for finding the biserial r[4] to determine the relationship of each response to each

---

[3]Henry D. Rinsland, A Basic Vocabulary of Elementary School Children (New York: The MacMillan Company, 1945) p. 37.

[4]Joy P. Guilford, Fundamental Statistics in Psychology and Education (New York: McGraw-Hill Book Company, Incorporated, 1956), p. 297.

individual total score. These relationships are shown in detail in Table XXII, pages 104 and 105 of Appendix I. On the basis of the judgmental process, two items which showed high correlations were eliminated because they appeared too easy for the majority of the children. The instrument was then refined to include eight visual-motor coordination items, fifty-four vocabulary items, and twenty-four basic general concepts items.

Results. Frequency distribution charts for each major section of the test and for total test scores were made. Figures 2 and 4 show some skewedness to the left and Figure 3 shows a distinct skewedness which would tend to indicate that few children made low scores. These charts are included on pages 106 through 109 in Appendix I.

The range, mean, and standard deviation for the pilot group were calculated and the results are shown in Table VI on the following page.

The test-retest method was used to determine the reliability of the experimental instrument. The reliability coefficient was computed through the use of the Pearson r from raw data using thirty cases selected at random. The results are shown in Table IV on the following page.

TABLE III

MEANS, STANDARD DEVIATIONS AND RELATED DATA

	Possible Score	Range	Mean	S.D.
Section 1 Visual-Motor	8	8-0	5.8	1.9
Section 2 Vocabulary	53	52-22	43.8	5.5
Section 3 Concepts	25	25-6	21.1	4.0
Total Test	86	83-28	72.2	9.2

TABLE IV

INTERCORRELATION COEFFICIENTS AND RELIABILITY COEFFICIENTS

Variable	Section 2	Section 3	Total Score	Reliability Coefficient
Section 1 (Visual-Motor)	.87	.52	.51	.55
Section 2 (Vocabulary)		.51	.87	.66
Section 3 (Concepts)			.82	.52
Total Score				.83

Each section of the experimental instrument was compared with every other section and with total scores in order to establish intercorrelation coefficients. The Pearson r was used to compute these coefficients. The results are shown in Table IV on the previous page.

Section 3 of the experimental instrument consists of five variables which were included in one section under the title of Basic General Concepts. The investigator believed these to be separate variables with little relationship to each other and included them in a section for convenience of scoring only. Table V below shows the intercorrelations of these variables.

TABLE V

INTERCORRELATIONS OF CONCEPTS IN SECTION 3

Variables	Number	Size	Likeness	Difference	Position
Color	.14	.16	.02	.26	.18
Number		.19	.10	.16	.27
Size			.36	.24	.26
Likeness				.23	.11
Difference					.28

Conclusions. Several generalizations may be made concerning the results obtained in the pilot study.

1. The sample used is representative of the Modesto population as a whole.

2. Most of the vocabulary words used are included in the written vocabulary of first graders.

3. The experimental instrument does not appear to have enough "top" for use with first graders. Modification of the scoring through use of a weighting system would provide a more symmetrical form of raw score distribution.

4. According to the general verbal description of coefficients used by Guilford,[5] the reliability coefficients show a modest relationship among the sections and a moderately strong correlation with total scores.

5. In comparing the test sections with each other, Section 3, Basic General Concepts, correlates least well, showing a mild relationship with the other sections but a moderately strong degree of relationship with total score. Section 1, Visual-motor Coordination, shows a modest correlation with total score but a moderately strong relationship with Vocabulary.

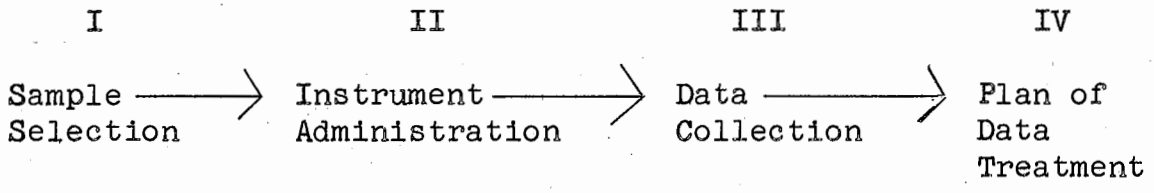
---

[5] Ibid., p. 145.

CHAPTER IV

DESIGN OF EXPERIMENT

This chapter is concerned with the design of the study. This includes description of the method of collection of data and the data treatment plan. Actual treatment of the data will be described in Chapter V. The investigation with the experimental group began in the spring following the pilot study described in Chapter III. The steps to be described in Chapter IV are illustrated below:



I. THE SAMPLE

An early problem to be encountered by an investigator is the determination of his population and the selection of the sample. An error at this stage of development could so bias the results as to make them invalid. Many statisticians cite the error made by the Literary Digest public opinion poll during the 1936 presidential campaign as a prime example

A comparison was made between the occupations of the employed fathers of the children included in the sample and the occupations of the Modesto male wage-earners as reported by the 1960 census.[2] Total numbers in each occupation category used by the census report were converted into percentages, as were the total numbers of employed fathers of children in the sample, in order to make meaningful comparisons. The actual percentages, listed according to census categories, are found in Table II on page 36.

The resultant comparison shows that the study population was considerably below the Modesto percentage in the Managerial and Operatives categories and above the percentage reported for the categories General Laborers and Occupations Not Reported. In all other areas the percentage distribution of male wage-earners included in the sample was within two percentage points of the total wage-earners as reported. A comparison of the study population fathers and total Modesto population would indicate that the sample population is fairly representative of the total male labor

---

[2] United State Census of Population, 1960, California, General Social and Economic Characteristics (Washington, D. C.: Government Printing Office, 1960), p. 6-316.

force in Modesto.

As the study progressed the size of the sample changed. No additional kindergarten pupils were added, but fifty-five were eliminated because they had moved from the school district and a follow-up was impossible. In each phase of the investigation the number of pupils used in computations consisted of the subjects who were still enrolled in the same school in which they were at the beginning of the study.

II. EXPERIMENTAL INSTRUMENT

The revised experimental instrument, entitled the Kindergarten Screening Test, which includes the test description, the directions for administration, test pictures, a scoring key, and a record form for recording responses, is located on pages 111 to 144 in Appendix II.

The test is divided into three sections. Visual-motor Coordination section consists of eight designs which are arranged according to difficulty. Six of the designs were developed by the investigator and two of the designs, the square and the triangle, are found in the Revised Stanford-

Binet, Form L-M.[3]

Fifty-four vocabulary words comprise Section Two. These words are presented in pictorial form. Four sets of pictures representing words as separate entities are comparable in size and located together on four successive pages. Page 121 depicts fruits and vegetables, pages 122 and 123 show common articles, and page 124 is comprised of pictures of tools. Pages 125 through 132 consist of single pictures. Vocabulary words relating to these four pictures require the ability to discriminate details as well as to identify prominent objects.

Section Three, Basic General Concepts, is designed to measure subject matter and informational items which are considered by the investigator to be within the experiential background of most first graders. The twenty-four items in this section include: (1) identification of five colors, (2) counting of five groups of dots, (3) recognition of the size of four objects, (4) identification of three objects which are different from others in a series, (5) identifi-

---

[3]Lewis Terman and Maud Merrill, Stanford-Binet Intelligence Scale (Boston: Houghton Mifflin Company, 1960), pp. 82 and 87.

cation of three objects which have identical mates, and (6) locating the position in space of four objects.

The full-scale instrument, comprising eighty-six items, was used in the investigation with the study population of 337 kindergarten pupils.

III. DATA COLLECTING TECHNIQUES

The Kindergarten Screening Test was administered to 337 second semester kindergarten pupils in three schools in the Modesto City Schools district. The tests were scored, and then the raw scores were tabulated both by sex and by school. Scoring norms devised by the investigator were used to convert the raw scores into classifications denoting expected achievement levels. Table VI on the following page lists the scoring levels and the descriptive classification of each.

The Lee-Clark Reading Readiness Test was next administered by the investigator to 324 kindergarten pupils who had been included in the original study population. The thirteen pupils who were not tested at this time were either absent because of prolonged illness or had moved from the district. These tests were scored and the raw scores

recorded in the same manner as had been done with the experimental instrument. Norms for use with kindergarten pupils tested late in the second semester were used to convert the raw scores into expected achievement levels. The norms used are listed in Table VII on the following page.

TABLE VI

KINDERGARTEN SCREENING TEST READINESS LEVELS

Score	Classification
0-55	Low
56-63	Low Average
64-75	High Average
76-86	High

The Kindergarten Screening Test was re-administered by the investigator two weeks after the initial administration of the instrument to a random sample of thirty kindergarten pupils. The sample was obtained by the assigning of numbers to all pupils originally tested and drawing thirty numbers at random. Once a number had been selected it was

not returned to the total group. The thirty retests were scored according to the same standards as originally used. Retesting was done in order to obtain raw scores for use in computing Pearson reliability coefficients.

TABLE VII

LEE-CLARK READING READINESS NORMS

GROUP II: END OF KINDERGARTEN

Score	Classification
0-32	Low
33-39	Low Average
40-51	High Average
52-64	High

All testing and scoring completed at this time had been done by the investigator. The question now arose as to whether or not an untrained examiner would obtain similar results when administering and scoring the Kindergarten Screening Test. Since the Kindergarten Screening Test had been originally devised for use by kindergarten teachers it was felt it should be tested by such teachers. A certifi-

cated kindergarten teacher who was on leave of absence administered the Kindergarten Screening Test to twenty pupils who were selected at random from School "B", representing the low socio-economic group. Her testing was restricted to one school at the request of the school district administration. Pupils at School "B" who had previously been retested by the investigator were eliminated from the selection. Thus the sample was limited to kindergarten pupils from one school who had not already been retested on the experimental instrument. A random sample with the above restrictions was obtained. As previously, when a number was drawn it was not returned to the original group. After administering the Kindergarten Screening Test the kindergarten teacher scored the protocols. The teacher completed the administration and scoring without having received any previous instructions other than those given in the printed test directions.

The compiled raw data now included:

1. Raw scores and general classification listing obtained for 337 pupils on the Kindergarten Screening Test.

2. Raw scores and general classification listing obtained for 324 pupils on the Lee-Clark Reading Readiness Test.

3. Raw scores obtained for thirty pupils on the Kindergarten Screening Test re-administered by the investigator.

4. Raw scores obtained for twenty pupils on the Kindergarten Screening Test re-administered by a kindergarten teacher.

IV. PLAN OF DATA TREATMENT

The first treatment of the accumulated data was to determine the means and standard deviations for each section of the Kindergarten Screening Test, and to determine the relationships between each test section with every other test section and between each test section and total score.

Reliability coefficients for each test section and for total score were computed using the raw scores obtained by a random sample of thirty kindergarten pupils in a test-retest situation. The Pearson r was used for this computation.

A reliability coefficient was established to determine the relationship between the raw scores of twenty kindergarten pupils who were tested first by the investigator and three weeks later retested on the same instrument, the Kindergarten Screening Test, by a certificated kindergarten teacher. The relationship between the two sets of raw scores was obtained by using the Pearson r.

In this exploratory validation study the validity of the Kindergarten Screening Test was investigated by determining the relationship between it and a recognized readiness test. Lee-Clark Reading Readiness Test total scores were compared with total scores obtained on the Kindergarten Screening Test through the use of the formula for computing a Pearson r.

In order to determine the relationship between the Kindergarten Screening Test and kindergarten teacher judgment a coefficient of contingency was computed between total scores and kindergarten teacher judgments. The relationship between total scores on the Lee-Clark Reading Readiness Test and kindergarten teacher judgment was established through the use of the same statistical method.

Since the experimental instrument was devised as an objective aid to the kindergarten teacher in her decision-making process it was necessary to determine the relationship between the Kindergarten Screening Test and student success in a first grade program. Reading group placement, whether the pupil was in the top, middle, or bottom group or not yet reading, needed to be compared with total scores on the Kindergarten Screening Test. The coefficient of contingency

was used to compute this relationship. The coefficient of contingency was also computed between actual reading group placement and total scores on the Lee-Clark Reading Readiness Test in order to determine whether the experimental instrument could predict first grade success as well as an established test.

The final comparison was to establish the relationship between kindergarten teacher judgment and actual first grade reading group placement. A chi square was used for this computation since both evaluations were categorized.

Other productive comparisons were developed which were incidental to the main study but which were instructive in nature. These comparisons were made and were included in the study since they added to the original investigation.

CHAPTER FIVE

PRESENTATION OF DATA

Data treatment was rather complex because of the variety of forms of raw data obtained. With the Kindergarten Screening Test or the Lee-Clark Reading Readiness Test raw scores could be used. However, both kindergarten teacher judgment and first grade post-readiness reading group placement were reported in four categories or levels of achievement. Therefore data relating to either of these two areas had to be investigated with statistical methods which would be most appropriate for the particular situation.

The raw scores obtained by the study population on the Kindergarten Screening Test were interpreted first. Charts showing the distribution of raw scores on each test section and on total test scores were devised to illustrate raw score dispersion. These distributions are found in Figures 1 through 4 on pages 106 through 109 in Appendix I. Some negative skewing was present in Section Two, Vocabulary, and Section Four, Total Score and definite skewing in Section Three, Basic General Concepts which in turn influenced the total score toward negative skewing. Section One, Visual-

motor Coordination, showed a more symmetrical form of distribution.

The means and standard deviations for each test section and for total test scores were computed. Additional data, including total possible scores and the range of raw scores, were compiled and included in Table VIII below. This information was used to establish the categories or levels for the verbal description of test norms.

TABLE VIII

MEANS, STANDARD DEVIATIONS AND RELATED DATA FOR THE KINDERGARTEN SCREENING TEST

	Possible Score	Range	Mean	S.D.
Section 1 Visual-Motor	8	0 - 8	4.0	1.9
Section 2 Vocabulary	54	11 - 53	42.8	5.9
Section 3 Concepts	24	6 - 24	17.8	3.6
Total Test	86	14 - 83	64.7	9.6

The raw scores of Section One, Visual-motor Coordination, ranged from a perfect score of eight to zero correct

reproductions. The mean for this section was 4.0 and the standard deviation was 1.9. One standard deviation above and below the mean included raw scores from 2.1 through 5.9. Two standard deviations above and below the mean included raw scores from .2 through 7.8. Sixteen pupils were unable to reproduce successfully any design and scored a total of zero on this section. Nine girls and seven boys made up this group, thus indicating that reproduction of designs is relatively difficult for both boys and girls. One boy and one girl from School "A", the highest socio-economic level, scored zero. Five boys and four girls from School "B", the lowest socio-economic level, also were completely unable to reproduce designs and scored zero on this test section.

Section Two, Vocabulary, contained a possible perfect score of fifty-four. The raw scores actually ranged from a high score of fifty-two to a low score of eleven. The mean for Section Two was 42.8 and the standard deviation was 5.9. Raw scores ranging from 36.9 through 48.7 fell within one standard deviation above and below the mean. Two standard deviations above and below the mean included raw scores from 31.0 through 54.4. Nine raw scores fell below two standard deviations below the mean and no scores fell above two

standard deviations above the mean. One boy in School "A", the highest socio-economic level, fell into this designation. He had previously been identified by a school psychologist as being emotionally disturbed. The remaining four boys and girls were from School "B", the lowest socio-economic group. Two of these pupils who scored eleven and sixteen points had been identified by a school psychologist as mentally retarded. No raw scores for students from School "C", the middle socio-economic level, fell below two standard deviations from the mean.

Raw scores of Section Three, Basic General Concepts, ranged from six through twenty-four; a perfect score was twenty-four. The mean for this test section was 17.8 and the standard deviation was 3.6. Scores ranging from 14.2 through 21.4 fell within one standard deviation from the mean. Two standard deviations included scores from 10.6 through 25.0. Since the total possible score was twenty-five, no pupils were able to score beyond two standard deviations above the mean. The scores of sixteen pupils fell more than two standard deviations below the mean. Schools "A" and "C", representing the highest and middle socio-economic levels of the experimental group, contained a total of three girls and

four boys. Two boys and two girls were from School "A" and two boys and one girl were from School "B." This would indicate that there was little difference between the number of pupils below two standard deviations in these two schools. The raw scores of five boys and three girls fell below two standard deviations from the mean in School "B", the lowest socio-economic level.

The same pattern as was evidenced in each test section was also apparent in comparing total raw scores. The total possible score for the complete test was eighty-six and the raw scores ranged from fourteen through eighty-three. The total test mean was 64.8 and the standard deviation was 9.6. Scores falling within one standard deviation above and below the mean ranged from 55.2 through 74.4 Two standard deviations included raw scores between 45.6 and 84.0. No raw scores fell within three standard deviations above the mean. In School "A" one girl fell within three standard deviations below the mean. This pupil was withdrawn by her parents at a later date upon the recommendation of the kindergarten teacher. In School "C", the middle socio-economic level, one boy and one girl fell below two standard deviations below the mean. In one instance the teacher had rated the boy as

expected low average achievement in first grade and the girl as expected high average achievement. Three boys and six girls in School "B" obtained total scores which fell within three standard deviations below the mean. All nine pupils were rated by their kindergarten teacher as "4" or expected inability to achieve in a regular first grade program.

In order to interpret the relationship between each section of the Kindergarten Screening Test and every other section, and with total test scores, a Pearson r was computed in each instance. The intertest correlation coefficients are shown in Table IX below.

TABLE IX

INTERCORRELATION COEFFICIENTS OF THE KINDERGARTEN SCREENING TEST

Variable	Test 2	Test 3	Total Test
Test 1	.40	.49	.67
Test 2		.59	.78
Test 3			.84

ship. Section Three, Basic General Concepts, in addition to a low relationship with Visual-motor Coordination and a moderate relationship with Vocabulary shows a more substantial relationship with Total Score. The coefficient of correlation between Section Three and Total Score was .84. This was the highest relationship established.

After establishing the interrelationships among the test sections, the next step was that of establishing a reliability coefficient. The Kindergarten Screening Test was initially administered to the total study population; a retest was administered three weeks later using thirty random cases. A Pearson r was computed using the raw scores obtained from the test-retest situation with the results shown in Table X on the following page.

The reliability coefficients in Table X show a moderate relationship existing between raw scores upon readministration of the Kindergarten Screening Test.

Section Three showed the least reliability with a coefficient of .56. Sections One and Two had stronger reliability coefficients of .76 and .70 respectively. The reliability coefficient for Total Score was .68 suggesting a modest, positive relationship.

showing a substantial relationship. This result would indicate that the Kindergarten Screening Test could be effectively administered and scored by an untrained examiner who had experience in teaching children of this age group.

The basic purpose as defined in this study was to establish an initial validation on a limited population thus the problem was to determine what were the relationships among the Kindergarten Screening Test, the Lee-Clark Reading Readiness Test, kindergarten teacher judgment, and first grade post-readiness reading group placement. This problem was stated in terms of the null hypothesis. The comparisons were made and presented as coefficients of correlation. They are found in Table XI below.

TABLE XI

COEFFICIENTS AMONG THE KINDERGARTEN SCREENING TEST, THE LEE-CLARK READING READINESS TEST, KINDERGARTEN TEACHER JUDGMENT AND ACTUAL READING GROUP PLACEMENT

	Lee-Clark	Kindergarten Teacher Judgment	Reading Grouping
Kindergarten Screening Test	.74	.58	.52
Lee-Clark		.49	.59
Kindergarten Teacher Judgment			.51

The first relationship was established between the Kindergarten Screening Test and the Lee-Clark Reading Readiness Test. The raw scores of the study population of kindergarten pupils were used to compute the Pearson product-moment coefficient of correlation. This method of obtaining two sources of information and correlating the two criteria is a common means of establishing concurrent validity. The validity coefficient obtained between the experimental instrument and an established readiness test, the Lee-Clark Reading Readiness Test, was .74. This coefficient would indicate a moderate correlation and so the existence of a definite relationship between the two instruments. A "t" test for determining the significance of a coefficient of correlation was computed using Fisher's formula.[2] Applying the formula to a Pearson r of .74 and an N of 325 yielded a "t" of 19.76. A "t" this high would indicate that the hypothesis that the population correlation was zero could be rejected at the .01 level of confidence. We may further state that there is a definite relationship between the experimental instrument and an established readiness test,

[2] Ibid., p. 219.

thus demonstrating concurrent validity.

We may thus assume that there is a definite relationship between functioning on the Kindergarten Screening Test and the Lee-Clark Reading Readiness Test and that similar functions are being measured on both instruments. It would also appear that the Kindergarten Screening Test might be used as an alternate for the Lee-Clark Reading Readiness Test.

The second validity study was made by comparing the raw scores of the study population with judgments made by kindergarten teachers. The judgments were made by the teacher on each class and were reported to the investigator in the form of a rating. "1" indicated an expectation of high achievement in Grade One; "2" indicated an expectation of high average achievement; "3" meant expected low average achievement, and "4" meant expected low achievement or an inability to achieve success in a first grade program. The kindergarten teachers had no knowledge of test scores obtained from either the Kindergarten Screening Test or the Lee-Clark Reading Readiness Test. Since the comparison to be made was between raw scores and ratings, the formula for

the coefficiency of contingency was used.[3] Raw scores were converted into six categories and teacher rating into four categories which limits the maximum coefficient, C, which might be obtained to between .91 and .866. The coefficiency of contingency between the Kindergarten Screening Test and kindergarten teacher judgment was .58. An early step in computing the coefficiency of contingency was to find chi square, which in this instance was 171.1. Table XII below shows the observed frequencies used in computing the chi square. Using twenty-one degrees of freedom the significance level of this chi square was significant beyond the .01 level.

TABLE XII

OBSERVED FREQUENCIES OF KINDERGARTEN SCREENING TEST SCORES AND KINDERGARTEN TEACHER RATING

		Scores 14-55	56-62	63-65	66-68	69-71	72-83
Rating	1	1	7	11	11	18	46
	2	5	17	15	23	28	28
	3	26	25	18	9	10	3
	4	20	14	1	1	0	0

[3] Ibid., p. 316.

The raw scores on the Lee-Clark Reading Readiness Test were compared with kindergarten teacher judgment using the same ratings as described in the paragraph above and in the previous computations. The Lee-Clark raw scores were also divided into five categories. The coefficient of contingency obtained between the Lee-Clark scores and kindergarten teacher judgment was .49. The obtained chi square was 101.5 which is beyond the .01 level of significance. Observed frequencies are shown in Table XIII below.

TABLE XIII

OBSERVED FREQUENCIES OF LEE-CLARK READING READINESS TEST AND KINDERGARTEN TEACHER RATING

		Scores 10-35	36-41	42-50	51-55	56-64
Rating	1	3	6	29	26	28
	2	8	20	38	31	14
	3	34	21	18	9	4
	4	22	8	5	0	0

In order to determine whether the correlation obtained between the Kindergarten Screening Test and kindergarten

teacher judgment was significantly higher than between the Lee-Clark Reading Readiness Test and teacher judgment, a $\bar{z}$ test of difference between two z coefficients was computed.[4] The standard error of the difference between the two coefficients was .077, and the $\bar{z}$ was 1.298. This difference is not significant at the .05 confidence level so the assumption may be made that there is no significant difference between the two coefficients of correlation.

The Kindergarten Screening Test and the Lee-Clark Reading Readiness Test show similar relationships with kindergarten teacher judgment. The relationship between the Kindergarten Screening Test and teacher judgment was .58 and between the Lee-Clark Reading Readiness Test and kindergarten teacher judgment was .49.

The experimental instrument did not show a significantly higher relationship with kindergarten teacher judgment than did the Lee-Clark Reading Readiness Test. Both instruments showed a moderate relationship, thus leading to the tentative conclusion that predictions of first grade success made by kindergarten teachers are based on information which

---

[4] Ibid., p. 194.

might not be obtainable on objective-type tests. Such information might include work habits, social skills, and attitude toward school.

Two questions then arise. How well do certain objective tests predict first grade success? Are they better predictors of first grade success than kindergarten teacher judgment?

The next step was to determine the relationship between the Kindergarten Screening Test and actual first grade post-readiness reading group placement. The data regarding reading group placement was obtained in late October, 1964. The first grade teachers had six weeks in which to determine the reading group in which each child could work best. Thus a period of five to six months lapsed between the time of test administration and the final stage of data collecting. An inevitable condition of the study was that first grade teachers had to be given access to Lee-Clark scores prior to making groupings. This introduction of contamination between the Lee-Clark and first grade grouping cannot be assessed and may be assumed to affect correlations between these two specifically. First grade teachers had no access to Kindergarten Screening Test scores so no contami-

nation exists here.

Reading groups in the schools used in the investigation were divided into levels. Numerical ratings were assigned by the investigator to each reading group or level. The highest or "top" reading group was assigned the number "1"; the middle reading group was "2", the bottom group was designated as "3" and the group which was still in a readiness program after six weeks of school was called "4." Those children who were retained in kindergarten were also placed in category "4."

The raw scores on the Kindergarten Screening Test were compared with the actual first grade reading group placement of the individuals involved in the study using the same ratings as were described in the previous paragraph. The Kindergarten Screening Test raw scores were divided into six categories which limited the maximum obtainable "C" to .91. The coefficiency of contingency obtained between the Kindergarten Screening Test raw scores and actual reading placement was .52. The obtained chi square was 101.0 which, using fifteen degrees of freedom, is beyond the .01 level of significance. Table XIV on the following page shows the observed frequencies used in obtaining the chi square.

TABLE XIV

OBSERVED FREQUENCIES OF KINDERGARTEN SCREENING TEST AND READING GROUP PLACEMENT

		Scores 14-55	56-62	63-65	66-68	69-71	72-83
Rating	1	2	8	8	11	19	44
	2	7	10	13	16	19	17
	3	11	17	8	7	7	2
	4	20	14	4	4	0	1

The raw scores on the Lee-Clark Reading Readiness Test were then compared with actual reading group placement. The same ratings were used to describe reading group placement as were previously described. The coefficient of contingency between Lee-Clark scores and reading group placement was .59. The chi square was 144.1. The observed frequencies are shown in Table XV on the following page. Using twelve degrees of freedom, the obtained chi square is beyond the .01 level of significance.

The next computational step was to determine whether or not the obtained correlations of .52 between the Kinder-

garten Screening Test scores and actual reading group placement and .59 between Lee-Clark scores and reading group placement were significantly different. A $\bar{z}$ test of the difference between two coefficients[5] was computed. The standard error of the difference between the two z coefficients was .087 and $\bar{z}$ was 1.15. This difference is not significant at the .05 level thus leading to the assumption that there is no significant difference between the obtained coefficients.

TABLE XV

OBSERVED FREQUENCIES OF LEE-CLARK READING READINESS TEST AND READING GROUP PLACEMENT

		Scores 10-35	36-41	42-50	51-55	56-64
Rating	1	1	3	27	32	30
	2	8	13	37	14	9
	3	17	18	11	6	0
	4	28	10	3	1	1

[5] Ibid.

The correlations of .52 and .59 are not significantly different thus leading to the assumption that the Kindergarten Screening Test does not predict first grade post-readiness reading group placement to any greater degree than does the Lee-Clark Reading Readiness Test. If both instruments show only moderate predictiveness then we may hypothesize that the Kindergarten Screening Test could be used as an alternate for the Lee-Clark Reading Readiness Test. Since the Kindergarten Screening Test requires ten minutes to administer, it would appear that the Kindergarten Screening Test would save teacher time when only one or two students are involved.

It should be observed that there is a possibility that the predictive utility of the Kindergarten Screening Test may actually be superior to that of the Lee-Clark. This is because the first grade teacher had access to Lee-Clark scores whereas Kindergarten Screening Test scores were not available. Therefore, the correlation between the Lee-Clark and first grade teacher ratings may be spuriously high and so not actually be comparable to the Kindergarten Screening Test correlation with teacher rating.

The final question was to determine whether or not

kindergarten teacher judgment would more effectively predict first grade post-readiness reading group placement than the two instruments.

The relationship between kindergarten teacher judgment and actual first grade post-readiness reading group placement was established. Since both sets of data had been divided into four categories, an eight cell chi square was first computed yielding a 95.0 which was significant beyond the .01 level. Table XVI below shows the observed frequencies in each cell.

TABLE XVI

OBSERVED FREQUENCIES OF KINDERGARTEN TEACHER RATING AND READING GROUP PLACEMENT

		4	3	2	1
Rating	1	2	6	23	48
	2	5	20	33	35
	3	20	18	19	8
	4	17	8	6	1

Using the chi square for computations, a .51 coefficient of correlation resulted.

The relationship of .51 between kindergarten teacher judgment and first grade post-readiness reading group placement is moderate. We may then assume that kindergarten teacher judgment is as good a predictor as the Kindergarten Screening Test or the Lee-Clark Reading Readiness Test. One might then question the advisability of using either instrument. As far as predictive validity is concerned this is questionable. However, as was previously stated in another chapter, kindergarten teachers are reluctant to pass or fail questionable cases on the basis of subjective judgment alone. An objective test providing quantified data can thus be used by the kindergarten teacher to enhance or support her judgment. In addition, tests provide the teacher with a relatively objective criterion which is easily explained to parents. The Kindergarten Screening Test could be effectively used in this situation since it provides objective data. The Kindergarten Screening Test has an additional value in that it can be administered in a single sitting of ten minutes whereas the Lee-Clark Reading Readiness Test requires two sittings if the children appear to be tiring during the second test. The latter requires three times as long to administer, although it may be assumed until broader

standardization of the Kindergarten Screening Test occurs to have higher reliability.

Multiple correlations were next computed in order to establish the relationships among a dependent variable, post-readiness reading group placement, and three independent variables, kindergarten teacher judgment, the Kindergarten Screening Test and the Lee-Clark Reading Readiness Test.[6] The coefficients of correlation used to establish validity were also used in these computations. Table XIV on page 63 lists these obtained coefficients.

TABLE XVII

MULTIPLE CORRELATIONS AMONG FOUR VARIABLES*

Variables	X_2 and X_3	X_2 and X_4	X_3 and X_4
X_1	.58	.64	.90

* X_1 post-readiness reading group placement

X_2 kindergarten teacher judgment

X_3 Kindergarten Screening Test

X_4 Lee-Clark Reading Readiness Test

[6] Ibid., p. 393.

The correlation between post-readiness reading group placement, as the dependent variable, and kindergarten teacher judgment coupled with the Kindergarten Screening Test was .58. The correlation between post-readiness reading group placement, the dependent variable, and kindergarten teacher judgment coupled with the Lee-Clark Reading Readiness Test was .64. These correlations show a definite moderate relationship, thus leading to the tentative conclusion that either test could be used with the same effectiveness to enhance kindergarten teacher judgment. The advantage of using the Kindergarten Screening Test would be that it is a quick screening device and would require a shorter time to administer.

It is interesting to note that the Kindergarten Screening Test coupled with the Lee-Clark shows a relationship of .90 with post-readiness reading group placement, leading to the tentative conclusion that a combination of the two tests and kindergarten teacher judgment is the best predictor. It should be noted that even though this is a high correlation, the length of time required to administer both tests might make this process undesirable. There also remains the possibility of a spuriously high correlation

between the Lee-Clark and actual first grade post-readiness reading group placement which would affect the above tentative conclusions.

In addition to the primary concern of the investigations, the establishment of the relationships among the Kindergarten Screening Test, the Lee-Clark Reading Readiness Test, kindergarten teacher judgment and first grade post-readiness reading group placement, some interesting observations could be made comparing the total test score means of the various subgroups within the study population on the Kindergarten Screening Test. These subgroup means are shown in Table XVIII below.

TABLE XVIII

KINDERGARTEN SCREENING TEST SUB-GROUP MEANS

School	A.M. Classes	P.M. Classes	Teacher A	Teacher B	Boys	Girls
A	67.8	67.7	68.7	66.8	66.6	67.8
B	60.5	59.4	61.5	58.5	58.2	61.3
C	65.3	67.2	66.1	65.8	65.8	65.8
Total	64.8	64.6	65.6	65.3	63.4	65.1

There was relatively little difference between morning and afternoon classes. Pupils had been assigned to classes in order of enrollment, the principals involved usually filling the morning classes first. The means for morning classes were 67.8 in School "A", 60.5 in School "B" and 65.3 in School "C." The afternoon means were 67.7 in School "A", 59.4 in School "B" and 67.2 in School "C." The total means for morning classes were 64.8 and 64.6 for the afternoon classes, bearing out the success of the aim of nondiscriminatory grouping. The differences here were so obviously close that a statistical test was not performed.

The Kindergarten Screening Test means of classes varied according to teacher quality as estimated by principals. The principals' opinions were obtained by interview.

Table XIX on the following page suggests on an inspectional basis that there is a possibility that Kindergarten Screening Test means might vary according to teacher quality but the N is too small and the limited data inconclusive.

In Schools "A" and "B", which represented the upper and lower socio-economic levels of the experimental group, girls scored higher than boys. The girls' mean in School "A"

was 67.8 and the boys' mean was 66.6. In School "B" the girls' mean was 61.3 and the boys' mean was 58.2. In School "C", which represented the middle socio-economic group, the boys' and girls' means were both 65.8. The total means for boys were 63.4 and 65.1 for girls.

TABLE XIX

CLASS VARIATION ON THE KINDERGARTEN SCREENING TEST

School	Teacher	Mean	Principals' Opinions
A	a	68.7	Excellent
A	b	66.8	Good
B	c	61.5	Excellent
B	d	58.5	Good
C	e	66.1	Average
C	f	65.8	Good

Through observation one may readily see that the means for School "A", the upper level of the study population, were highest, 67.8 and 67.7, School "C", the middle level, had means of 65.3 and 67.2, whereas the means of School "B", the lower level, were 60.5 and 58.4.

In summary, the relationships among the Kindergarten Screening Test, the Lee-Clark Reading Readiness Test, kinder-

garten teacher judgment, and first grade reading group placement were all moderate positive relationships. The closest relationship of .74 was established between the Kindergarten Screening Test and the Lee-Clark Reading Readiness Test. The experimental instrument showed a definite relationship with kindergarten teacher judgment, a coefficient of .58, as did the Lee-Clark and teacher judgment with a correlation of .49. Definite relationships were also shown between reading group placement and the other variables.[7]

Operationally speaking, it would appear from the results obtained in this study that the Kindergarten Screening Test could be employed as a screening device either at the end of the kindergarten period or during the first few weeks of first grade during early readiness. The reliability coefficients obtained using the test-retest method with the kindergarten group were quite similar to those obtained by the same method with beginning first grade pupils. These

---

[7]

Variable	Reading group placement
Kindergarten Screening Test	.52
Lee-Clark Reading Readiness Test	.59
Kindergarten teacher judgment	.51

findings would suggest the possibility of using the Kindergarten Screening Test at either time with equal success.

The Kindergarten Screening Test was designed as an additional means for evaluating pupil progress other than teacher judgment. The kindergarten teacher would not necessarily use this appraisal device with an entire class but rather with a few pupils about whom she was concerned. Some reasons why she might use the test could include (1) she might wish the additional data because she was undecided as to what action to take (2) the parents might be unwilling to accept teacher judgment alone and wish additional data or (3) the school executive might desire more objective data about the pupil before recommending either repeating the kindergarten experience or remaining at home for an additional year before beginning first grade.

The kindergarten teacher would administer the test on an individual basis, in a locale away from group activity and according to the directions given in Part Two of the Test, found in Appendix II of this text. The test would be scored according to the directions found in Part Three. If the pupil's score fell below fifty-five, he would be rated as "low", representing a rating of anticipated lack of success in a first grade program.

CHAPTER VI

CONCLUSIONS AND RECOMMENDATIONS

The purpose of this investigation was one of exploratory validation. This was accomplished through establishing the relationships among the Kindergarten Screening Test, the Lee-Clark Reading Readiness Test, kindergarten teacher judgment, and actual first grade reading group placement. The problem was expressed in terms of null hypotheses that there were no significant relationships between the Kindergarten Screening Test and the other means of evaluating readiness for First Grade.

The experimental instrument, the Kindergarten Screening Test, was designed by the investigator as an additional means at the disposal of the teacher for the evaluation of pupil progress and maturation at the end of the kindergarten period.

I. CONCLUSIONS

Null hypotheses. The hypothesis that there would be no significant relationship between scores on the Kindergarten Screening Test and the Lee-Clark Reading Readiness

Test was rejected at the .01 confidence level. In rejecting the hypothesis that the population correlation was zero, we may assume that factors other than chance are in operation. At this exploratory stage the .74 coefficient of correlation obtained suggests the hypothesis that there is a relationship between functioning on the Kindergarten Screening Test and the Lee-Clark Reading Readiness Test and that similar functions are being measured on both instruments. The Kindergarten Screening Test would also appear to be a useful alternate instrument for this sample and for similar selected samples. Broader validation would demonstrate whether it would be equally relevant for other populations.

The hypothesis that there would be no significant relationship between scores on the Kindergarten Screening Test and kindergarten teacher judgment was rejected at the .01 confidence level. In rejecting the hypothesis that the correlation was zero we may assume that factors other than chance are in operation. The relationship of .58 between the Kindergarten Screening Test and kindergarten teacher judgment would indicate a definite relationship. It should be noted that kindergarten teacher judgment is traditionally used as the basis for initial reading group placement.

The hypothesis that there would be no significant relationship between the Kindergarten Screening Test and first grade reading group placement was rejected at the .01 confidence level. It should be noted that reading group placement was made after an intervening variable, a readiness program, was introduced. Additional variables which influenced post-readiness reading group placement were the recommendations of kindergarten teachers and the kindergarten scores obtained on the Lee-Clark Reading Readiness Test, both of which were available to first grade teachers. There were also maturational factors which were in operation in the early first grade year which were not testable in kindergarten and which were not designed to be tested by the Kindergarten Screening Test. The coefficient of correlation of .52 would indicate that the Kindergarten Screening Test would not appear to be appropriate for use in the post-readiness stage of First Grade. The Kindergarten Screening Test as other performance tests, including the lower levels of the Stanford-Binet, may eventually prove to be similar to other young-child tests. It may correlate better with pre-first grade maturation and less well with abstract functioning following some first grade instruction. The Kindergarten

Screening Test appears to be moderately useful in kindergarten assessment but less useful for first grade placement following readiness training.

Additional relationships. Additional relationships not relating directly to the null hypotheses but pertinent to the total investigation were established. These relationships led to interesting conclusions as follows.

A relationship of .49 was established between the Lee-Clark Reading Readiness Test and kindergarten teacher judgment. It is interesting to note that the relationship between the Kindergarten Screening Test and kindergarten teacher judgment was .58 thus leading to the conclusion that, with the population used, the two instruments showed similar modest correlations with teacher judgment. It would thus appear that for this population and for similar selected samples the Kindergarten Screening Test may provide data which could be as effectively used to enhance teacher judgment as could data provided by the Lee-Clark Reading Readiness Test. It would be interesting to verify this with a larger and more representative population.

A possibly contaminated coefficient of contingency of .59 between the Lee-Clark Reading Readiness Test and first

tests can predict better than kindergarten teacher judgment. The results tend to indicate the generalization that, just as would ability and achievement tests, teacher judgment is similarly affected by pupil maturation and training.

The evident fact is that the Kindergarten Screening Test, the Lee-Clark Reading Readiness Test and kindergarten teacher projected ratings were all obtained while the pupils were still in kindergarten. The intervening variables of maturation, knowledge of test scores and kindergarten teacher judgment should then affect post-readiness reading group placement to a comparable extent.

Multiple correlations showed a relationship of .58 between post-readiness reading group placement and the Kindergarten Screening Test coupled with kindergarten teacher judgment. These correlations would indicate that either test could be used to enhance teacher judgment with the same effectiveness. Use of the Kindergarten Screening Test might prove more desirable because of the shorter time required for administration.

Characteristics of the experimental instrument. Intertest correlations within the Kindergarten Screening Test

varied considerably. The correlation between Visual-motor Coordination and Vocabulary was .40 and between Visual-motor Coordination and Basic General Concepts was .49. The correlation between Vocabulary and Basic General Concepts was .59. It would appear that, as is desirable with batteries, there is a mild, positive correlation between the various test sections.

A more definite relationship was established between each test section and Total Score. The coefficient of correlation between Visual-motor Coordination and Total Score was .67. The correlation between Vocabulary and Total Score was .78 and between Basic General Concepts and Total Score was .84. These results suggest that although the relationships among test sections are desirably low, the relationship of each section to total score is, as is to be anticipated, higher.

Reliability coefficients established by the test-retest method showed a moderate relationship of .68 existing when the test-retest was administered by an experienced examiner and a coefficient of .77 when the retest was administered by an untrained examiner who had no instructions other than those given in the printed directions. These

coefficients would indicate that the Kindergarten Screening Test could be administered and scored by an untrained examiner as well as by a trained examiner. It would also confirm that a brief screening test would produce less stable scores than a longer instrument.

Scores of students in the classes of teachers who had been rated "excellent" were higher than those of students who were in classes of teachers rated as "good" and "average." Student scores in the classes of "average" teachers were lower than those of the two higher classifications.

The Kindergarten Screening Test tended not to discriminate well among the better students, indicating a lack of "top" to the test. The major concern of the design of the instrument in the investigation, however, was with the distributions at the lower end of the scale which were susceptible to fine discrimination by the nature of the curve.

There was no appreciable difference in raw scores on the experimental instrument between morning and afternoon classes or between boys and girls.

Observation of the Kindergarten Screening Test subgroup means presented in Table XVIII on page 78 would indicate

that differences in test scores existed among the three schools, implying a difference in raw scores according to socio-economic level. This would tend to confirm that comprehensive validation should be done after an extensive provisional try with the experimental instrument.

II. RECOMMENDATIONS

Follow-up studies in future investigations should include administration of the experimental instrument to a large representative sample of the general kindergarten population.

A weighting system may be desirable prior to use of the instrument with a stratified random sample. This would provide a more symmetrical form of raw score distribution for better discrimination at the upper end.

Focus should be specifically placed on the reliabilities of the Kindergarten Screening Test. More experimental comparisons should be made with longer established instruments.

The instrument might be used on a pilot basis in early First Grade as a retest or alternate form for other established instruments.

Future investigation should include use of the instrument as a means of screening for maturation of the economically deprived child prior to his entrance in kindergarten in order that special provisions might be made for his enrichment prior to entering school.

Further investigation should explore the use of the instrument to test the emotionally disturbed, the neurologically handicapped and the slow-learning or so called educationally handicapped child.

Exploratory usage beyond Grade One should be attempted with the seriously educationally handicapped child.

The instrument should be used as an entrance screen for first grade students new to a district both before and after reading readiness.

Further experimental work should focus on use of the instrument in accelerating the judgments of able teachers.

Experimental work should also focus on enhancing the judgments of less able teachers, particularly in dealing with the less mature kindergarten child.

Extension of the experimental instrument through the insertion of brief subtests on the functions, (1) a performance test of comprehension by the following of directions and

(2) gross to finer scale of phonetic discrimination by audio methods, should be included. The entire augmented battery should then be tested.

BIBLIOGRAPHY

Passow, A. Harry (ed.). Curriculum Crossroads. New York: Bureau of Publications, Teachers College, Columbia University, 1962.

Risland, Henry D. A Basic Vocabulary of Elementary School Children. New York: The MacMillan Company, 1945.

Solomon, Herbert (ed.). Studies in Item Analysis and Prediction. Stanford, California: Stanford University Press, 1961.

Spindler, George (ed.). Education and Culture: Anthropological Approaches. New York: Holt, Rinehart and Winston, 1963.

Strickland, Ruth G. The Language Arts in the Elementary School. Boston: D. C. Heath and Company, 1951.

Strang, Ruth. An Introduction to Child Study. New York: The MacMillan Company, 1951.

Teacher's Guide to Education in Early Childhood. Sacramento, California: State Department of Education, 1956.

Warner, Ruby H. The Child and His Elementary School World. Englewood Cliffs, New Jersey: Prentice-Hall, Incorporated, 1957.

B. PERIODICALS

Davis, Frederick B. "Item Analysis in Relation to Educational and Psychological Testing," Psychological Bulletin, 49:97-118, March, 1952.

Howard, Kenneth I. and Garlie A. Forehand. "A Method for Correcting Item-Total Correlations for the Effect of Relevant Item Inclusion," Educational and Psychological Measurement, 22:731-35, Winter, 1962.

Jastak, Joseph. "An Item Analysis of the Wechsler-Bellevue Tests," Journal of Counseling Psychology, 14:88-94, April, 1950.

Kimball, Solon T. "Those First School Years," The National Elementary Principal, 40:18-32, September, 1960.

Meyers, Charles T. "The Relationship Between Item Difficulty and Test Validity and Reliability," Educational and Psychological Measurement, 22:565-71, Autumn, 1962.

Spodek, Bernard. "Developing Social Science Concepts in the Kindergarten," Social Education, 25:5, May, 1963.

Strong, Edward K. "Good and Poor Interest Items," Journal of Applied Psychology, 46:269-75, August, 1962.

Williams, Gertha. "The Kindergarten and Reading: Another Viewpoint," Childhood Education, October, 1963, pp. 77-78.

C. GOVERNMENT PUBLICATIONS

United States Bureau of the Census, United States Census of Population, 1960, California, General Social and Economic Characteristics. Washington: Government Printing Office, 1960.

D. TEXTBOOKS

Bond, Guy L. Three of Us, Play With Us. Sacramento, California: California State Department of Education, 1954.

Buchanan, Cynthia Dee. Programmed Primer. New York: McGraw-Hill Book Company, Incorporated, 1963.

Burton, William H. and others. Days of Fun. Sacramento, California: California State Department of Education, 1954.

Clark, John R. and others. Growth in Arithmetic, Grade I. New York: Harcourt, Brace and World, Incorporated, 1962.

Dawson, Don T. and Lovelle Downing. Arithmetic I. Englewood Cliffs, New Jersey: Prentice-Hall, Incorporated, 1963.

Eicholz, Robert E. and others. Elementary School Mathematics: Primer. Palo Alto, California: Addison-Wesley Publishing Company, 1963.

Gray, William S. and others. The New Pre-Primers. Chicago: Scott Foresman and Company, 1951.

Greater Cleveland Mathematics Program. Chicago: Science Research Associates, 1962.

Gunderson, Agnes G. and others. Learning to Use Arithmetic. New York: D. C. Heath and Company, 1962.

Gundlach, Bernard H. Arithmetic, Primer and Grade I. River Forest, Illinois: Laidlow Brothers, 1963.

Scott, Louise Binder and Virginia S. Pavelko. Time for Phonics. St. Louis: Webster Publishing Company, 1962.

Shane, Harold G. and Kathleen B. Hester. Tales to Read and Stories to Remember. River Forest, Illinois: Laidlow Brothers, 1960.

Russell, David H. and others. Under the Apple Tree and Play with Us. Boston: Ginn and Company, 1952.

E. TESTS

Ammons, Robert B. and H. S. Ammons. Ammons Picture Vocabulary Test. Missoula, Montana: Psychological Test Specialists, 1948.

_____. Quick Test. Missoula, Montana: Psychological Test Specialists, 1962.

Bender, Loretta. Visual Motor Gestalt Test. New York: American Orthopsychiatric Association, Incorporated, 1946.

Goodenough, Florence. Measurement of Intelligence by Drawings. New York: World Book Company, 1926.

Kent, Grace. Kent E. G. Y. Scales. New York: Psychological Corporation, 1946.

Lantz, Beatrice. Easel Age Scale. Los Angeles: California Test Bureau, 1955.

Lee, J. Murray and Willis W. Clark. Lee-Clark Reading Readiness Test. Monterey, California: California Test Bureau, 1962.

Pintner, R., B. U. Cunningham and W. N. Dorost. Pintner-Cunningham Primary Test. Yonkers-on-the-Hudson: World Book Company, 1940.

Readiness for Formal Work Test. Western Australia: Education Department, 1961.

Slosson, Richard L. Slosson Intelligence Test. East Aurora, New York: Slosson Educational Publications, 1963.

Sullivan, Elizabeth T., Willis W. Clark and Ernest W. Tiegs. California Test of Mental Maturity, Primary Form. Los Angeles: California Test Bureau, 1957.

Terman, Lewis M. and Maud A. Merrill. Stanford-Binet Intelligence Scale, Form L-M. Boston: Houghton-Mifflin Company, 1960.

Van Alstyne, Dorothy. Van Alstyne Picture Vocabulary Test. New York: Harcourt, Brace and World, 1961.

Wechsler, David. Wechsler Intelligence Scale for Children. New York: Psychological Corporation, 1949.

APPENDIX I

TABLE XX

DISTRIBUTION OF OCCUPATIONS OF MALE WAGE-EARNERS IN MODESTO*

Occupation	Modesto Total	Modesto Percentage
Professional, Technical, etc.	1,233	14%
Farmers and Managers	119	1
Managers, Officials, Proprietors	1,496	18
Clerical	490	5
Sales Workers	1,128	12
Craftsmen, Foremen	1,302	15
Operatives	1,206	13
Private Household	0	0
Service	515	6
Farm Laborers	187	2
Other Laborers	527	7
Occupation not reported	533	7

*This does not include an approximate population of 30,000 within the immediate environs.

TABLE XXI

CLASSIFICATION OF VOCABULARY WORDS USED IN UNREFINED EXPERIMENTAL INSTRUMENT

Word	Symbol	Word	Symbol	Word	Symbol
carrots	1b5	something burning	2b	toy	1a3
cherries	2a	something sharp	2b	wire	2b
grapes	2b	something to eat	1a2	someone sleeping	2b
beans	2a	saw	1a1	umbrella	1b5
bananas	2a	hammer	2b	book	1a2
orange	1a2	screwdriver	---	hitting	---
corn	1a5	pliers	---	wheel	1b5
roots	3a	shovel	2a	hammock	---
seeds	1b4	rake	2b	motorcycle	3b
bunch	2a	ruler	3b	something to wear	1a3
train	1a2	nail	3a	something to drink	1a4
lamp	3b	sprinkler	---	candle	1b5
dust pan	---	suitcase	3b	globe	---
vase	2b	telephone	2a	pocket	2a
flower	1b1	ribbon	2a	picture	1a3
propellers	---	feet	1a4	bicycle	1b1
leaves	1a4	hair	1a3	fence	2a
something steaming	---	necklace	---	mower	---
feather	3b	dial	---	post	2b
faucet	---	cap	1a4	someone leaning	---
envelope	2b	handle	2b	spokes	---
plate	2a	leg	1b4	motor	2a
match	2a				
light	1a5				

NOTE: The symbols used in the table are based on the University of Oklahoma Study conducted in 1945 classifying words used in original written stories.

Interpretation of symbols used:

Symbol	Group found within
1a1 to 1a5	1st 100 of the 1st 500 of the 1st 1,000 words
1b1 to 1b5	1st 100 of the 2nd 500 of the 1st 1,000 words
2a to 5b	1st $\frac{1}{2}$ of the 2nd 1,000 words

TABLE XXII

RELATIONSHIP BETWEEN TEST ITEMS

Test Section	Item	Correlation	Test Section	Item	Correlation
I	square	.57	II	something to eat	.50
	triangle	.67		saw	.36
	reversed p	.44		hammer*	0
	lower case b	.62		screwdriver	.44
	pie	.55		pliers	.59
	ovals	.26		shovel	.69
	diamond	.58		rake	1.01
	W	.52		ruler	.38
II	carrots	.36		nail	.60
	cherries	.44		sprinkler	.66
	grapes	.34		suitcase*	0
	beans	.55		telephone*	0
	bananas	.42		ribbon	.38
	orange	.59		feet	.24
	corn	.20		hair	1.10
	roots	.42		necklace	.29
	seeds*	.16		dial	.43
	bunch*	.04		cap	.49
	train*	.49		handle	.50
	lamp	.51		leg*	.17
	dustpan	.22		toy	.48
	vase	.57		wire	.85
	flower	.56		someone sleeping	.30
	propellers	.47		umbrella	.63
	leaves*	.11		book	.86
	something steaming	.47		hitting	.37
	feather	.14		wheel	.37
	faucet	.30		hammock	.32
	envelope	.57		motorcycle*	0
	plate*	0		something to wear	.37
	match	.30		something to drink	.29
	light	.36		candle	.64
	something burning	.40		globe	.46
	something sharp	.40			

TABLE XXII (continued)

Test Section	Item	Correlation	Test Section	Item	Correlation
II	pocket	.46	III	three	.43
	picture	.35		eight	.51
	bicycle*	0		largest	.66
	fence	.22		smallest	.69
	mower*	0		tallest	.47
	post*	0		shortest	.51
	someone leaning	.21		sand pail*	1.06
	spokes	.30		wagon	.46
	motor	.47		gingerbread boy	.46
III	red	.67		sailboat	.35
	blue	.76		triangle*	0
	orange	.66		cube	.24
	green	.94		moon	.36
	yellow	.92		flower	.37
	two	.67		in front of	.44
	one	.45		in back of	.49
	four	.65		on top of	.70
	six	.65		inside	.35

NOTE: Relationship was determined by use of the bi-serial r formula: $\frac{Mp - Mt}{t} \times \frac{P}{Y} =$

*Indicates discarded items

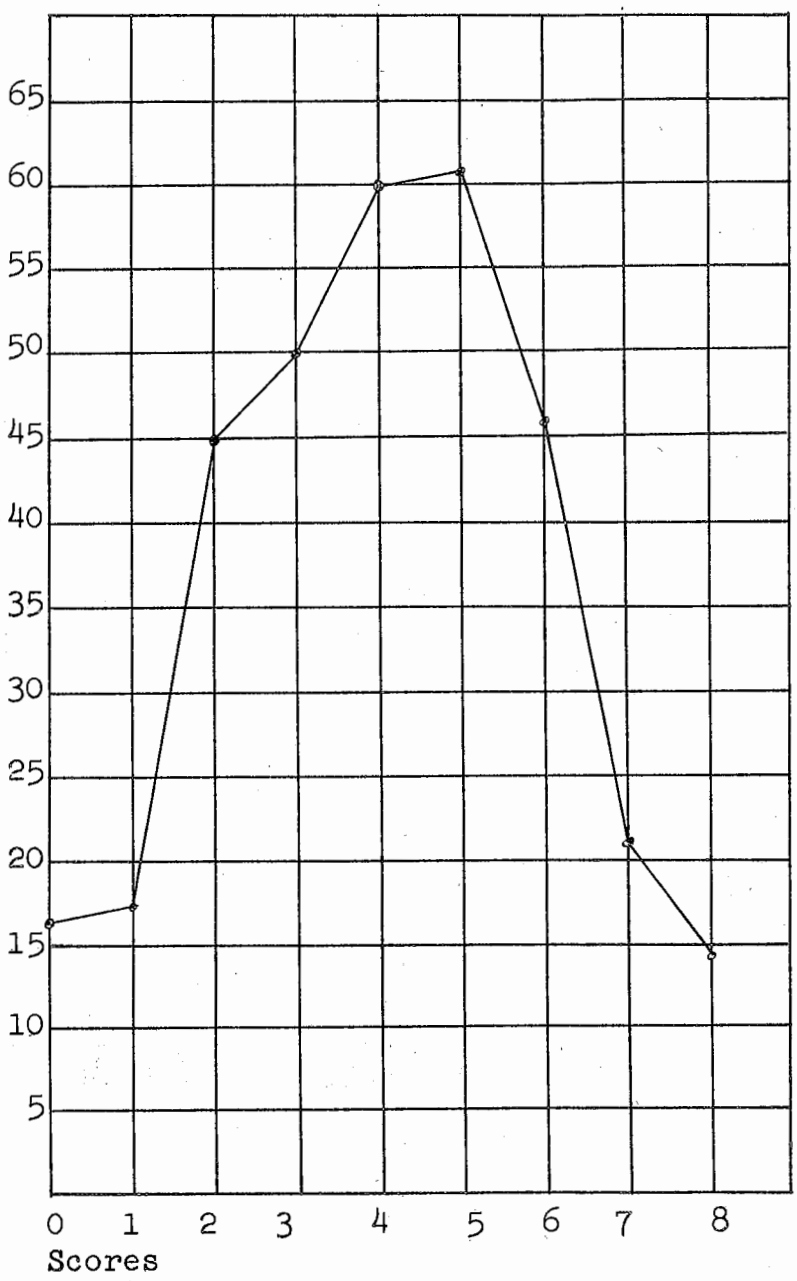


FIGURE 1

DISTRIBUTION OF RAW SCORES ON KINDERGARTEN SCREENING TEST
SECTION ONE: VISUAL-MOTOR COORDINATION

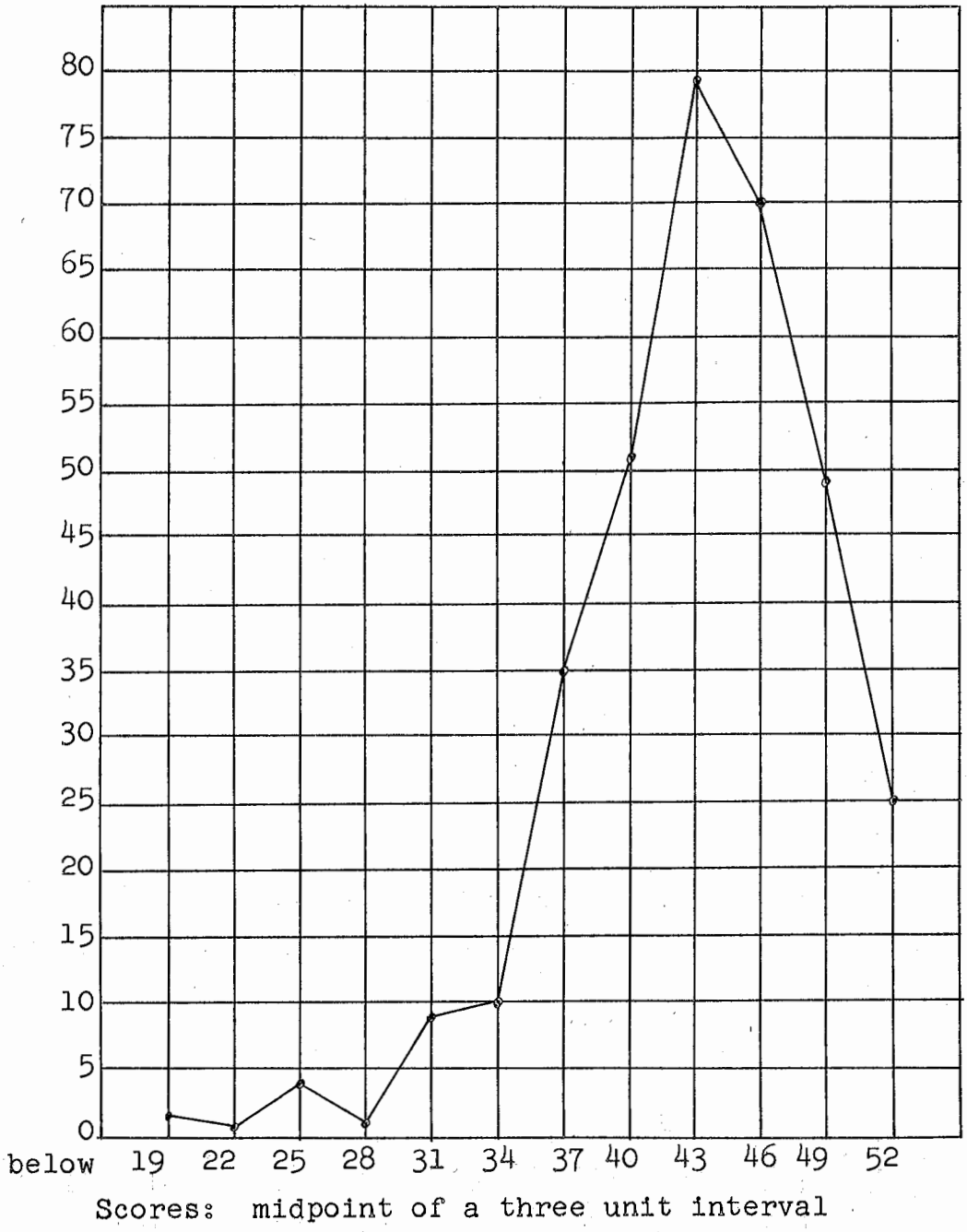


FIGURE 2

DISTRIBUTION OF RAW SCORES ON KINDERGARTEN SCREENING TEST
SECTION TWO: VOCABULARY

APPENDIX II

A KINDERGARTEN SCREENING TEST

DEVISED BY MARY McGUIRE

1964

PART ONE

TEST DESCRIPTION

Section 1: consists of eight drawings arranged according to difficulty. Each design is individually presented and the child reproduces the design as it appears to him. This subtest involves gross motor coordination, eye-hand coordination, and form perception.

Section 2: is composed of eight pictures. Each picture consists of several pictured vocabulary items and some pictures are complete. This subtest is based on the premise that identification of vocabulary words through recognition may be more efficiently measured by the child's pointing to the word than through verbal response. The use of complete pictures requires the ability to discriminate details.

Section 3: consists of four pictures. These items measure the child's understanding of the concepts of color, number, size, differences, similarities, and position in space.

This instrument may be administered in approximately ten minutes. Administration should occur in a quiet place with only the child and the examiner present. In administering the test, directions should be presented exactly as they appear in the Directions for Administration and no additional help should be given.

Materials needed:

(1) Record Form

(2) Primary pencil

(3) Manual of Instructions

(4) Set of eight pictures for design reproduction

(5) Set of eight vocabulary pictures

(6) Set of four concept pictures

PART TWO

DIRECTIONS FOR ADMINISTRATION

Make sure all necessary information has been recorded on the Record Form before starting to administer the test.

Section 1: Designs should be arranged in numerical order with the blank card covering all designs. Each design will be presented singly and will remain in front of the child while he is reproducing the design. Place the Record Form and pencil in front of the child and say: I am going to show you some pictures and I want you to draw what you see. Uncover Design 1 and say: See this box? Make one like it. Make it here. (Indicate upper left hand box.) Next uncover Design 2 and say: Now make one like this. Make it here. (Point to next box.) Continue in the same manner for the remainder of the designs.

Section 2: Say: That was fine. Now I am going to show you some more pictures. (pause) I will say a word out loud and I want you to put your finger on that word in the picture. Present Card 1 and say: Carrots. (pause) Put your finger on the carrots. Bananas. Show me the bananas. Continue either of the above directions as long as the child needs the specific direction in order to respond. Many children will respond to word three on its presentation. Each card is presented separately and is removed from the child's field of vision after use.

VOCABULARY LIST

Picture One

1. carrots
2. bananas
3. grapes
4. corn
5. orange
6. cherries
7 beans
8. roots

Picture Two

9. flower
10. something steaming
11. lamp
12. dust pan
13. propellers
14. vase

Picture Three

15. something burning
16. feather
17. faucet
18. match
19. light
20. something to eat
21. something sharp

Picture Four

22. shovel
23. sprinkler
24. nail
25. rake
26. saw
27. pliers
28. screwdriver
29. ruler

Picture Five

30. necklace
31. hair
32. feet
33. wire
34. toy
35. ribbon
36. handle
37. dial
38. cap

Picture Six

39. book
40. umbrella
41. wheel
42. someone sleeping
43. hitting
44. hammock

Picture Seven

45. candle
46. something to wear
47. pocket
48. picture
49. something to drink
50. globe

Picture Eight

51. fence
52. spokes
53. someone leaning
54. motor

Section 3: Present the first Concepts Card and pointing to the "red" kite say: What color is this kite? Continue with: What color is this one? Move to the next row and say: How many dots are there in this box? Continue with: How many in this one? Point to all three balls in the next row and say: Which ball is largest? Which one is smallest? Point to all three trees and say: Which tree is tallest? Which one is shortest?

Present Concept: Differences Card. Point to all four pails and say: Which sand pail does not look like the others? Continue with: Which wagon is not like the others? Which gingerbread boy is not like the others? Which sailboat is not like the others?

Present Concept: Likenesses Card. Point to the square, triangle and circle, then to the triangle at the left and say: Which one looks like this one? This item is not scored. Continue using the same directions.

Present Concept: Position in Space Card and say: Look at this picture carefully and tell me the name of something that is in front of the house. Tell me something in back of. Something on top of. Something inside.

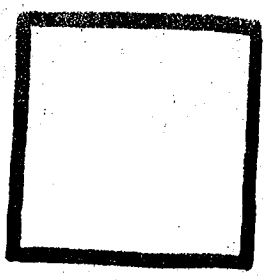


FIGURE 1

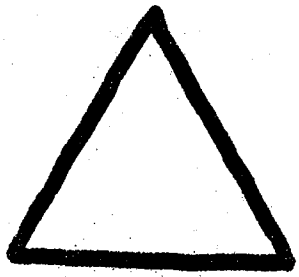


FIGURE 2

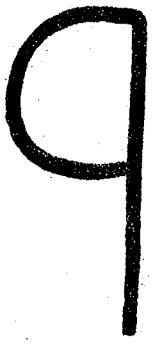


FIGURE 3

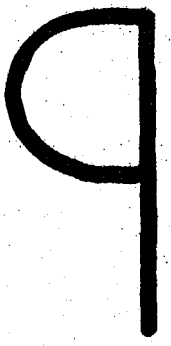


FIGURE 4

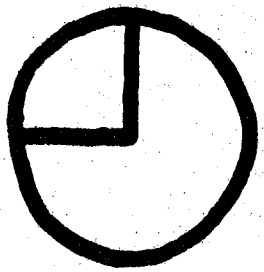


FIGURE 5

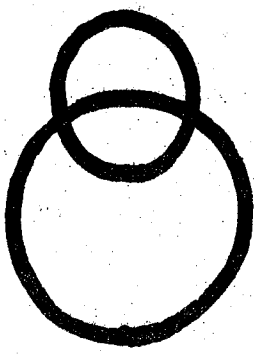


FIGURE 6

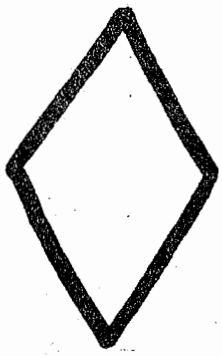


FIGURE 7

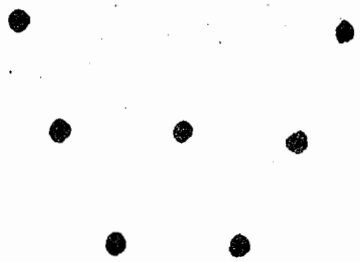


FIGURE 8

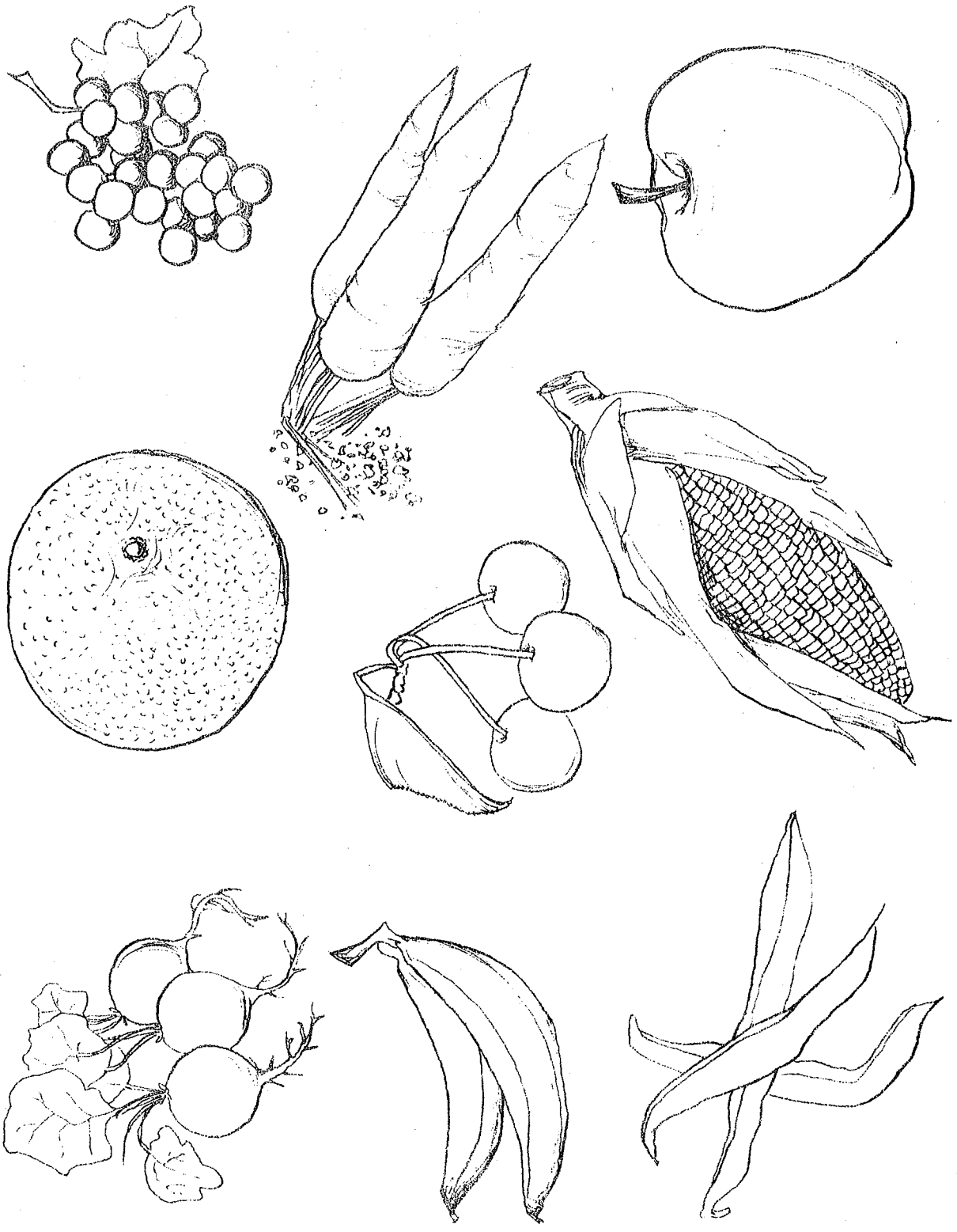


FIGURE 9

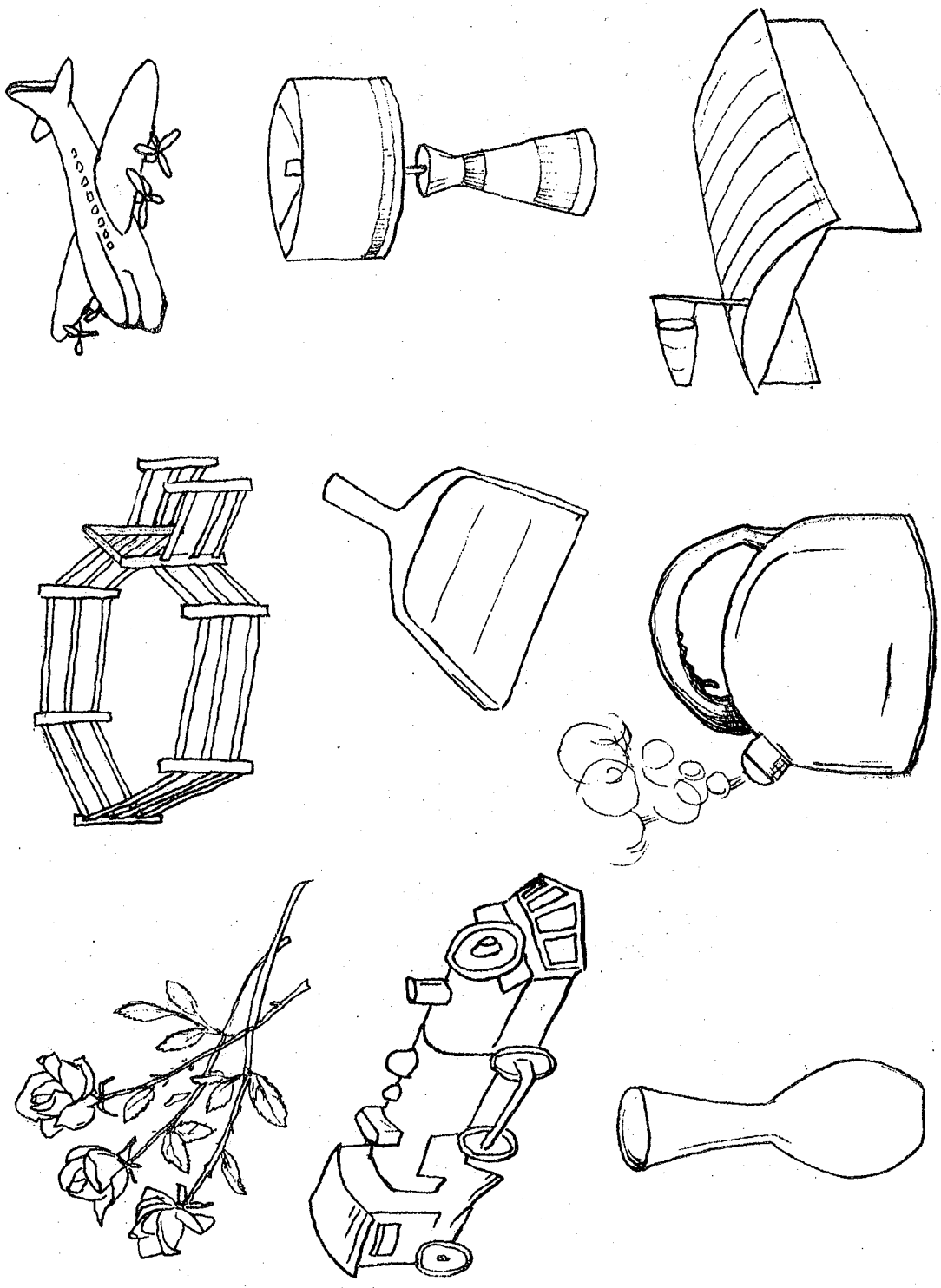


FIGURE 10

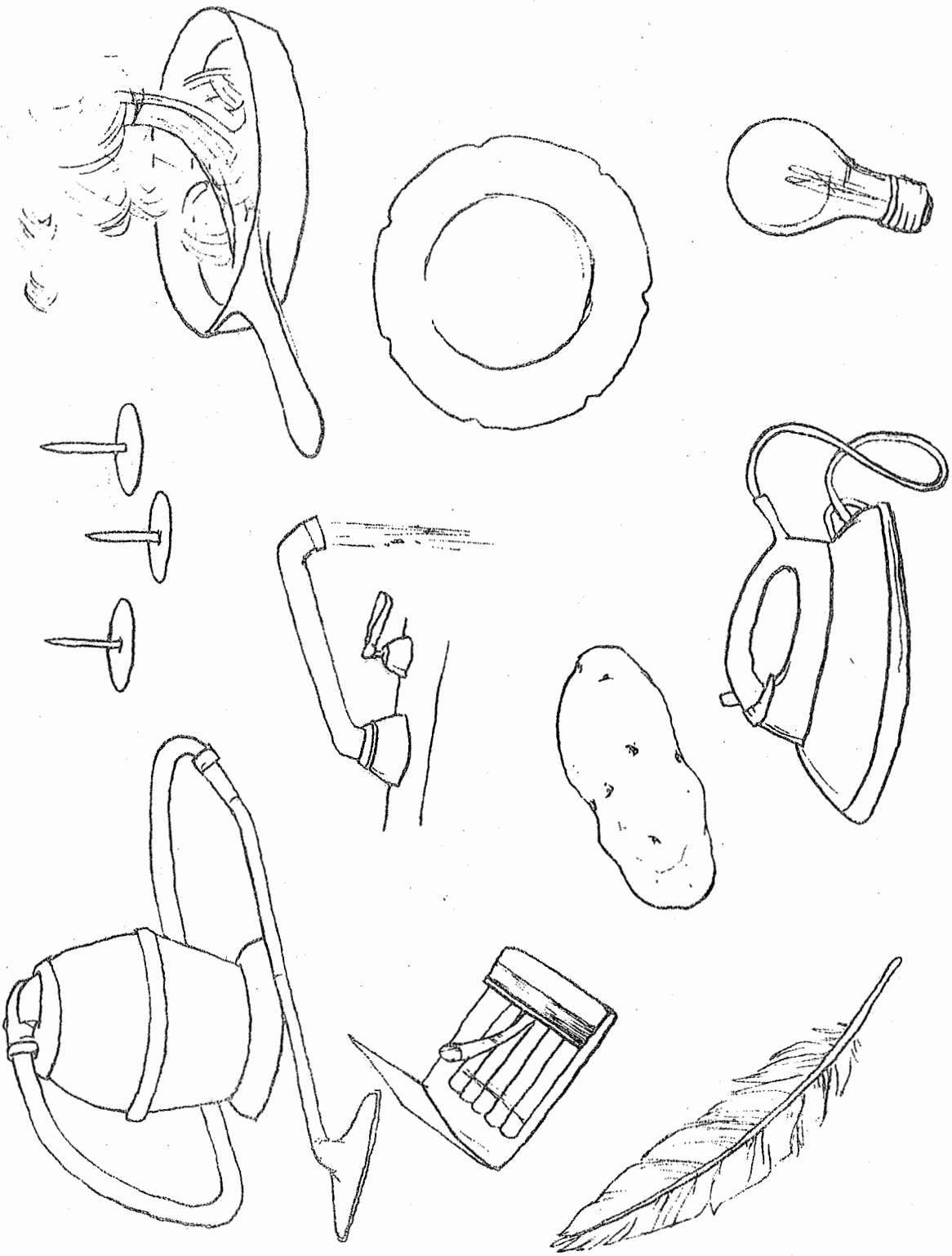


FIGURE 11

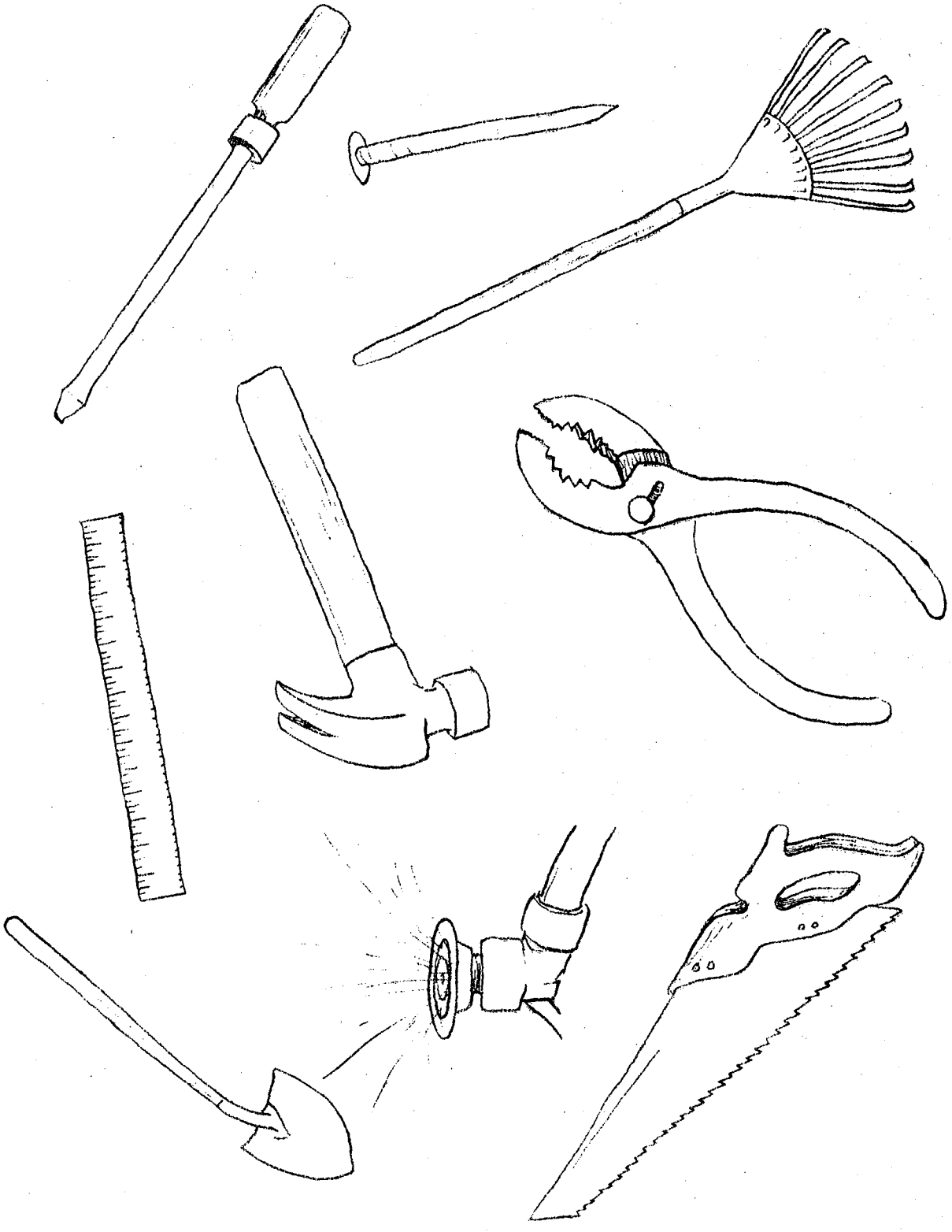


FIGURE 12

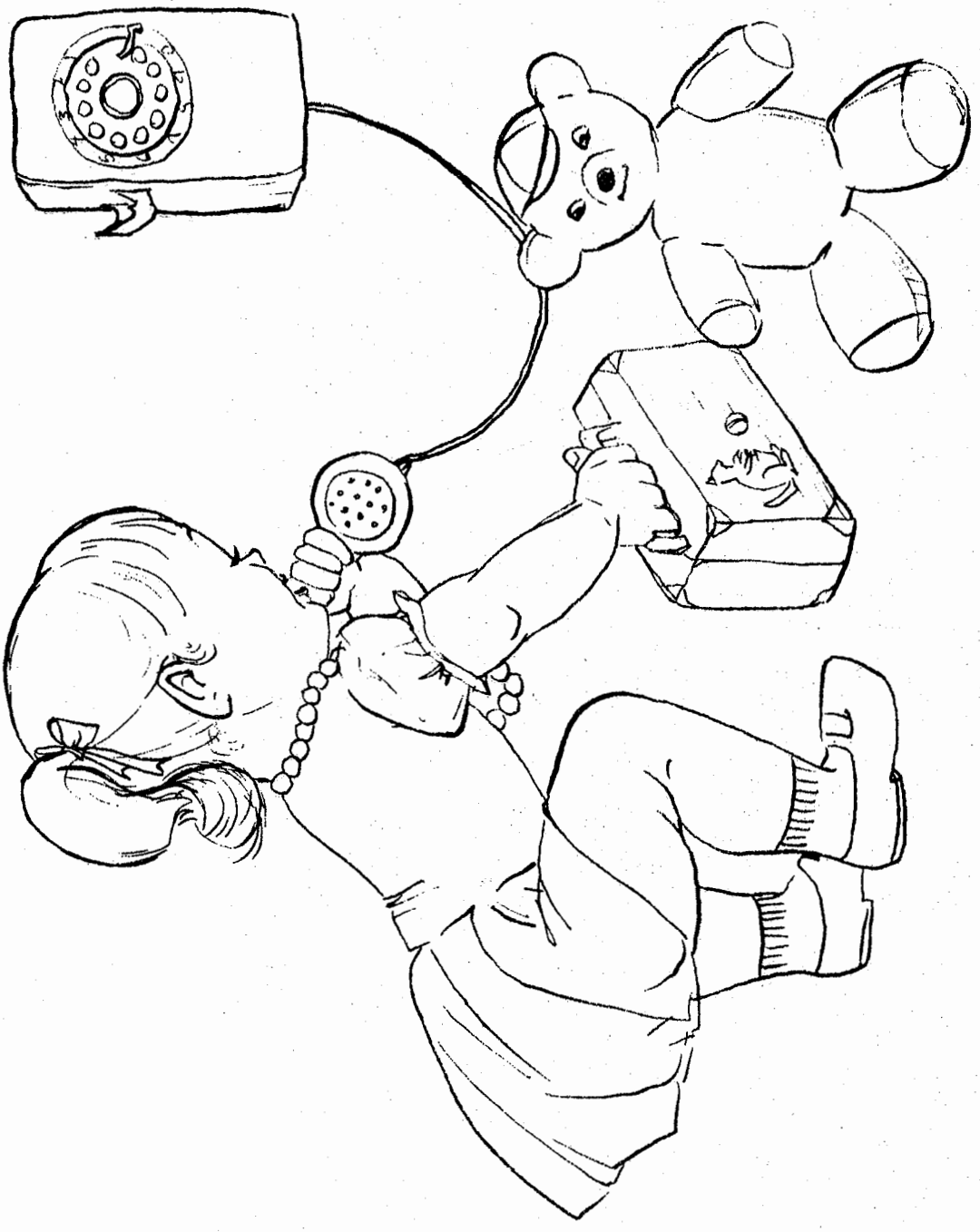


FIGURE 13

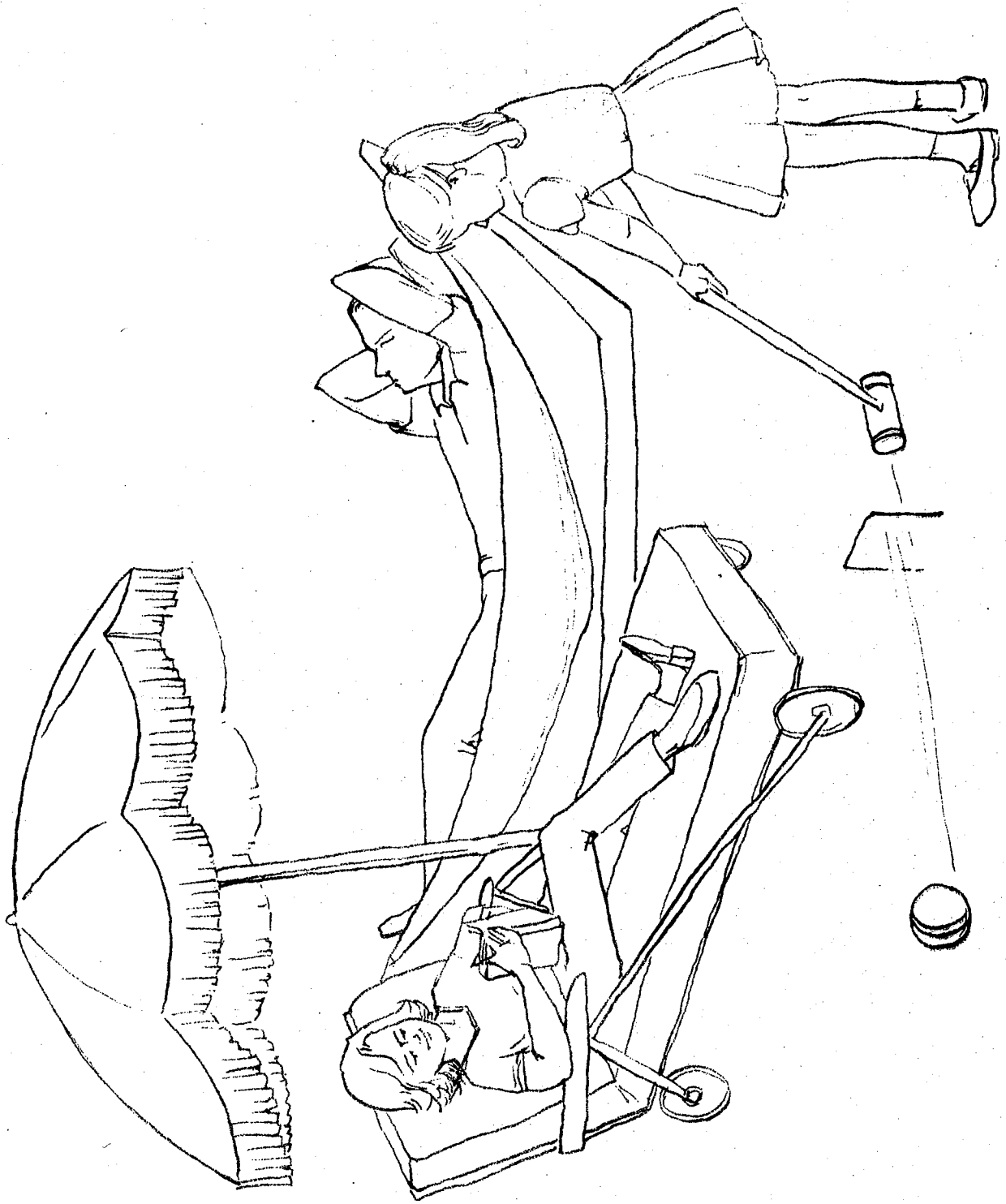


FIGURE 14

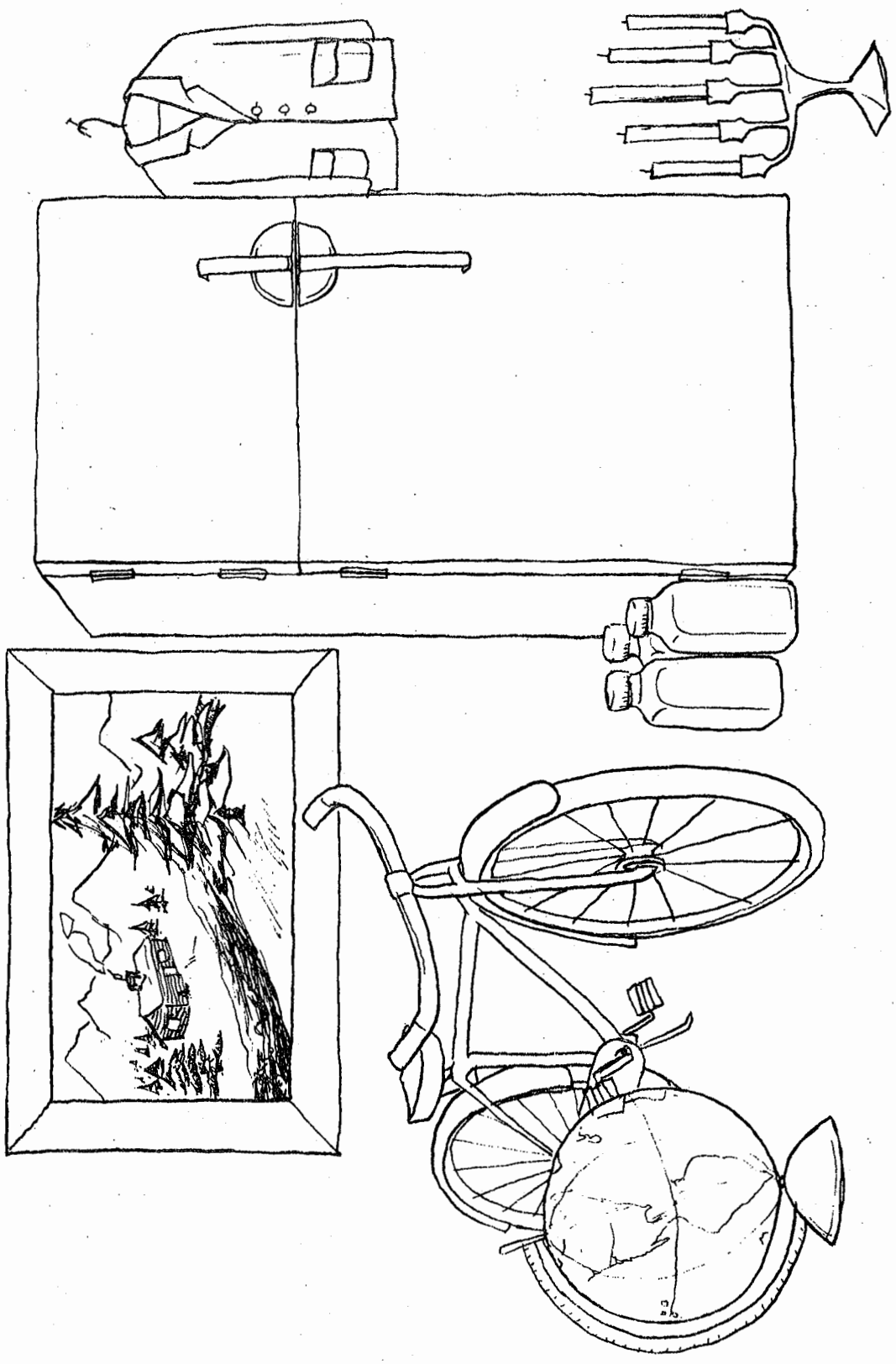


FIGURE 15

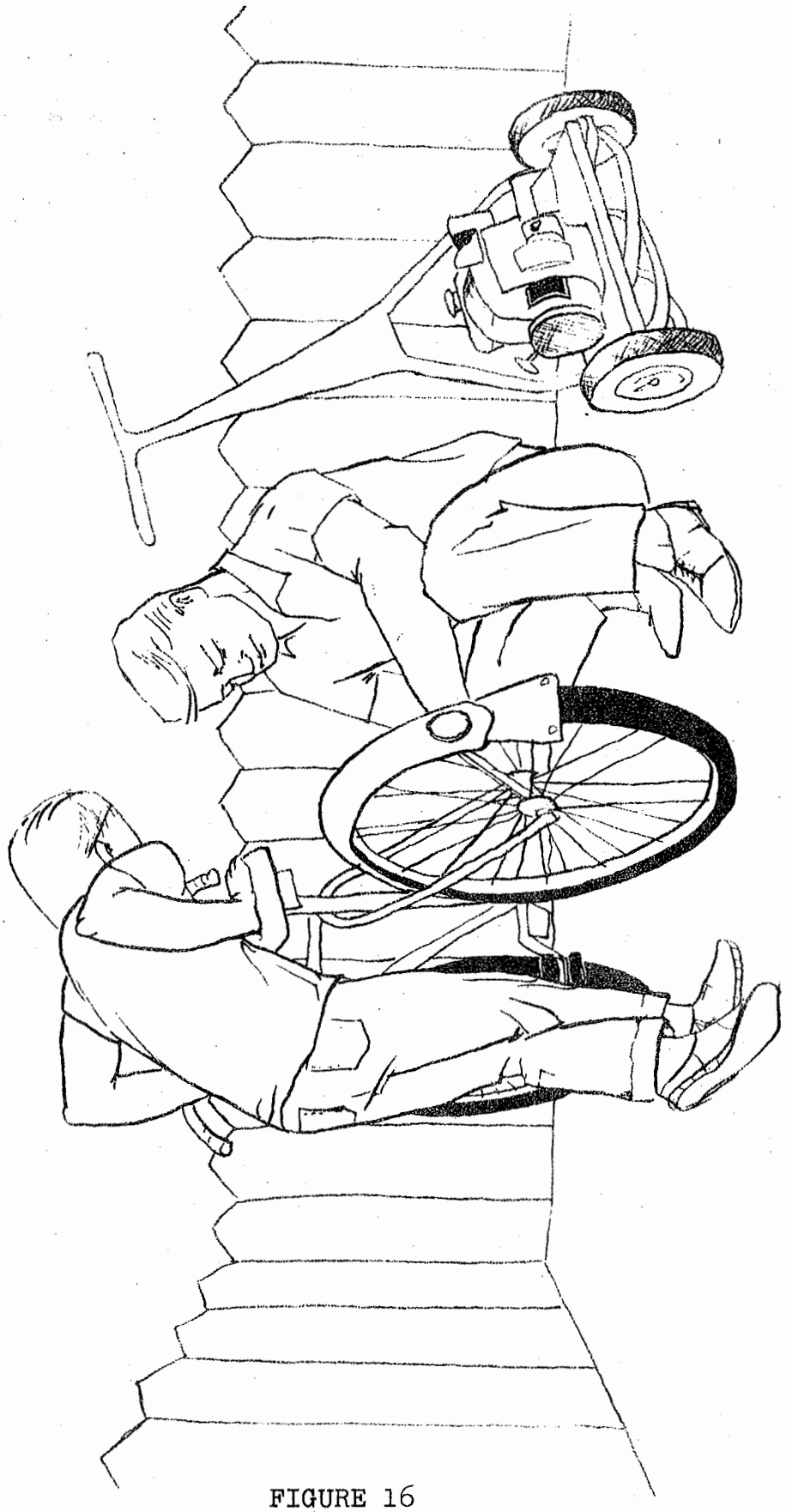


FIGURE 16

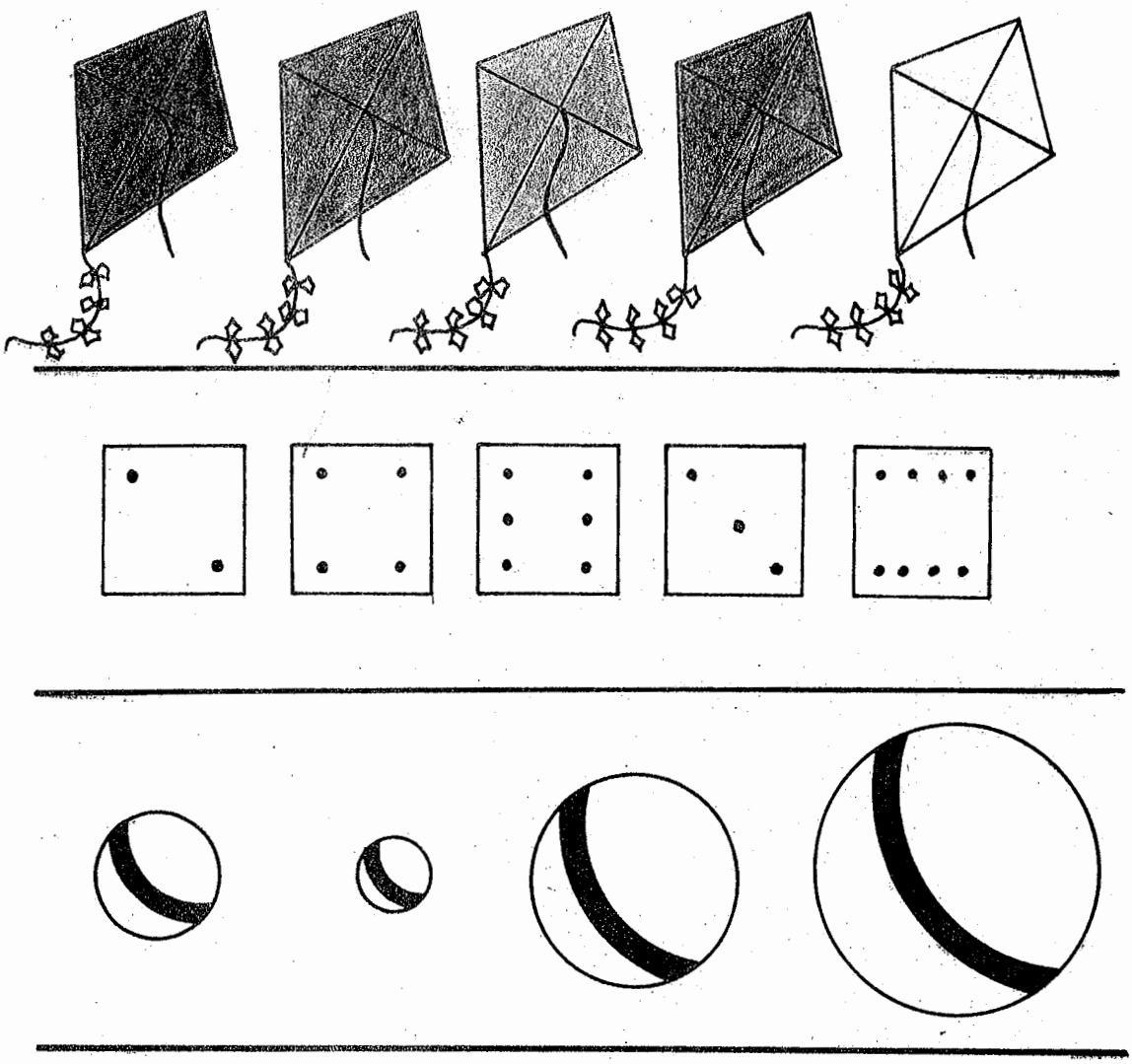


FIGURE 17

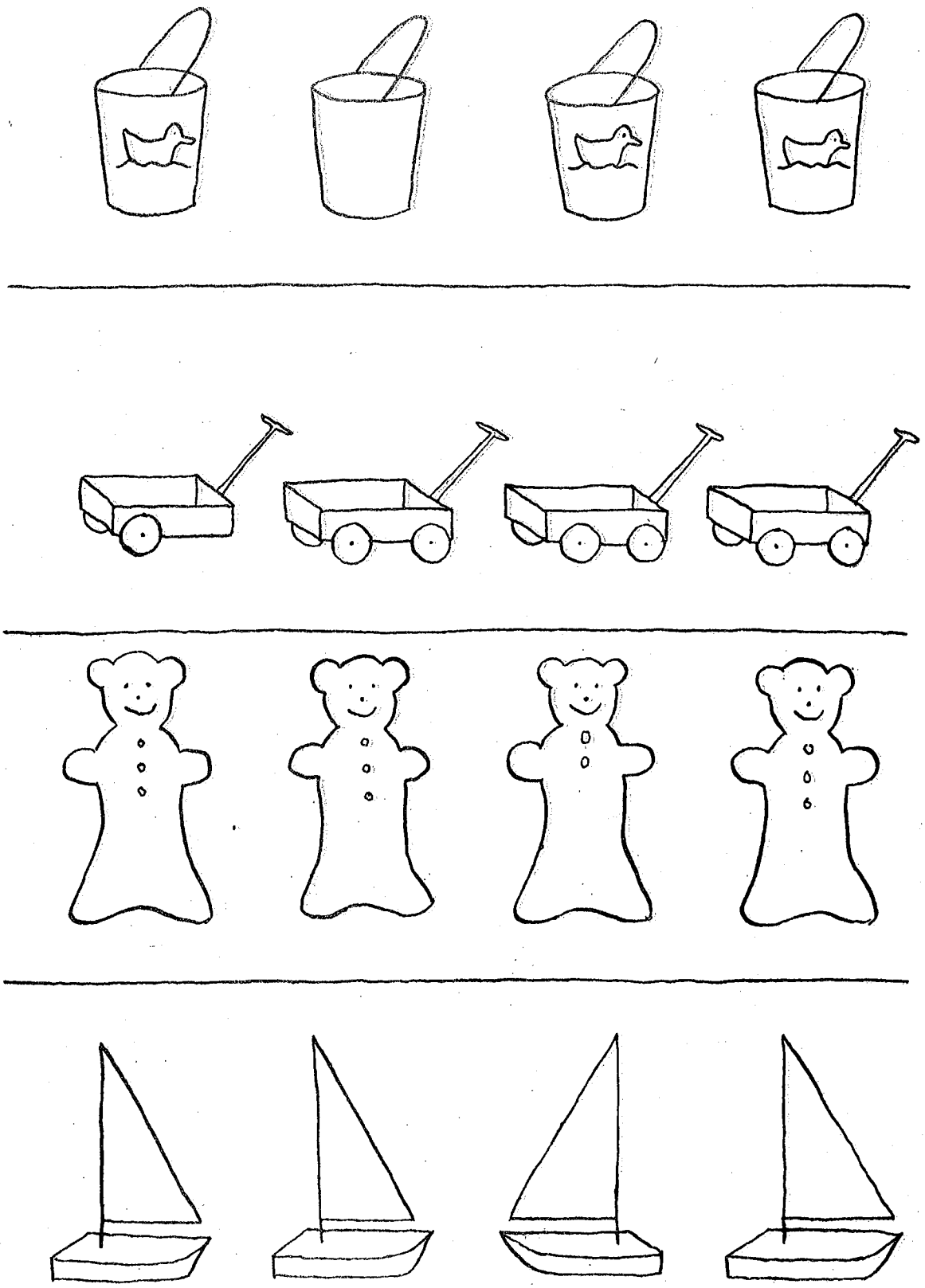


FIGURE 18

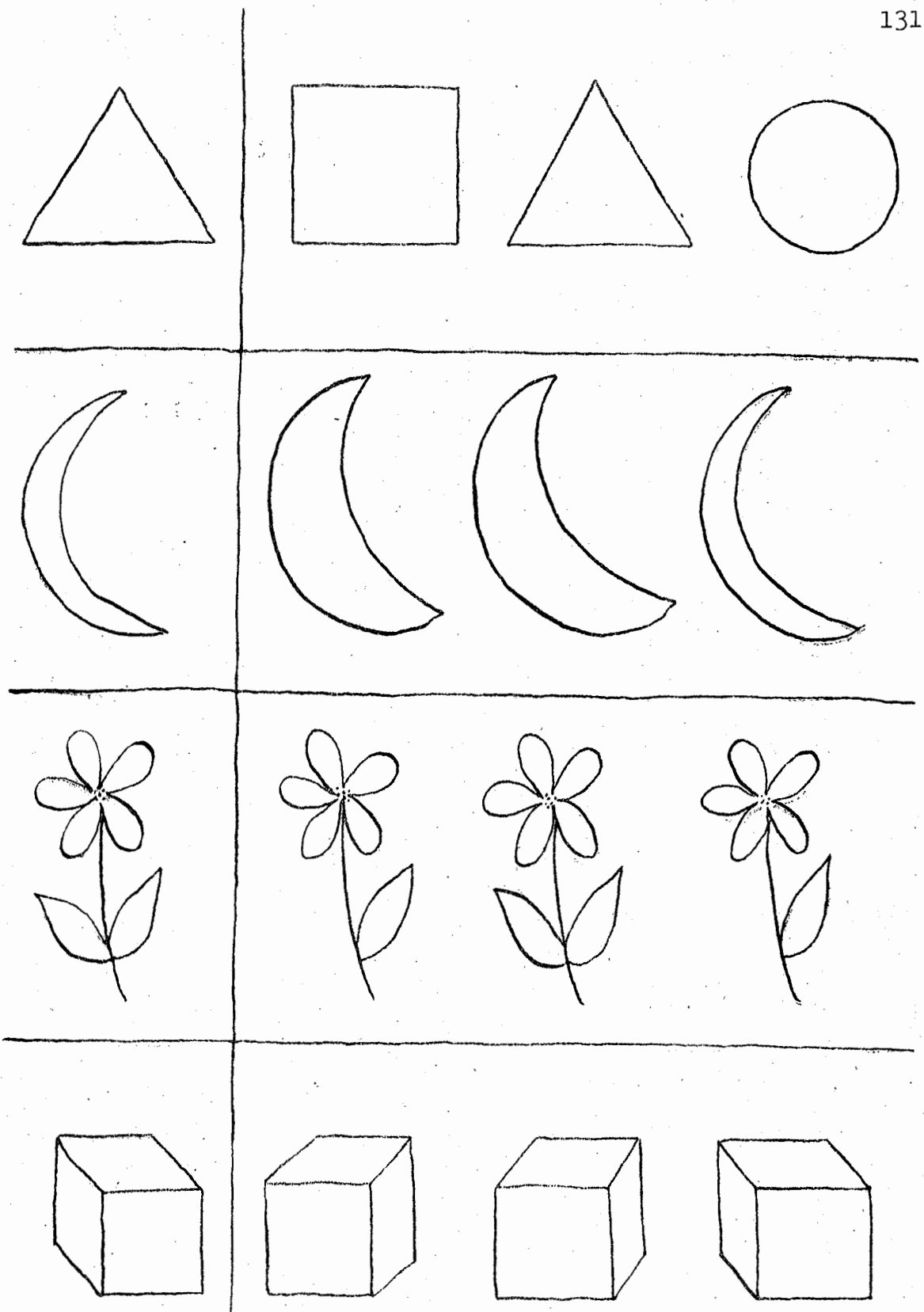


FIGURE 19

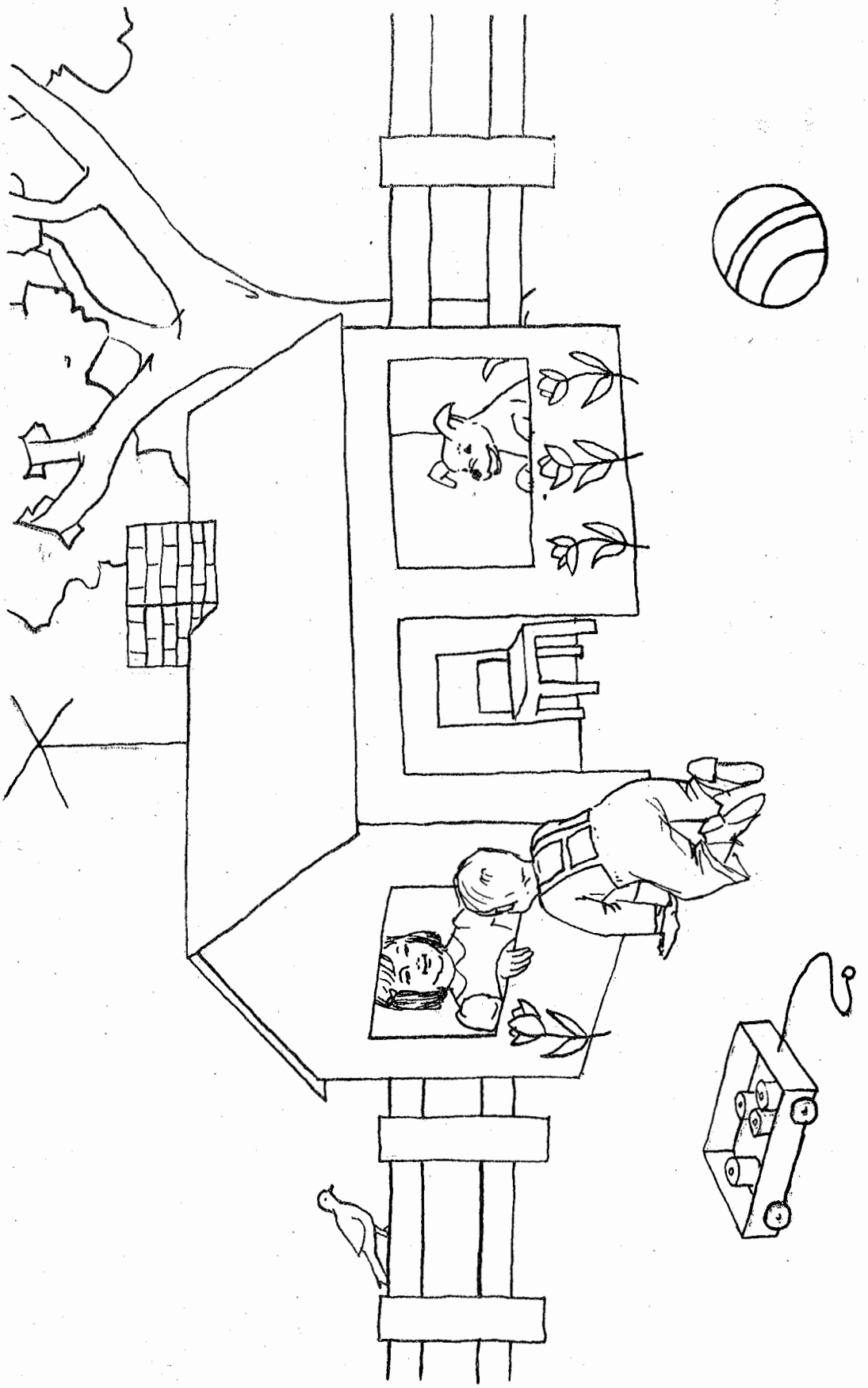


FIGURE 20

PART FOUR

SCORING

Record the correct response on the Record Blank by placing a check mark after each correct response except Concept: Position in Space where the child's response is recorded.

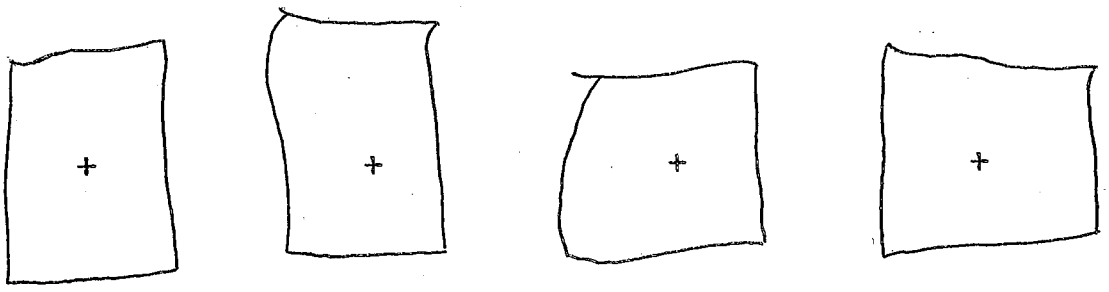
Each correct answer is scored as one point:

Possible Scores

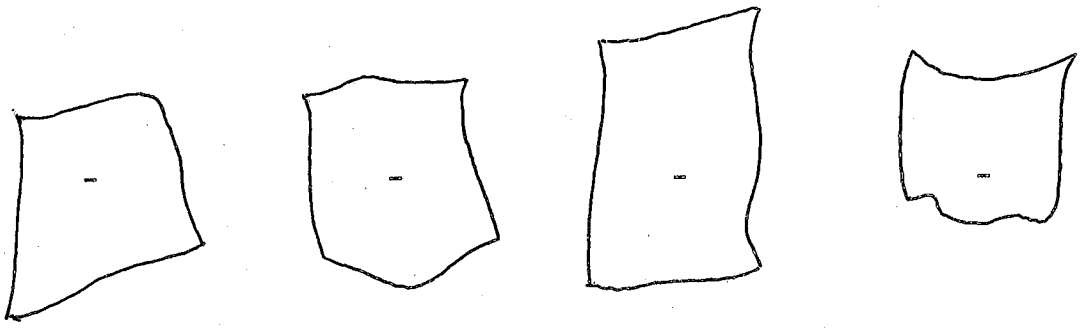
Section 1:	8
Section 2:	54
Section 3:	24
Total:	86

Section 1:

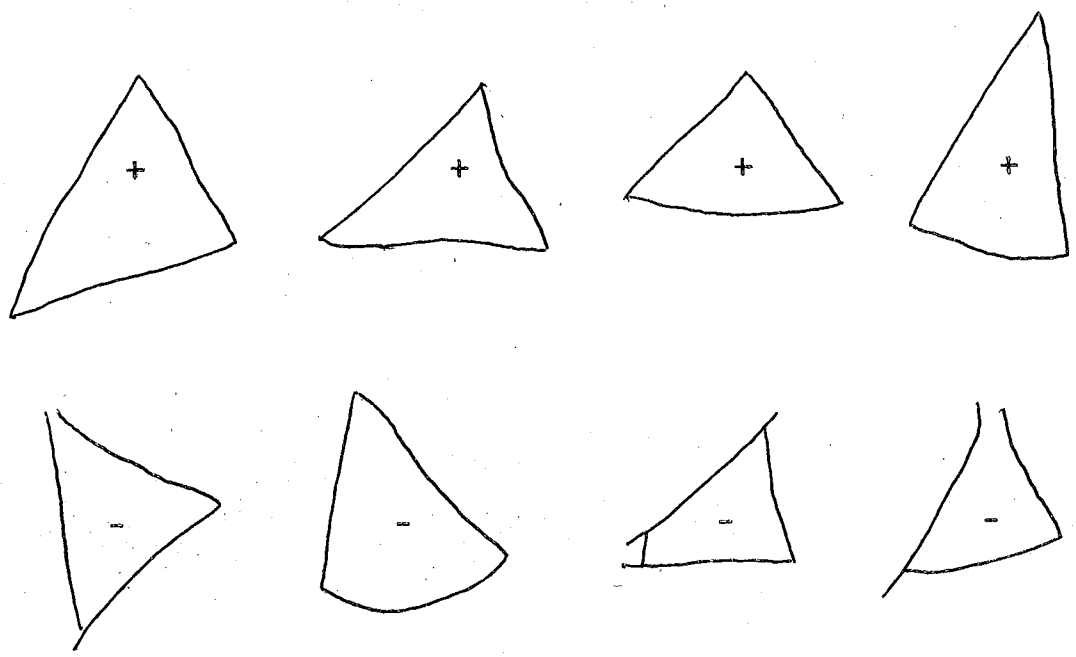
1. Square * - Lines should be unbroken and the figure should not be more than half again as long as it is wide. Corners should approximate right angles and not be rounded.
   Examples:



* This design is found in the Stanford-Binet Intelligence Scale, Form L-M, Year 5, and is scored in a similar manner.

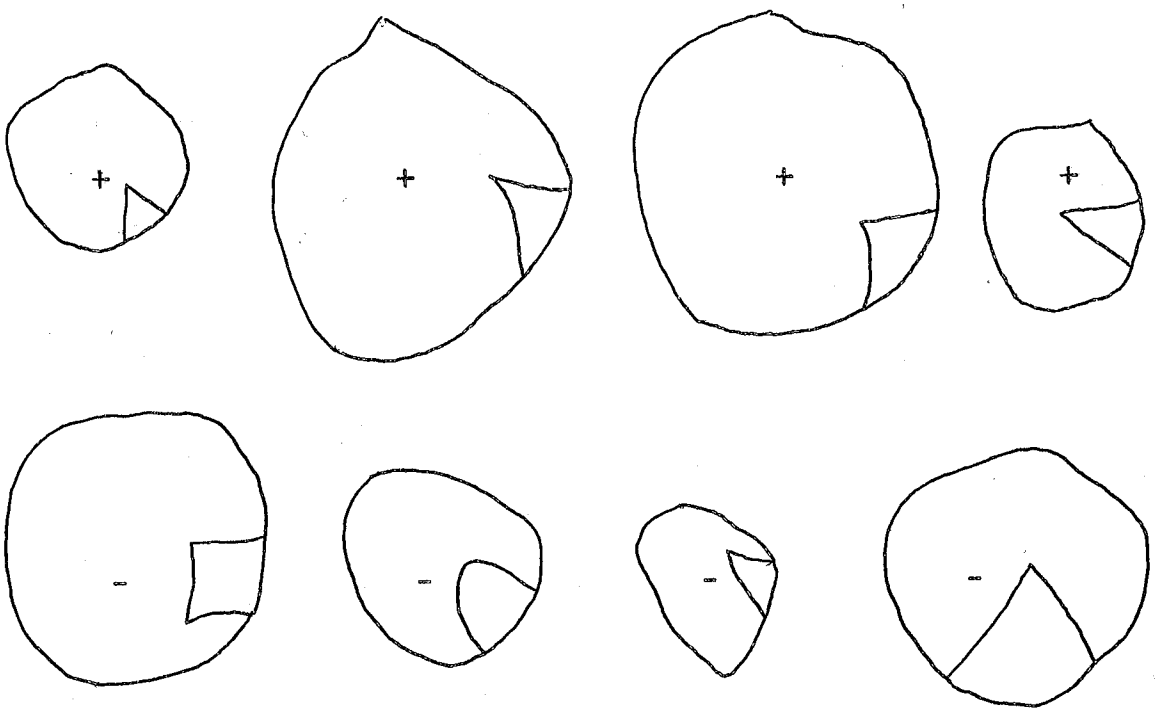


2. Equilateral triangle - Lines should be unbroken and of the same approximate length. Angles must be at each line intersection. The base line does not have to be exactly straight but should not be rounded.

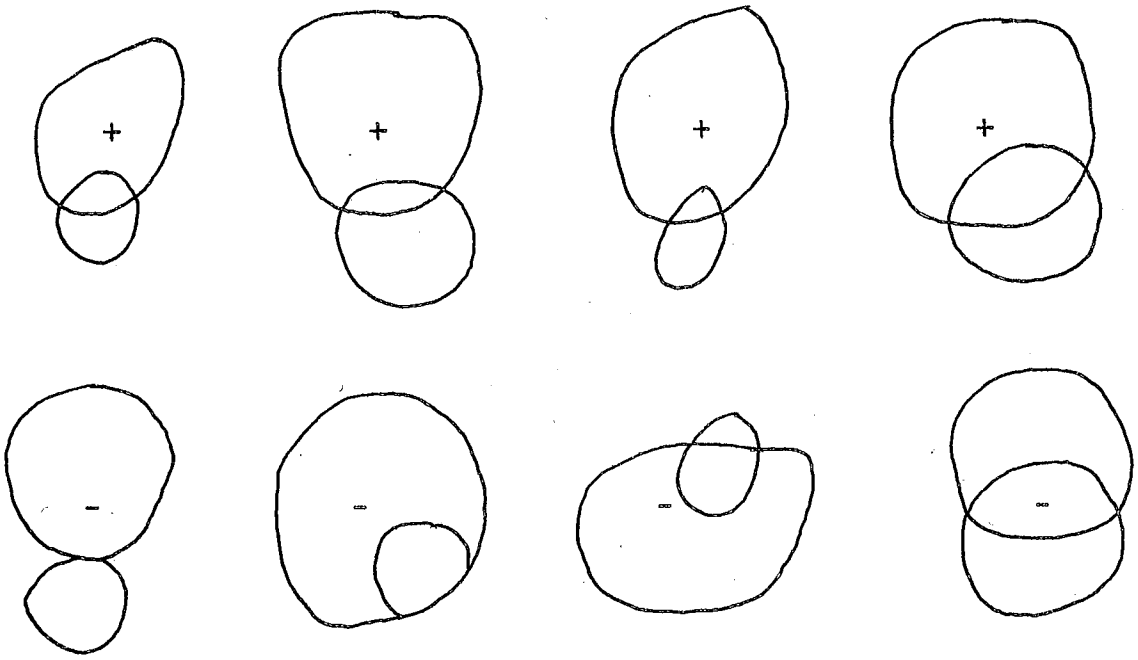


3. and 4. - Reversed P and lower-case b - These designs are included to indicate reversals. The placement of the oval on the line is considered rather than size of oval.

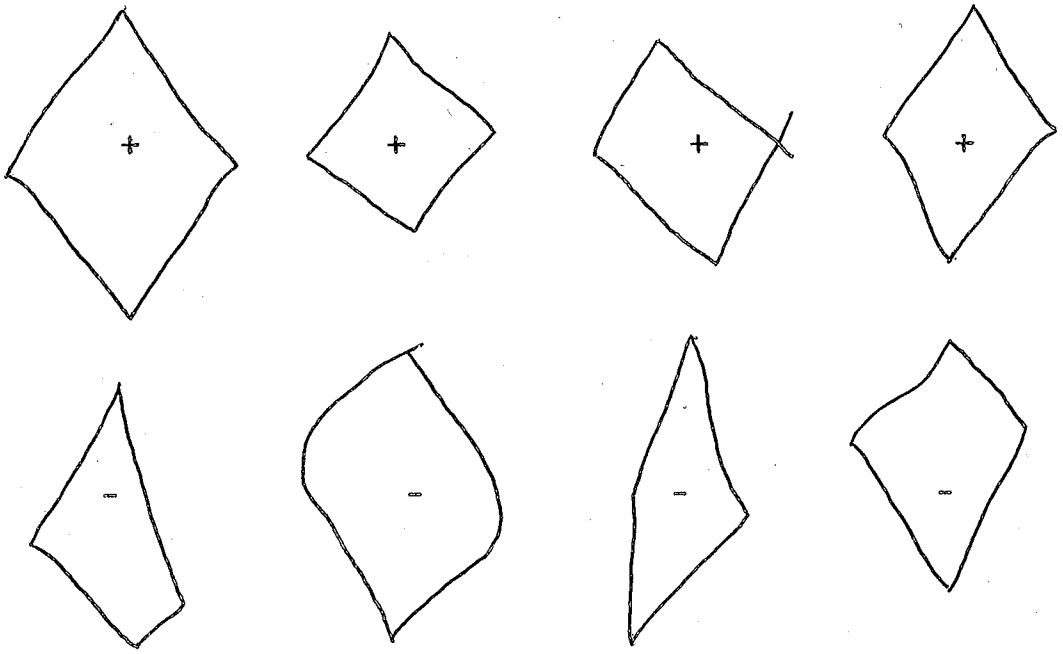
5. Pie - The outside line should be continuous and round rather than oval-shaped. The segment does not have to be one-fourth of the circle but must be wedge-shaped and within the prescribed area.



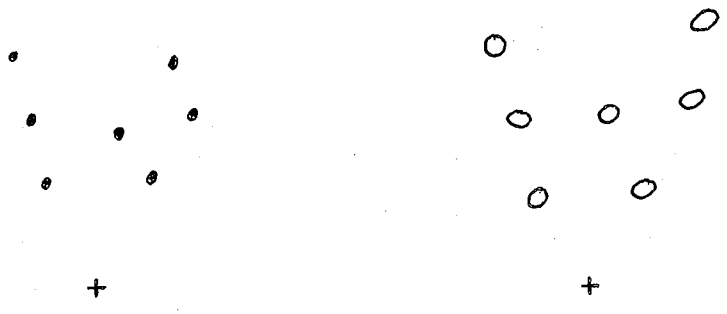
6. Overlapping ovals - This design may be circular as well as oval. The top oval must be larger than the bottom and overlapping must be present.



7. Diamond * - This design must have four well-defined angles: each pair of angles should be approximately opposite. It must be more diamond-shaped than square or kite-shaped.

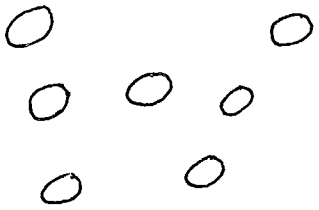


8. W - This design may consist of dots or small circles but must be in the shape of the design and consist of the correct number of dots.



* This design is found in the Stanford-Binet Intelligence Scale, Form L-M, Year 7, and is scored in a similar manner.

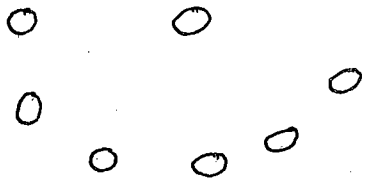
+



+

-

-



-

-

Section 2:

1. carrots
2. bananas
3. grapes
4. corn
5. orange
6. cherries
7. beans
8. beets or carrots
9. flowers
10. teakettle
11. lamp
12. dust pan
13. airplane
14. vase
15. frying pan
16. feather
17. running faucet
18. match
19. light bulb
20. potato
21. thumb tacks
22. shovel
23. sprinkler head
24. nail
25. rake
26. saw
27. pliers
28. screwdriver
29. ruler
30. child's necklace
31. hair
32. child's feet
33. telephone wire
34. animal
35. ribbon on girl's hair
36. handle on lunch box or telephone
37. telephone dial
38. cap on animal
39. book
40. lawn umbrella
41. wheel on chaise lounge

42. man
43. girl or mallet
44. hammock
45. any candle or candle holder
46. coat
47. pocket in coat
48. picture on wall
49. milk bottle
50. globe of world
51. any part of fence
52. bicycle spokes
53. boy on left
54. motor on mower

Section 3:

1. Color: red, blue, orange, green, yellow

2. Number: 2, 4, 6, 3, 8

3. Size (ball) largest - at extreme right
smallest - second from left
tallest - second from left
shortest - extreme left

4. Differences: second sand pail
first wagon
third gingerbread boy
third sailboat

5. Likenesses: triangle (sample)
third moon
second flower
third cube

6. Position in Space:

in front of: toy, ball, boy, flower
in back of: fence, bird, tree, antenna
on top of: chimney, roof, bird, antenna
inside: girl, chair, dog

PART FIVE

KINDERGARTEN SCREENING TEST

(Individual)

PUPIL'S RECORD FORM

Name_____ Circle One: Boy Girl
Last First Middle

Date of Test_____School_____

Date of Birth_____Examiner_____

Pupil's Age_____

	Possible Score	Pupil's Score
Visual-Motor Coordination	8	_____
Vocabulary	54	_____
Basic Concepts	24	_____
Total	86	_____

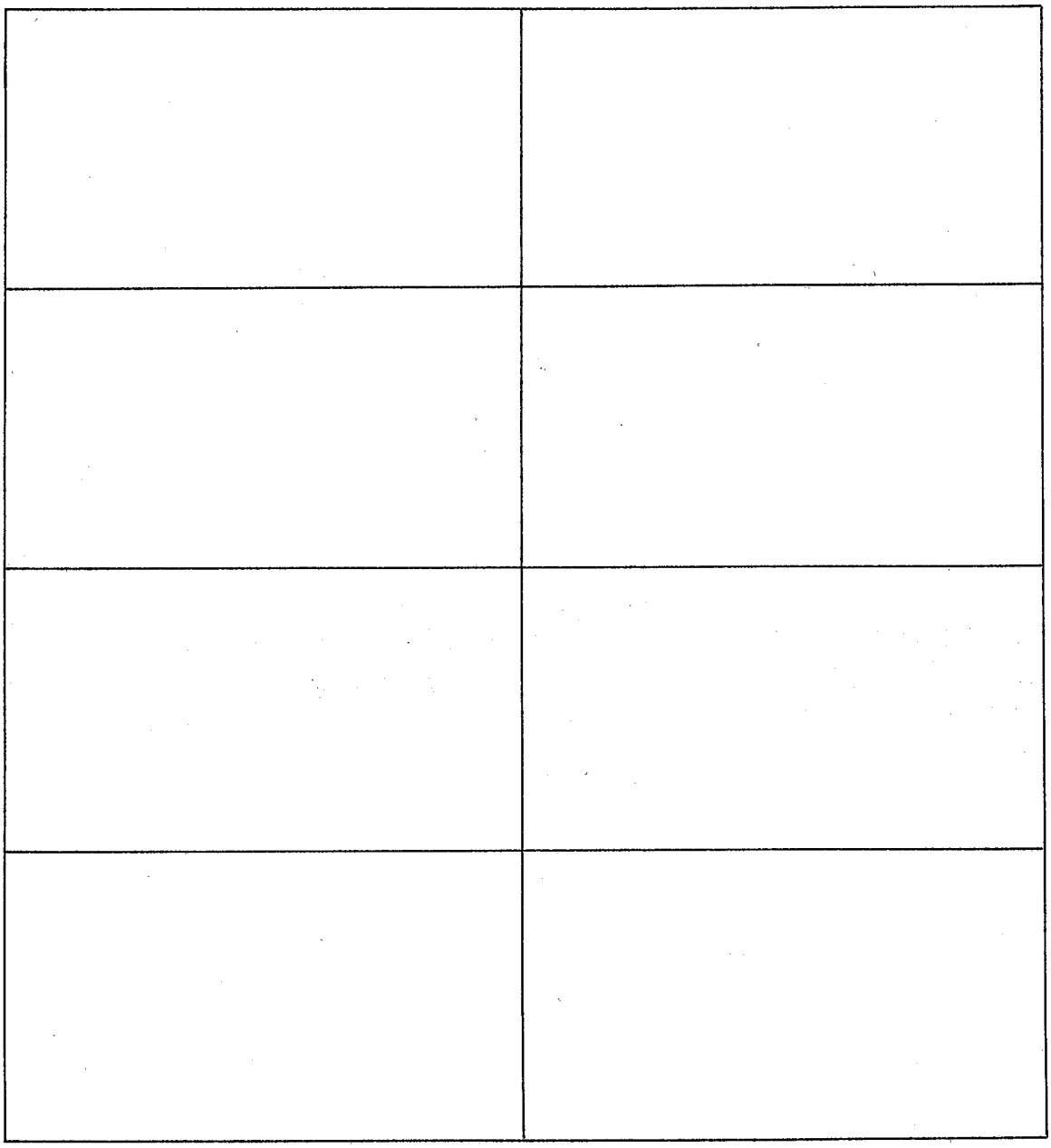
Rating Scale	Total Score	
	86 - 76	High
	75 - 64	High Average
	63 - 56	Low Average
	55 - 0	Low

Pupil's Rating_____

SECTION ONE

VISUAL-MOTOR COORDINATION

Score_____ Number of correctly reproduced designs



SECTION TWO

VOCABULARY

Score_____ Check each correct response.

Picture One

_____ 1. carrots
_____ 2. bananas
_____ 3. grapes
_____ 4. corn
_____ 5. orange
_____ 6. cherries
_____ 7. beans
_____ 8. roots

Picture Two

_____ 9. flower
_____ 10. something steaming
_____ 11. lamp
_____ 12. dust pan
_____ 13. propellers
_____ 14. vase

Picture Three

_____ 15. something burning
_____ 16. feather
_____ 17. faucet
_____ 18. match
_____ 19. light
_____ 20. something to eat
_____ 21. something sharp

Picture Four

_____ 22. shovel
_____ 23. sprinkler
_____ 24. nail
_____ 25. rake
_____ 26. saw
_____ 27. pliers
_____ 28. screwdriver
_____ 29. ruler

Picture Five

_____ 30. necklace

_____ 31. hair

_____ 32. feet

_____ 33. wire

_____ 34. toy

_____ 35. ribbon

_____ 36. handle

_____ 37. dial

_____ 38. cap

Picture Six

_____ 39. book

_____ 40. umbrella

_____ 41. wheel

_____ 42. someone sleeping

_____ 43. hitting

_____ 44. hammock

Picture Seven

_____ 45. candle

_____ 46. something to wear

_____ 47. pocket

_____ 48. picture

_____ 49. something to drink

_____ 50. globe

Picture Eight

_____ 51. fence

_____ 52. spokes

_____ 53. someone leaning

_____ 54. motor

SECTION THREE

CONCEPTS

Score_____ Check each correct response

Color	Size	Likenesses
_____ red	_____ largest	(sample) triangle
_____ blue	_____ smallest	_____ moon
_____ orange	_____ tallest	_____ flower
_____ green	_____ shortest	_____ cube
_____ yellow		

Number	Differences	Position
_____ two	(sample) pail	_____ in front of
_____ four	_____ wagon	_____ in back of
_____ six	_____ boy	_____ on top of
_____ three	_____ sailboat	_____ inside
_____ eight		