6-1959

A Developmental Study of Performance on the Hunter-Pascal Concept Formation Test

Ernest R. Larsen
University of Tennessee - Knoxville

To the Graduate Council:

I am submitting herewith a dissertation written by Ernest R. Larsen entitled "A Developmental Study of Performance on the Hunter-Pascal Concept Formation Test." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Psychology.

Gerald R. Pascal, Major Professor

We have read this dissertation and recommend its acceptance:

William O. Jenkins, Raymond R. Shrader, Harold Holloway, Virgil E. Long

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

July 10, 1958

To the Graduate Council:

I am submitting herewith a thesis written by Ernest R. Larsen entitled "A Developmental Study of Performance on the Hunter-Pascal Concept Formation Test." I recommend that it be accepted in partial fulfilment of the requirements for the degree of Doctor of Philosophy, with a major in Psychology.

Gerald R. Pascal
Major Professor

We have read this thesis and recommend its acceptance:

W. O. Jenkins

R. R. Shrader

V. E. Long

Harold D. Holloway

Accepted for the Council:

Dale Wantling
Dean of the Graduate School

A DEVELOPMENTAL STUDY OF PERFORMANCE ON THE HUNTER-PASCAL CONCEPT FORMATION TEST

A THESIS

Submitted to
The Graduate Council
of
The University of Tennessee
in
Partial Fulfillment of the Requirements
for the degree of
Doctor of Philosophy

---

by

Ernest R. Larsen

June 1959

ACKNOWLEDGEMENT

I would like to extend my regards and appreciation to the members of my committee, Doctors Gerald R. Pascal, William O. Jenkins, Raymond R. Shrader, Harold D. Holloway, and Virgil E. Long for their aid and guidance throughout the course of this study.

A special debt of gratitude is owed Mr. Curtis G. Gentry, Director of Instruction, Mr. Thomas B. Smith, Supervisor of Elementary Education, and Mr. Roy R. Anderson, Supervisor of Secondary Education for their counsel and willingness to supply the experimental subjects. And, last but most certainly not least of all, to those Principals, Mr. Cecil Kelly, Miss Helon Brixie, Mr. William Davis, to Assistant Principal Worrell Bosse, and to the many teachers, who by virtue of the length of the list must go unnamed, who displayed such keen interest and a ready willingness to cooperate and interrupt their very busy schedules, I express my most sincere thanks.

Ernest R. Larsen

TABLE OF CONTENTS

CHAPTER PAGE

APPENDICES PAGE

LIST OF TABLES

TABLE PAGE

CHAPTER I

INTRODUCTION

When psychology was closely allied to philosophy, it was defined as the study of "consciousness" or "mind." Its' methodology was investigating introspection, and by necessity, this limited any experimentation to the human species. In Europe, the initial research emphasis was on sensation and perception.

Although this philosophical orientation was questioned by a few investigators, this prevailing theoretical climate was irreversibly changed in America with the arrival and wide acceptance of Watson and his Behaviorism (37). Behavior then became the proper object of interest, and since all organisms behave, the object of study of psychology increased a thousand-fold.

In America, psychology's interest turned to the field of learning. Prior to Watson's main publication in 1913, some experimenters were engaged in animal research. In the 1890's Thorndike had used cats in his puzzle-boxes. In 1901, W. S. Small started a lasting friendship when he introduced the laboratory white rat to a small scale reproduction of the Hampton Court Maze (2). This was the era of the creation of new techniques and apparatus for studying behavior.

Learning and The Phylogenetic Scale

Soon experimenters began investigating how far down the phylogenetic scale an organism could be placed and still be capable of learning. The fact that mammals are capable of some type of learning is indisputable. Munn (29) concludes that in simple discrimination and maze problems, lower vertebrates do not demonstrate significantly greater ability than some higher invertebrates.

Within the invertebrate classification arthropods (crabs, bees, ants, and cockroaches), octupi, snails, worms, and starfish can learn; and there is some questionable evidence that microorganisms like paramecium and stentor show those changes of behavior we term "learning" (29).

Thinking

As time passed, some investigators became interested in a still more complex aspect of the living organism, "thinking." The investigation of this problem within the framework of early American Behaviorism and animal subjects posed some challenging questions of theory and methodology.

One of the earliest attempts at incorporating "thinking" into a theoretically sound and experimentally measureable procedure was proposed by Hunter. He used the two techniques of delayed reaction and double alternation to study "symbolic" behavior. The definition of these two problems and their experimental history will be pursued in the next chapter.

These two problems became a standard technique of the comparative psychologist and they were used in many phylogenetic studies. As the missing details of the phylogenetic variable were filled in, these two problems were used ontogenetically with human subjects. It became apparent that the full possibilities of these two problems had not been utilized. To this end a test was created by Pascal and Jenkins (30) to more fully utilize the proven merits of these problems.

Statement of Problem

The present experiment is a developmental study of the ability of human organisms to solve the delayed reaction, double alternation, and serial reaction problems which constitute the Hunter-Pascal Concept Formation Test. It extends the phylogenetic study of the ability of animals to form concepts to the human organism, using essentially the same techniques and problems.

CHAPTER II

SURVEY OF LITERATURE

As maze studies became increasingly popular, the question arose as to what controlled or mediated the rat's ability to learn and follow the correct path. Watson conceptualized the process as a chaining of stimulus-response relationships; Hunter was to question this formulation through a protracted series of papers.

The Delayed Reaction Problem

The delayed reaction problem was conceived by Harvey Carr and first used by Hunter under Carr's direction (28). It consisted of presenting a discriminating stimulus to a subject (*S*), preventing the *S* from responding immediately, withdrawing the stimulus, and then allowing the *S* to respond in the physical absence of the stimulus. Since all behavior must result from previous stimulation, and since the original discriminating stimulus is no longer present, some substitute stimulus must be involved.

An example from the earliest work of Hunter (11) may clarify this procedure. Hunter conditioned rats to run to a light above a door opening in a three choice maze while the light remained on. After this association was well-established, the animal was allowed to see the light go on but was

detained in its starting-box until the light was turned off. The animal was then released after varying periods of time.

Lower Vertebrates

The initial phase of the delayed reaction research investigated the lower vertebrates with the white rat being most often the experimental S. Although attempts were made to establish absolute time limits of delay for the rat, variations in experimental procedure made the various results of different experimenters (Es) non-comparable. The shortest maximal delay for rats of ten seconds was found by Hunter (11); McCord (25), using a different method, extended the delay to four minutes with the rats reaching a high level of accuracy. McCorquodale (26) replicated McCord's experiment and obtained similar results.

Hunter (11) established twenty-five second maximal delays with raccoons and five minute maximal delays with dogs. Using a modified apparatus, Adams (1) found cats able to delay from three to seventeen hours.

Primates

The experimental results with primates generally show longer periods of successful delay, but they also show high variability. Yerkes (39) found a brief maximal delay of only two minutes for some gorillas and monkeys, although other gorillas could delay for forty-eight hours. Tinklepaugh (36) found his monkeys making a high percentage of correct responses

after delays of from fifteen to twenty hours.

Humans

Hunter (11) (12) and Skalet (34) studied normal children in the age ranges of thirteen months to eight years. The maximal delays ranged from a scant eight to twelve seconds (one of Hunter's (12) 13 month old Ss) to one to thirty-four day delays (Skalet's (34) Ss). Skalet found a correlation of +.478 between maximal delay and chronological age. Using a group of male mental defectives, Pascal et al (32) found a correlation of +.61 between ability to delay and Mental Age. Collins (3) and Tapp (35) used a series of problems involving delayed reaction and single alternation with children ranging in age from five to sixty months. They obtained a smooth, negatively accelerated function when problem performance was related to chronological age. Their function appeared to rise rapidly until age two and one-half where some slowing down appeared. Using a similar age range and a more extended series of problems, Kaiman (21) discovered a high positive correlation between problem solution and chronological age.

Although there have been different maximal delay times reported on the delayed reaction test with different Ss and different apparatus and techniques, Heron (9) concluded that ability to delay corresponds with placement on the phylogenetic scale.

The Double Alternation

Feeling unconvinced that the work with the delayed reaction had established his ideas regarding the "symbolic process," Hunter (13) began working with the double alternation problem in 1918, launching a research program which has continued aperiodically for thirty years. He conceived of the double alternation temporal maze as a better instrument to measure symbolic ability.

Lower Vertebrates

It was a well-demonstrated fact that the white rat could master a single alternation pattern in a spatial maze. However, the pattern of right, left, right, left (RLRL) could be mediated through kinesthetic cues. The stimulus for the first right turn could be the consequences of being placed in the maze or starting box. The stimulus for the next left turn could be the kinesthetic and proprioceptive stimuli as a consequence of having made a right turn. In this fashion, transversing each portion of the maze could set-up a differential stimulus pattern for making the next turn.

In the double alternation temporal maze the rat transverses a figure-eight shaped path in a T maze whose arms lead back to the starting position. To be successful on this maze requires the _S_ to turn right twice at the choice point followed by two left turns in a RRLL pattern. In this series the sensory consequences of turning right would

lead to a following right turn half the time and a left turn the other. Hunter (14) states,

> It is impossible for one and the same stimulus to cause first one response and then another unless it is supplemented by some other factor either inside or outside the organism. . . the experimental situation rules out the possibility of a supplement from outside the subject's body. . . .

Hunter (14) theorized that the ability to perform double alternation could be mediated by cumulative neural effects or a symbolic process. Although the evidence was never very clear-cut, he preferred the latter explanation (17).

The data regarding the rat's ability to learn the double alternation response is complex (13) (15) (17) (18). In Hunter's original temporal apparatus only a portion of his rats could learn the pattern; this required a long series of training trials of a very special nature. Furthermore, the degree of mastery was never complete, and the rat could not extend the series beyond that upon which he was trained (RRLL).

More recently, Schlosberg and Katz (33) modified the locomoting response into a lever-pressing habit. They found that the rat could learn the RRLL pattern in fewer trials and with a higher degree of accuracy. However, since their apparatus was quite different from Hunter's, it is not directly comparable to the other studies in the series. Schlosberg and Katz's results were in direct opposition to

an earlier experiment of Jenkins (19); he also used a lever pressing response in a Skinner box, and his _S_s were not able to learn the double alternation pattern.

In 1928 Hunter (14) used raccoons as _S_s. Although learning the series proved to be a very difficult undertaking which pushed the _S_s' capacity to the limit, some _S_s could learn, and Hunter judged raccoons' ability to be slightly superior to rats. Here again, this acquisition required an extremely long training period.

Other investigators followed-up this problem using other types of _S_s. Karn (22) found cats able to master the maze with a high degree of mastery, but his _S_s could not extend the series beyond the length upon which they were trained. Karn and Malamud (23) also found dogs able to learn quicker than cats, but the dogs also were unable to extend the series.

Primates

Gellerman (5) used rhesus monkeys and found them superior to rats, cats and dogs. However, he was unable to attempt a test of their ability to extend the series and as a result concluded that the maze was unsuitable for use with monkeys.

As a result of his difficulties, Gellerman (7) built an apparatus more suitable for the monkeys since it was limited to forearm manipulation; and, with this change,

he clearly established that the monkeys could extend the series beyond that of their training.

Human Subjects

In 1931 Gellerman (6) completed the lower phylogenetic phase and made the logical extension of the problem to human subjects. He used children, from age three to thirteen, and college sophomores in a walk-through temporal maze. The adults learned significantly faster than the children, and the children's performance surpassed that of the other vertebrates. In this study some other information was available regarding sophomores: their grades in psychology and scores on The Thurstone Psychological Examination for College Freshmen. Gellerman obtained a correlation of +.58 between quickness to learn and intelligence, and a coefficient of +.21 between successful maze performance and course grades.

Hunter and Bartlett (16) attempted a follow-up of Gellerman to determine more precisely at what age the double alternation problem can be solved by children. They broke with the temporal maze and modified Gellerman's apparatus for monkeys (7). They constructed two boxes with lids which were attached to a front wall and base. Sitting behind the wall the E had access to the inside of the boxes by means of holes cut through the front wall. Their youngest S to solve double alternation to a criterion of three errorless

trials (of eight responses each) was 3 years and 7 months old. This child was also able to extend the series beyond his training length. As a result of the change in apparatus, Hunter's Ss learned the pattern more quickly than Gellerman's children using the walk-thru maze.

In 1955, Pascal and Zax (31) questioned the use of standard intelligence tests for a Cerebral Palsy population. They stated, "The need for a standardized test of intellectual capacity relatively free from dependence upon verbal and motor skills is clear." They speculated regarding the ability of the double alternation test to predict the ability of Cerebral Palsy children to profit from special training. To this end they constructed two rating scales of educability and cooperativeness. Correlating double alternation performance against a combined score on these two scales they obtained a correlation of +.64. A bi-serial r for Pass-Fail versus behavioral rating yielded a value of +.73. Two judges, working independently, used existing Stanford-Binet IQ scores plus clinical judgment of the test scatter to rank order these twenty-four Ss. When these composite rankings were correlated against the behavioral ratings, the correlation was +.57. Pascal and Zax concluded that the double alternation performance correlated higher with the behavioral ratings than did the IQ tests. They also stated,

> The results of this study suggest many possibilities for the development of an entire series of problems utilizing the apparatus used here which can be

> substituted for the standard intelligence tests in working with children who are impaired in verbal or motor areas or even those who are deaf and possibly blind. The apparent need is, of course, for a series of graded problems among which double alternation would be a more complex one. The simplest problems are those utilizing delayed reaction. . . Next are the problems where reward is always found behind the same door, followed by the single alternation problem and then the double alternation problem. Finally extremely complex serial learning problems may be used to measure in the highest ranges of intelligence.

A special apparatus was used in this study which extended the two box apparatus of Gellerman into a flat wall with five doors. This apparatus was created for the earlier Pascal et al (32) study. Except for minor differences in dimensions and construction details, the apparatus used in these two studies is that used for the Hunter-Pascal Concept Formation Test (30).

Pascal and Jenkins began work on this series of problems in 1956. A preliminary form of the test was used by Leslie (see below), and a further modification was done in 1957 (30). More information on this test will be given in the next chapter.

Hodges (10) used a large sample of 240 male elementary school children ranging in age from six to twelve years. He varied the procedure by using five playing cards rather than the five door apparatus of Pascal et al (32). Hodges found significant relationships between the number of trials to the criterion and number of errors versus chronological

age and Mental Age.

In his discussion, Hodges noted a tendency for Ss who seemed emotionally disturbed to do poorly on the double alternation test. This lead was followed-up by three investigators.

Psychological Deficit and Alternation Ability

Worley (38) used Hodges' card procedure in the attempt to discriminate between varying degrees of severity of mental illness. When double alternation performance was compared to judges' ratings of severity of illness, only a moderate correlation was obtained, but when it was compared to the psychotic vs. non-psychotic dichotomy, a high positive relationship was found. This points out that those factors which accompany mental illness are associated with a loss in the ability to conceptualize.

When a group of alcoholics were compared with controls matched for intelligence and education, Davis (4) discovered the controls surpassed the alcoholics.

Leslie (24) modified Worley's study by returning to the five door apparatus and extending the complexity of the alternation test. He used the Pascal-Jenkins series of four problems: Single Alternation, Double Alternation, and a six and eight response pattern of Serial Alternation composed of single and double alternation. A group of hospitalized neurotics and psychotics were matched for intelligence and

education. Leslie found a significant difference between the neurotics and psychotics performance when compared to both the judges' quantitative rating of severity of illness and the psychotic versus non-psychotic dichotomy; the latter comparison resulted in a more significant discrimination than the former.

Summary

The H-P Test is a logical extension of a procedure which has traditionally been used in the psychological laboratory in the study of symbolic behavior. Various vertebrates have been studied and their ability to perform the test corresponds perfectly with their placement on the phylogenetic scale. The following organisms have been systematically studied and their ability from little to much is ranked as follows: rats, raccoons, cats, dogs, monkeys, children, and adults. Some ontogenetic research has been carried out with human Ss showing an increase in capacity to solve the problems with increasing maturation. A graded series of these alternation problems also shows promise as a potential measure of psychological impairment.

In the beginning a temporal maze was the favored apparatus. This has been modified into a two box problem, a five card test and finally into the five door apparatus which is used in the present study.

CHAPTER III

METHODS AND PROCEDURES

Experimental Design

Ten students, five male and five female, were selected from each of the even numbered grades, second through twelfth, resulting in a total of sixty Ss. Each S was administered the H-P Test on an individual basis. The relationship between performance on the Hunter-Pascal Concept Formation Test (H-P Test) and chronological and mental age was studied. The relationship between ability to perform the test and academic achievement was a secondary problem.

Selection of the Subjects

The experimental Ss were students from an elementary school, junior high school, and high school in the Knoxville city school system. Ten students from each of the even numbered grades, second through twelfth, were chosen. The selection of grades from particular schools and children within particular grades required some consideration of sampling problems.

On the basis of some limited experimental work (6) (10) (16) (31) some of the problems in the H-P Test are known to be correlated with intelligence. The substantial

relationship of occupation to intelligence has been well established. Therefore, since occupation, and more directly income, is a major aspect of socio-economic status, we could predict that, to some extent, success on the H-P Test would be correlated with socio-economic status.

Certain Knoxville schools were known to draw their students from very homogenous socio-economic groups. Other schools, because of their geographic school district, tapped a heterogeneous sample. It appeared quite conceivable that sampling variations between schools could be so extreme as to greatly override the maturational variable. For example, a sample of school A's eighth grade students of high socio-economic status could have been compared with a sample of school B's tenth grade students of low socio-economic background. It was conceivable that considerable overlap of these two distributions could have resulted because of sampling variation.

The purpose of a developmental study is to investigate the relationship between the ability to perform a task and chronological age, that is to duplicate the effect of a repeated re-testing of the same group of children over an extended period of time. Therefore it appeared best not to confound the problem by introducing uncontrolled variations of socio-economic status between different school grades.

In addition, such sampling variations could easily have modified the resulting graphic function. This could

be especially true within the older age ranges where some leveling-off of the function might be expected. By an uncontrolled variation of socio-economic factors, that is using different types of school districts, any empirical evidence of a leveling-off of the function could be the result of sampling variations.

Therefore, a feeder-system relationship was required of the three schools which were selected. They should all draw their students from the same geographic location, funneling them upward; all the graduates from the grammar school should go to the same junior high school, and so forth. These qualifications were explained to the school board and the three schools were chosen upon the basis of their knowledge and experience. The school district which was selected had a population which was heterogeneous, but it tended to have disproportionately less members of the lower socio-economic class and was therefore a slightly higher than average sample of the city's population taken as a whole.

The selection criteria for picking a sample of ten experimental Ss from a class population of twenty to forty students was the next issue. The primary focus was on chronological age, and grade units were chosen as a practical method of dividing the Ss into experimental sub-groups. The three schools all had more than one class of each grade, but there was no known selection factors involved in the assignment

of students to particular teachers, or particular home rooms.

After a class was selected on a chance basis, it was divided into two rosters, boys and girls. Each sex was ranked according to their chronological age. Limits of plus and minus twelve months from the median age were used to define the greatest degree of acceptable variability. Any children above or below these cut-off points were considered to be grossly deviant with regard to age; they were also generally deviant with regard to their school achievement. This exclusion seemed necessary because the study was concerned with "normal" children; almost all of the Ss eliminated were of excess age.

The median age and the limits of plus and minus twelve months from the median age furnished three reference points for each sex. Those Ss whose chronological ages were the closest to these reference points were selected, and this resulted in placing three Ss in the sub-group. The last two reference points were the median age plus and minus six months, and those Ss closest to these age values were picked to complete the experimental sub-group of five boys and five girls, ten Ss in all. This procedure attempted to spread the ages of each sex through a twenty-four month range in each grade group.

This tentative list was then shown to the Ss' teacher, and she was asked if any Ss were deviant with regard to their classroom behavior. This procedure was possible at the

grammar and junior high school levels, for at these levels the teachers had considerable contact with each student. At the high school level the criterion for exclusion was changed to any deviant behavior which would bring a student into disciplinary contact with the principal. No students were excluded from the sample for either of these two reasons. The E had the benefit of the school medical records to eliminate any children with impaired and uncorrected losses in hearing or vision.

In grades second through eighth, seven changes were required from the original list. The substitutions were made for the following reasons: four Ss were absent on the date of testing, one S left school, one S had previously failed a grade, and one S was chronically ill. This was a total S change of 18 per cent. At the high school level the E was only allowed to test the Ss during their study hall period. Therefore, a S had to have a study period on a day of the week that the E was free for testing. As a result of this requirement, considerable substitution was necessary. However, the substitutes were usually quite close in age and no systematic bias was apparent in Ss having study halls on particular days of the week. The general characteristics of the experimental population are given in Table I.

TABLE I

SUMMARY TABLE FOR CHARACTERISTICS
OF THE EXPERIMENTAL POPULATION
(N=60)

	School Grades					
	2nd	4th	6th	8th	10th	12th
Mean Chronological Age	7-7	9-5	11-9	14-0	16-0	17-9
Mean Mental Age	7-5	10-2	12-5	14-3	14-5	15-10
N	10	10	10	10	10	10

The Hunter-Pascal Test

The Hunter-Pascal Test (H-P Test) was developed by Pascal and Jenkins (30). At present the test is available only for research purposes, and a preliminary manual has been written which may be obtained from the authors. Much of the following material can be found in greater detail in this manual.

Apparatus

The apparatus for the Hunter-Pascal Test is a modification of a device originally used by Pascal et al (32) in a delayed reaction study. As the S faced the apparatus he saw a white wall twelve inches high and eighteen inches wide. This flat white surface was broken by five door openings, two and one-forth by three and one-forth inches, equidistant from each other; the openings were closed from behind by five black doors. The doors were hinged at the top and opened upward and inward with a light touch from the S's fingers. The sides of the front wall extended away from the S and back toward the E thirteen inches. Thus these three surfaces functioned to conceal the movements of the E's hands and arms as he placed the reward behind a predetermined door. To further prevent the Ss from locating the reward through extraneous factors, shields were placed between the doors so that it was not possible for a S to open a door and see the area behind an adjacent door. See Appendix C for further

details concerning the construction and dimensions of the apparatus.

The Reward Placement Technique

The E placed several thicknesses of cloth beneath the apparatus to dampen any auditory cues arising from the placement of the candy. The younger children, grades second through sixth, could not see the movement of the E's arms and shoulders because of their shortness of trunk (while seated) in relationship to the height of the apparatus. With taller Ss a particular system was used to prevent their utilizing visual cues. The pile of candy was centrally located behind the front wall. The E took a piece with either the right or left hand -- whichever was appropriate for the particular presentation -- and pushed both hands forward into the space behind the middle door. Then at the same time the left hand went into door two, the right hand entered door four. This was followed by a similar simultaneous entrance into door one and door five with the left and right hand respectively. The candy could be released behind any of the doors. With a minimum of practice this procedure developed into a quick and rhythmic process. The E did not follow the movement of his hands with his eyes or check visually to see if the candy was correctly placed. The procedure was followed of maintaining contact with the S's eyes throughout the placement. Since the S was supposed to

attempt to grasp the plan lying behind the placements, any attempts on the part of the E to work too rapidly would minimize the time for rehearsal. Therefore a moderate rate, which could be easily handled by the E, was used.

Reward

The reward consisted of a confection called "corn candy." It is a semi-hard orange, yellow and white candy shaped like a kernel of corn enlarged in size by a factor of two or three times. The Ss were informed they could keep all the candy they discovered, but they were allowed to eat only the first piece and encouraged to save the remainder to eat after the testing was completed. This procedure avoided any satiation effect.

The Problems

The two problems of single and double alternation used by Hunter with his animal Ss did not present a sufficiently broad range of difficulty for all ages; single alternation was too difficult for young children and double alternation was too easy for the older Ss. Therefore, in order to extend the age range and the intrinsic discriminatory ability of the alternation test, Pascal and Jenkins (30) placed additional problems above and below those of Hunter. This resulted in a series of eight problems, ranging from very easy to very difficult.

The initial task, Problem I, was one of Immediate

Reaction. Standing in front of the apparatus, the E, making certain he had caught the S's attention, placed the reward behind a door as the S watched. The S was then allowed to immediately search for the candy.

Problem II, Delayed Reaction 1 (Short), is similar to the first except that it interjects a short delay between watching the reward being placed and the S being allowed to search. Immediately after the reward placement, the S is turned (or with the older Ss asked to turn) around in a complete circle: the S was then allowed to search.

In Problem III, Delayed Reaction 2 (Long), the S was again allowed to watch the placement of the reward. Then he was led away from the apparatus and presented with a distracting stimulus for a sufficient period of time so that he could be returned to the front of the apparatus exactly sixty seconds after the placement of the reward. The distracting stimulus was looking at books suitable to the child's grade level.

Problem IV was Single Alternation. For this problem the E retired behind the box so that he could place the reward behind the doors without the S's awareness. The reward was placed behind the doors according to a 1 5, 1 5 pattern.

Double Alternation, Problem V, consisted of a 1 1 5 5, 1 1 5 5 sequence.

The sixth problem, Serial Reaction 1, was inserted as a transition problem to bridge the gap between Double Alternation and the more difficult Serial Alternations. It consisted of a 1 1 1 5 5 placement.

Problem VII, Serial Reaction 2, was a combination of double and single alternation. It's pattern was 1 5 1 5 5 5 .

And, the most difficult problem, Serial Reaction 3, utilized an eight unit combination of single and double alternation, 5 1 5 5 1 5 1 1.

Scoring

The H-P Test was composed of eight problems of increasing levels of difficulty. All the problems consisted of ten trials. A trial consisted of the entire series or pattern being given to a S one time. A trial is further subdivided into presentations. A presentation consists of one discovery of the reward. While there are always ten trials per problem, the number of presentations per trial varies with the complexity of the problem. On Problems I, II, and III, Immediate Reaction, Delayed Reaction (Short) and Delayed Reaction (Long), each trial consists of only one presentation. On Problems IV through VIII the number of presentations per trial is two, four, five, six and eight respectively.

Solution of a problem was defined as selecting the correct door throughout two errorless trials; these two

errorless trials constituted the criterion of learning. The various scores consisted of the sum of all units up to but not including those units in the two criterion trials.

One specific deviation was made from the foregoing. If a subject showed, by an errorless performance on trial ten, he was beginning to learn the pattern, he was given one additional trial. If he succeeded, he was credited with passing the problem and his score consisted of the number of presentations, trials, and errors up to but not including the criterion trials. If he made a mistake on trial eleven, his score was taken from the first ten trials and the eleventh was disregarded. This was designed to prevent a S's score from being penalized for a last-minute near solution.

Each S had ten trials in which to solve a given problem, and if he had not reached a solution at this point, he was considered to have failed the problem and testing was discontinued. Termination after only one problem failed, arose from the developmental work done by the authors, for they (30) established the fact that the problems are arranged in order of increasing difficulty.

The Presentation Score was the sum of all the presentations across problems. If a S did not attempt some problems because of a previous failure, he was credited with the maximal failing Presentation Score for each of the problems not attempted.

Dimensions of H-P Test Performance

Several possible aspects of the H-P Test can be quantified as dependent variables. In the section below the possibilities of these measures will be discussed.

Number of problems passed

The number of problems a S can solve is most obviously a key measurement. However, this variable lacks the power of discrimination of other closely related measures. An extreme example would be a S who solved a particular number of problems with a minimal number of presentations in contrast to another S who solved the same number of problems with the maximal allowable number of presentations. Both Ss would be credited with the same number of problems passed: their performance would be quite different.

Number of trials to the criterion

The number of trials is partially derived from the pass-fail variable. Each problem has a maximum of eleven trials; a solution can result in a minimal number of zero trials and a maximal number of nine trials. This measure adds more discrimination than pass-fail. However, the Presentation Score (a description of which follows) adds even more discrimimation, and since the trials score and Presentation Score are very highly correlated, the number of trials score adds little information to the Presentation Score.

Presentation Score

This score is the number of placements in each series (a trial) multiplied by the number of trials necessary to reach the criterion. From Problem IV, Single Alternation, through Problem VIII, Serial Reaction 3, the number of maximal presentations increases from twenty to eighty. It is this property which results in a theoretical range of scores from zero to two-hundred and eighty; therefore this characteristic "weighting" yields the greatest discrimination of all the score variables.

Error Score

The number of mistakes in performing a task is also a classical variable of psychology. Here again, the structure of the H-P Test prevents utilizing the error score in a straight-forward fashion. The number of errors a given S makes is contingent in part upon his number of opportunities to err. Since all Ss stop with the first problem failed, there will be wide variations in the number of problems attempted and the total number of opportunities for mistakes. An example can best illustrate the enormous range of possible differences.

Let us disregard for the purpose of example Problems I through III. Generally the greatest number of errors made on a single presentation is four. If a S solved Problem IV and failed Problem V he would experience 4 x 18 (maximal

number of passing presentations) plus 4 x 40 = 232 chances to err. If another S solved all problems except the last one, he would have a chance at making 4 x 18 plus 4 x 36 plus 4 x 45 plus 4 x 54 plus 4 x 72 = 900 errors. This factor has such a large effect upon the scores that the correlation between Presentation Score and error score is negative and moderately high ($r = -.48$). In other words, those Ss' who have the best (lowest) Presentation Scores tend to make more errors because they have more opportunities to err, since all Ss stop after the first problem failed.

Therefore, for errors to be a meaningful measure, each S's error score has to be corrected in some fashion with regard to his opportunities for making mistakes. It appeared reasonable to follow the Presentation Score procedure to obtain a corrected error score. Therefore a maximal error score was calculated for each problem. The total corrected error score was composed of the sum of the errors the S actually attained on the problems he attempted, plus a theoretical maximal error score of four errors per presentation on the problems he did not attempt. The value of four errors per presentation was selected because only in extremely rare cases did a S actually make more than four errors per presentation. This appeared to be a meaningful variable. However, when this procedure was followed there was an almost perfect correlation between Presentation Score and corrected error score.

A second attempt to utilize the error data was made. Rather than assuming maximal error scores for the problems not attempted, the Presentation Score for problems actually attempted by each S was divided into his actual total error score. This Error Index produced only a moderately high correlation with Presentation Score ($r = +.61$), and therefore, it appeared to supply additional information.

Test time

The time required to perform a task has long been a traditional measurement of psychology. The H-P Test is a broad range instrument designed to be used from age six months through old age. It is divided into a series of problems of eight levels of increasing difficulty. As a result the more difficult problems are definitely not appropriate for the Ss at the younger age levels. A standardized point of termination ends the testing when the S is unable to solve a problem within the limited number of ten trials; this prevents the E from expending unnecessary time and also avoids introducing an inevitable course of frustration and/or boredom to the S.

Therefore, even within a particular grade level, different Ss will work varying numbers of problems. Since the problems have different numbers of presentations, from ten to eighty, they are not directly comparable with regard to time.

In addition, a noticeable reduction in total time could result as the E attains more proficiency and dexterity with practice; this could result in a substantial degree of systematic variation.

Similarly, small daily variations in the speed of presentation by the E could introduce, through the accumulative effect of some two-hundred and eighty presentations, a considerable element of chance variability of time.

Therefore, neither time per problem nor total test time appear to be a useful variable, but they are recorded and included in the analysis.

H-P Test Performance and School Achievement

The primary question raised by the present study was the relationship between H-P Test performance and both chronological and mental age; but since other information about the Ss was also available, it too was analyzed.

Certainly one of the most meaningful aspects of the Ss'behavior in the school situation is their academic achievement. Grades assigned by teachers are definitely an imperfect measure of achievement, being subjective, unreliable, and multi-dimensional: well standardized achievement tests are more desireable. Unfortunately, achievement tests scores were only available for one-third of the sample, and the various tests were not comparable to allow for a direct

comparison across grades. However, the relationship between H-P Test performance and achievement for the fourth and sixth grades on the various tests was investigated.

By contrast, all Ss in the sample received grades, and these alphabetical symbols represented -- despite their shortcomings -- the only estimate of achievement common to the entire sample.

From an *a priori* viewpoint, two methods of arranging the grades seemed appropriate. It seemed reasonable to hypothesize that those numerous attributes and circumstances which determine a given S's grade in a given course vary through time. If the H-P Test is a measure sensitive to changes in contemporary efficiency, then the proper test of relationship would be between the H-P Test and the S's contemporary grades.

On the other hand, to the extent that the H-P Test is assessing a more stable variable, it would be expected to correlate more highly with the Ss' past grade point average. This measure would also have the advantage of being more reliable because it would average out random errors of teachers' gradings and encompass a greater sample of each S's behavior. However, the influence of constant errors on these grades was an important yet unknown factor.

Both of these relationships were investigated, but, unfortunately, the contemporary grades were not available for the tenth and twelfth grades. The past grade point average

was composed of from two to six series of final grades, depending on the extensiveness of the school's records and the number of years each S had been in school. Both series of grades were independent since the contemporary grades were not averaged into the past grades.

The distributions for both contemporary and cumulative grade point average are included in Appendix D.

Intelligence Test Scores

The scores on intelligence tests were an important aspect of this study. They yielded the Mental Ages which were related to the various H-P Test measures, and the IQ scores were also used in the correlations with grades.

The use of intelligence tests in the schools from which the sample was drawn was rather extensive. At the grade school level the children were initially tested with the Otis Self-Administering Intelligence Scale as early as the second grade and were routinely retested two years later with the same test. As the children became older they were tested with Form C of the Otis.

Some Ss had been tested only once, while other Ss had as many as three testings. The variability of some of the test re-test data was rather high. Its' most extreme case involved a double testing of the fourth grade when the mean score dropped fifteen points for an experimental

sub-group of ten Ss. This change was a significant difference (Wilcoxon Matched-Pair Rank T test (20), P between the .05 and .02 level). From a more general point of view, the six grades had a total of twelve testings. The distributions were all thrown in together and treated as independent groups. The Paitnik F_{range} analysis of variance (20) yielded a value of 3.7, P greater than .05.

Since there was some discrepancy between an individual S's score on two or more testings, the question arose as to which testing should be considered most representative. It was decided that whenever more than one IQ score was available for a given S that the mean score would be used. This would give a general policy for treating all cases and would also be more reliable since it was composed of a larger sample of each S's behavior.

The IQ distributions which were finally used are shown in Appendix F. The mean IQs for each grade are quite similar (F_{range} 2.5, degrees of freedom 48, P greater than .05), and as a whole, the distribution is skewed in favor of above average intelligence.

In calculating the relationship between the H-P Test and intelligence it was more appropriate to utilize the concept of Mental Age rather than an index of relative brightness (IQ). While an extremely bright second grader (IQ = 135) would stand out in his second grade classroom, it would seem unreasonable to expect him to do better on the

H-P Test than a sixteen year old high school student who would have only average intelligence, because the latter would have a Mental Age of sixteen years as compared to the second grader's Mental Age of ten years. Therefore, it was necessary to convert the IQ's into Mental Ages. The IQ table of the Stanford-Binet was entered and knowing CA and IQ, the appropriate Mental Age was determined.

Statistical Analysis

Two types of statistical techniques were used in the study: analysis of variance and correlations. The Kruskal-Wallis χ^2_H non-parametric analysis of variance was appropriate, because when ten subjects are used in each sub-group, it is not possible to determine whether or not parametric assumptions are fulfilled. For the same reason, the non-parametric Spearman rank correlation was the coefficient of choice.

CHAPTER IV

RESULTS

Number of Problems Passed

Table II presents the number of Ss passing specific problems on the H-P Test when the Ss are grouped by school grades. Casual inspection of the data shows that Problems I, II, and III did not discriminate between the experimental groups, since all Ss in all grades passed all three problems. Considering the advanced age of the present experimental population -- six through nineteen -- this was not unusual. The number of Ss solving each problem decreased as the problems became more difficult. The row sums in the right hand margin show the total number of problems passed by each grade group. This sum increased in an orderly fashion as the average age (school grade) increased. This regularity of change was broken only by the minor reversal of grade eight doing slightly better than grade ten. Within the table, the cell entries also reveal the orderliness of change. Inspection discloses that the above reversal was limited to only two problems, Serial Reaction 1 and 2.

The number of problems passed for individual Ss is shown in Table III. This table is a re-arrangement of the data in Table II, but Table III emphasizes the individual S's performance. It shows the same orderly increase in

TABLE II

NUMBER OF SUBJECTS IN EACH GRADE WHO PASSED SPECIFIC PROBLEMS ON THE H-P TEST TEN SUBJECTS PER GRADE

School Grade	H-P Problems								Total Problems Passed
	I	II	III	IV	V	VI	VII	VIII	
2nd	10	10	10	5	2	0	0	0	37
4th	10	10	10	6	5	4	1	0	46
6th	10	10	10	7	5	5	1	1	49
8th	10	10	10	9	8	8	5	3	63
10th	10	10	10	10	9	5	4	3	61
12th	10	10	10	10	10	10	7	4	71
Total Subjects Passed	60	60	60	47	39	32	18	11	327

TABLE III

NUMBER OF PROBLEMS PASSED
ON THE H-P TEST

Subject	School Grade					
	2nd	4th	6th	8th	10th	12th
1	5	7	8	8	8	8
2	5	6	6	8	8	8
3	4	6	6	8	8	8
4	4	6	6	7	7	8
5	4	5	6	7	6	7
6	3	4	4	6	5	7
7	3	3	4	6	5	6
8	3	3	3	6	5	6
9	3	3	3	4	5	7
10	3	3	3	3	4	6
Mean	3.7	4.6	4.9	6.3	6.1	7.1
Median	3.5	4.5	5.0	6.5	5.5	7.0
Range	2	4	5	5	4	2

ability to solve the problems as the Ss became older. The minor reversal of both measures of central tendency between grades eight and ten was present. The application of the Kruskal-Wallis non-parametric analysis of variance (20) yielded a Chi H Square of 25 (P less than .001).

Presentation Score

Table IV presents the Presentation Scores for individual Ss. This unit is the most basic measure of the H-P Test. The Presentation Score was very highly correlated with the number of problems passed (r = +.98), but, as a result of the former's extensive range within experimental sub-groups, it possesses a higher degree of discrimination and sensitivity. Because of these two reasons, only the Presentation Scores were used throughout the remainder of the analysis. Table IV possesses the same orderliness of change and the same small reversal between grades eight and ten. It is to be noted however, that the tenth grade had the lowest average Otis IQ score of all the six sub-groups. The Kruskal-Wallis analysis of variance yielded a Chi H Square of 25 (P less than .001).

Presentation Score vs. Chronological and Mental Age

In this study our main concern was of the maturation of the ability to solve the H-P Test problems. The foregoing section has pointed out that those differences in ability seen at the various school grade groupings were statistically significant. The maturation versus performance

TABLE IV

PRESENTATION SCORES ON THE H-P TEST

Subject	School Grade					
	2nd	4th	6th	8th	10th	12th
1	206	146	44	76	62	38
2	208	167	158	80	86	42
3	239	179	162	82	102	54
4	244	193	164	122	158	79
5	248	210	187	133	178	115
6	250	240	236	162	208	158
7	250	250	241	163	210	158
8	250	250	250	197	219	162
9	251	251	250	236	222	169
10	254	255	251	251	236	170
Mean	240.0	214.1	194.3	150.2	168.1	114.5
Median	249	225	211.5	147.5	193	136.5
Range	48	109	207	175	174	132

relationship was also investigated with correlational techniques.

Figure 1 depicts graphically the relationship between Presentation Score and Mental Age, chronological age, and IQ. The division of Ss into sub-groups was changed for this graph. In a developmental study the object is to extend the range of the variable being studied, and it was also necessary to use a year level breakdown to depict the relationship of the Mental Age variable. Therefore, the Ss were grouped by two year intervals rather than school grades. It appeared desireable to also present visually on the same figure the relationship between Presentation Scores and IQ. However, there was no direct relationship between IQ scores and the age units on the abcissa. Therefore, the IQ limits were arbitrarily located so as to superimpose the IQ function on top of the other two graphs.

It is obvious that all the functions showed a progressive improvement, although the IQ curve is the most erratic of the three. The orderliness of change and lack of reversals which characterize both MA and CA mean scores is especially noteworthy, for this high degree of smoothness was achieved using only eight Ss -- on the average -- for each point on the abcissa. The curves for CA and MA appear to be of a different nature, CA being negatively accelerated and MA being positively accelerated.

A Pearson product-moment r was calculated between

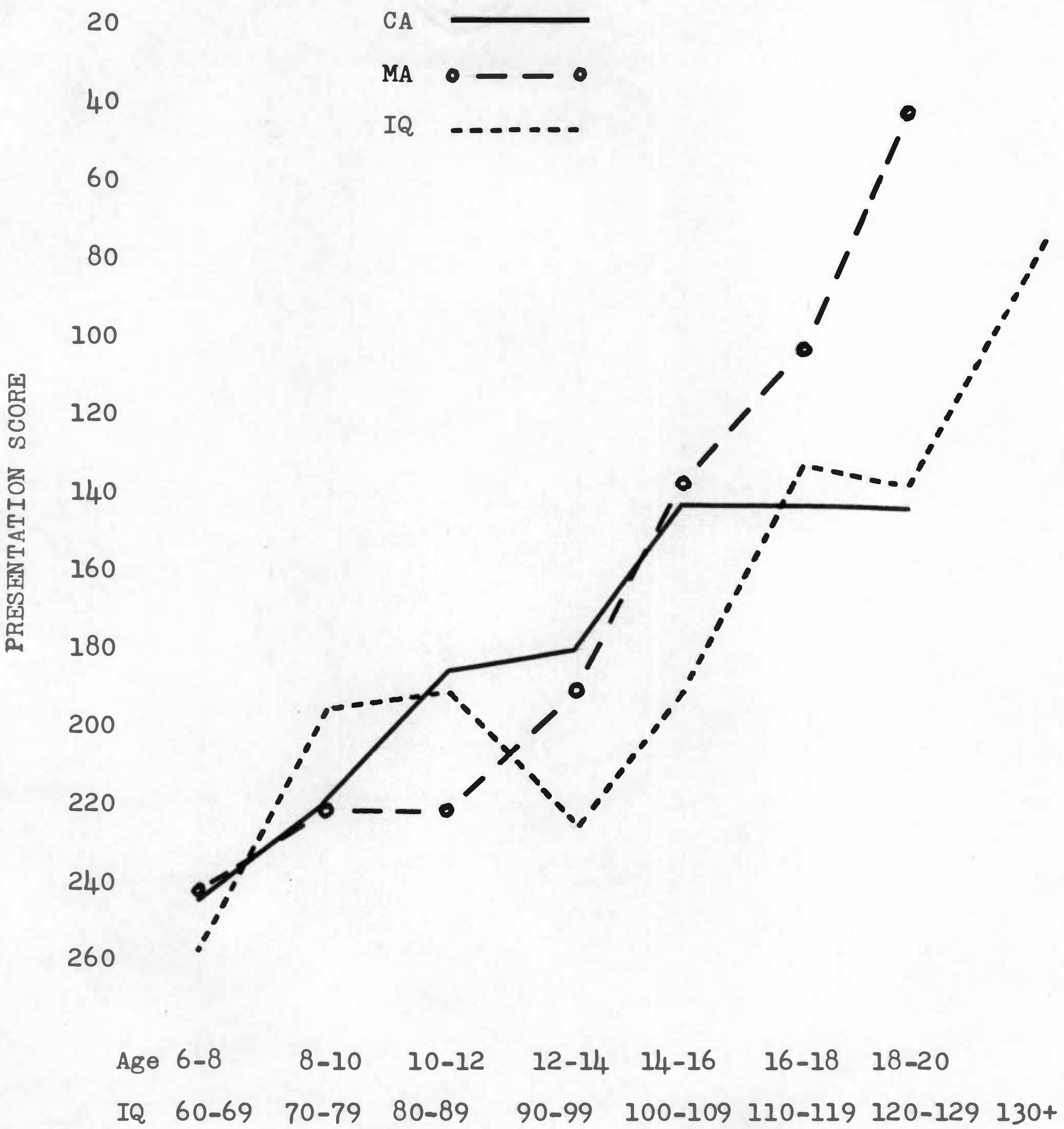

FIGURE I

PRESENTATION SCORE AS A FUNCTION OF IQ, CHRONOLOGICAL AGE AND MENTAL AGE

Presentation Score and both MA and CA; Presentation Score correlated +.76 with Mental Age and +.58 with chronological age. Both these correlations were considerably beyond the .01 level of significance. This raised the question as to which variable was most highly correlated in a purer measure of relationship. It made good experimental sense to conceive of selecting an experimental group holding CA constant and allowing MA to vary; similarly the opposite procedure seemed reasonable. Therefore, a mathematical duplication of this procedure seemed definitely appropriate. A partial correlation was calculated between Presentation Score and Mental Age holding CA constant and between Presentation Score and CA holding MA constant; the obtained values were +.60 and -.01 respectively. This pointed out that the moderately high relationship between CA and Presentation Score was a function of the high correlation (+.77) between CA and MA. The partial r of .60 between Presentation Score and MA, holding CA constant, was to some extent an underestimation of the degree of relationship. As McNemar (27) pointed out, any correlation involving only a portion of the possible range of a variable results in a smaller coefficient. This point will be pursued further in the chapter on discussion.

In addition to the relationships between Presentation Score versus CA and MA across grades, the same comparisons computed within grades were also investigated. Table V presents the rank-order correlation coefficients between

TABLE V

CORRELATIONS BETWEEN PRESENTATION SCORE VERSUS CA, MA, AND IQ FOR EACH GRADE

Comparison	School Grade					
	2nd	4th	6th	8th	10th	12th	Across all Grades
Presentation Score vs. CA	+.30	+.28	+.19	-.01	-.47	-.30	+.58
Presentation Score vs. MA	-.08	-.08	+.65	+.92	+.56	+.30	+.76
Presentation Score vs. IQ	-.23	-.17	+.47	+.83	+.62	+.31	+.38

Presentation Score versus MA, CA, and IQ for all six school grades. The initial impression gained from this table is of tremendous variability with the values ranging from -.30 to +.92. Closer inspection of the coefficients, however, shows an orderly process of change with the MA and CA variables relating in a complementary fashion. Below a certain critical point -- perhaps at about the sixth grade level -- the important variable necessary for the ability to perform on the H-P Test is CA. Then, when this crucial stage of development is passed, the MA variable quickly assumes major importance. In this connection it is helpful to refer back to Figure 1. Here the MA function was positively accelerated, starting slowly and then picking up considerable acceleration, and, by contrast, the CA function was negatively accelerated. This complementary relationship between CA and MA is further substantiated by Kaiman (21). He was able to give Stanford-Binet Intelligence Tests to only a portion of his Ss, and in this sub-sample he found the primary relationship was between Presentation Score and CA with MA being insignificantly related; these Ss ranged in age from two to four years.

Presentation Score vs. IQ

IQ scores are an index of relative brightness which show the position of an individual's mental development relative to others in the same chronological age range. Therefore, it made better theoretical sense to focus upon the relationship between Presentation Score and IQ within

grades rather than across grades. These coefficients within grades are also given in Table V. They show the same extreme range -- from -.23 to +.83 -- as the coefficients for the other variables and also the same orderly process of change. Since IQ is an index involving the theoretical ratio of MA to CA, it was not surprising that the IQ coefficients closely parellel those for MA. To complete the picture the coefficient for Presentation Score versus IQ across grades was also computed, and obtaining a value as high as +.38 under the theoretically questionable across groups procedure points out dramatically the primary relationship is between Presentation Score and MA, especially for the older Ss of the present study.

Error Score

Table VI presents the total number of errors for each S. All the distributions are highly similar with regard to both central tendency and variability. The Kruskal-Wallis Chi H Square of 1.99 shows that there is almost a non-chance (P between .80 and .90) degree of agreement. This very definite lack of difference occurred despite a highly significant difference with regard to the number of presentations (chances to err) for each grade group.

This result suggested re-arranging the data to equate the groups with regard to their chances of making errors. The Presentation Score is a combination of presentations actually performed plus the maximum number of presentations

TABLE VI

TOTAL ERROR SCORE ON THE H-P TEST

Subjects	School Grade					
	2nd	4th	6th	8th	10th	12th
1	43	74	33	50	46	38
2	47	46	47	88	66	45
3	58	65	56	39	33	54
4	63	86	51	82	95	53
5	66	72	95	62	44	55
6	47	56	33	35	44	72
7	48	26	45	33	39	50
8	53	23	38	50	71	47
9	43	35	40	34	47	74
10	30	45	49	42	45	50
Mean	49.8	52.8	48.7	51.5	53.0	53.8
Median	47.5	51.0	46.0	46.0	45.5	51.5
Range	36	63	62	55	62	36

on all problems located above a S's initial failure. Therefore, the credited number of presentations was subtracted from the Presentation Score, and the number of actual presentations was divided into the error score. The resulting index was multiplied by one hundred in conversion to percentages. These figures are located in Table VII.

These percentages presented an entirely different picture. As the Ss became older they made proportionately fewer errors and their performance tended to become less variable. The Kruskal-Wallis analysis of variance yielded a Chi H Square of 19 (P between .01 and .001). Thus there is a highly significant and orderly difference between the age groups with regard to this Error Index.

The table also suggested that there was a trend within particular grades for those Ss who have low Presentation Scores to have a lower percentage of errors. A rank correlation between Presentation Score and Error Index showed a substantial relationship (r = .611) between success at solution and fewer errors.

Error Index vs. Chronological and Mental Age

The relationship between the Error Index versus CA and MA was also investigated. A rank correlation between Error Index and age yielded values of +.56 and +.44 for chronological and Mental Ages respectively; both these values are significant beyond the .05 level.

TABLE VII

ERROR INDEX ON THE H-P TEST[a]

Subject	School Grade					
	2nd	4th	6th	8th	10th	12th
1	65	51	75	66	74	100
2	69	53	60	110	77	107
3	118	66	68	48	32	100
4	117	76	61	67	60	67
5	114	103	89	47	45	48
6	235	112	72	43	65	46
7	240	130	88	40	56	64
8	265	115	190	43	90	57
9	205	167	200	74	57	44
10	125	180	233	200	98	56
Mean	155.3	105.3	113.6	73.8	65.4	68.9
Median	121.5	107.5	81.5	57.0	62.5	60.5
Range	200	129	173	160	66	63

[a]The Error Index is obtained by dividing the error total by the actual presentations and converting to fractions by multiplying by one hundred.

A partial correlation holding CA and MA constant -- as was done with Presentation Scores -- seemed in order. When MA was held constant the relationship between Error Index and chronological age dropped to +.38; when CA was held constant the relationship between Error Index and MA dropped to +.02.

Thus, the relationship between error per cent and age was the opposite of the relationship between Presentation Score and age; success at solution was more dependent upon Mental Age, while avoidance of error was more dependent upon chronological age.

Test Time

In the process of working with the H-P Test, theoretical preconceptions were altered as information was being gathered regarding the performance of Ss on the various problems. Two changes were made in the original test, and approximately half of the Ss in grades second through sixth were re-tested on two or more occasions. As a result, their time scores are not comparable to the rest of the sample and are not included in the analysis.

In Table VIII the total time is given for the older half of the Ss. The three distributions are almost completely overlapping with almost exactly similar ranges: the differences between the means are so slight that no calculations are necessary to convince the reader of the lack of significant differences.

TABLE VIII

TOTAL TIME IN MINUTES TO COMPLETE THE H-P TEST

	TOTAL TIME			TIME/ACTUAL PRESENTATION		
Subject	8th Grade	10th Grade	12th Grade	8th Grade	10th Grade	12th Grade
1	30	21	17	.39	.34	.45
2	21	22	18	.26	.26	.43
3	21	32	15	.26	.31	.28
4	25	26	22	.20	.16	.28
5	25	18	29	.19	.18	.25
6	20	19	22	.24	.28	.14
7	20	15	21	.24	.21	.27
8	24	21	21	.21	.27	.26
9	15	18	33	.33	.22	.20
10	14	14	19	.67	.30	.21
Mean	21.5	20.6	21.7	.30	.25	.28
Median	21.0	20.0	21.0	.25	.26	.26
Range	16	18	18	.48	.18	.31

Since there was a statistically significant difference between both the number of problems solved and the Presentation Scores for the different grade groups, the total time for each S was divided by his actual number of presentations, as was done with the Error Index. These distributions are given in the right-hand panel of Table VIII.

The means and ranges do not show any orderly differences, and there is considerable overlap of the distributions. Therefore, this re-arrangement of the data shows that the grade groups were not significantly different with regard to the Time Index variable.

Sex Differences

The conclusion that girls develope more rapidly than boys is commonly accepted in the developmental literature. This relative superiority is especially characteristic in the younger age ranges, from one to ten (8). Therefore, it was important to discover if any sex differences were present in the H-P Test performance. The data were examined with regard to the two most important dimensions of the H-P Test -- successful performance and errors -- using Presentation Score and the Error Index as the specific measures.

Table IX presents this material in tabular form. The boys obtained a better performance score in three out of six grades, and a lower error score in four out of six grades.

TABLE IX

SEXUAL DIFFERENCES ON THE H-P TEST

			School Grade						
			2nd	4th	6th	8th	10th	12th	N=30
PRESENTATION SCORE	males	$\bar{X}$	237.4	231.8	214.4	122.8	198.0	105.4	185.0
		range	44	62	92	117	78	116	213
	females	$\bar{X}$	242.6	196.4	174.2	177.6	138.2	123.6	175.4
		range	46	104	207	175	160	132	216
			School Grade						
			2nd	4th	6th	8th	10th	12th	N=30
ERROR INDEX	males	$\bar{X}$	135.8	138.2	99.8	63.0	64.8	73.0	95.8
		range	200	104	130	67	38	61	222
	females	$\bar{X}$	174.8	82.4	127.4	84.6	66.0	63.2	99.7
		range	171	89	172	160	58	59	208

The possibility of seven out of twelve events going in one direction is almost a chance (P = .50) finding. Inspection of the mean scores for all males and females across grades, points to the same conclusion. The range of both measures is slightly over two hundred units and the mean differences are 9.6 and 3.9 for the Presentation Score and Error Index respectively; no statistical analysis is needed to conclude that this difference is not significant.

Although the measures of central tendency do not appear different, the females have a consistently larger degree of variability. In all six grade groups the girls' Presentation Score range exceeds the boys'. The binomial expansion shows that this degree of consistency occurs on a chance basis only three times in a hundred.

In conclusion, after analyzing the data no differences were found to indicate the need to separate the sample into sex groups for the analysis.

The H-P Test and School Achievement

Presentation Score and Grades

Appendix D presents the contemporary grade point average for almost all Ss in grades second through eighth. This data was not available to the E at the high school level. The means showed some fluctuation, but the analysis of variance yielded non-significance (Paitnik F_{range} = 3.31,

P greater than .05) A rank correlation between these contemporary averages and Presentation Score was +.34 (P less than .05, N = 39). Therefore there was a moderate and significant relationship between these two variables.

Appendix D presents the cumulative grade point average for almost all Ss. There was considerable difference between the distributions, but the analysis of variance showed that these differences in achievement only approached significance (Chi H Square = 10.2, P between .10 and .05). A rank correlation was computed between the cumulative grade average and Presentation Score. This yielded a r of +.10, which is a considerable drop from the correlation with contemporary grades. Since a slightly different set of Ss (N = 39 versus N = 57, for contemporary and cumulative grades respectively) was used in each correlation, another correlation was run to assess any unknown sampling variation. This correlation was computed between cumulative grade point average and Presentation Score, using only those cumulative Ss on whom the contemporary correlation was calculated. This resulted in a r of +.03 (N = 38), and therefore, it further substantiated the lack of relationship between performance on the H-P Test and extended school achievement as revealed through cumulative grade point average. The difference between the relationship of H-P Test performance to present and past academic achievement is most interesting. The indication that H-P Test performance is a better predictor of those variables

associated with contemporary functioning definitely should be followed-up.

Presentation Score vs. Achievement Tests

Table X presents the scores on two forms of the Stanford Achievement Test for most of the Ss in the fourth and sixth grades. When the Presentation Scores were correlated with these two measures of academic achievement, the coefficients consistently fell into a group of moderately high relationship. However, with only eight and ten Ss involved, very high coefficients are necessary to obtain statistical significance; and two of the values were significant beyond the .05 point while the third fell short of the .10 point.

Thus the Presentation Scores correlated much higher -- considering all three values -- with the Achievement Test Scores than with either past or present grade point average. In view of the criticisms which have been leveled against grades as a measure of academic achievement, it is unfortunate that the achievement test results were not available for a greater portion of the experimental population so that a more thorough investigation of this promising relationship could have been obtained.

Error Index and Grades

Both sets of grades were also inspected with regard to their relationship to the Error Index. A rank correlation yielded the following coefficients: contemporary grades

TABLE X

ACHIEVEMENT TEST SCORES
FOR MOST SUBJECTS IN
GRADES FOUR AND SIX

Subjects	4th Grade Stanford Achievement Test Form N Given 10-57	6th Grade Stanford Achievement Test Form K Given 10-55	
		Reading	Arithmetic
1	4.5	7.6	4.9
2	4.6	4.9	4.0
3	5.9	8.4	4.1
4	---	8.0	3.6
5	4.0	5.8	4.2
6	4.0	3.1	3.8
7	4.3	4.8	4.6
8	---	2.9	1.4
9	5.7	4.3	3.4
10	3.7	4.9	4.0
Correlation with Presentation Score	+.42	+.65	+.50
P	$P > .10$	$P < .05$	$.10 > P > .05$

versus Error Index, r = +.06 (N = 39) and cumulative grades versus Error Index, r = +.24 (N = 60). The magnitude of the latter coefficient approached the .05 level of significance.

IQ vs. Grades

For the purpose of comparison, some other relationships were evaluated. The IQ scores correlated with cumulative grade point average +.63 (N = 56) and with contemporary grade point average +.69 (N = 39). The relationship between the two sets of grades was quite high, +.88 (N = 37). Thus as a predictor of school achievement as reflected in grades, the H-P Test failed to come up to the level of both IQ test scores and past achievement. However, the relationship between the IQ variable versus the past grade variable and current grades was -- in one sense -- contaminated, for all the teachers had daily access to each S's file where this material was recorded.

Summary

For the purpose of a more organized over view of all of the preceeding material, two summary tables are presented. Since the number of problems passed is so highly correlated with Presentation Score (r = +.98), the coefficients between number of problems passed and the other measures will be highly similar to the coefficients between Presentation Score and these measures. Table XI is an

TABLE XI

INCOMPLETE CELL SUMMARY OF
H-P TEST INTERCORRELATIONS

	Number of Problems Passed	Presentation Score	Total Errors	Error Index	Total Time	Time Index
Number of Problems Passed	X	+.98	----	----	----	----
Presentation Score		X	-.48	+.61	-.24*	-.24*
Total Errors			X	-.09	+.61*	----
Error Index				X	----	----
Total Time					X	-.34*
Time Index						X

* N = 30

incomplete cell table depicting the relationship of selected H-P Test measures to each other. Table XII presents a summary of the relationship of the two main dimensions of the H-P Test with the variables of the secondary analysis.

TABLE XII

INCOMPLETE CELL SUMMARY OF
H-P TEST MEASURE WITH
VARIOUS CRITERIA

	Mental Age	Chronological Age	MA - CA **	CA - MA **	Contemporary Grade Average	Past Grade Average
Presentation Score	+.76	+.58	+.60	-.01	+.34*	+.10
Error Index	+.44	+.56	+.38	+.02	+.06*	+.24
IQ					+.69*	+.63
Past Grade Average					+.88*	
Mental Age		+.77				

* N = 40

**Partial Correlations

CHAPTER V

DISCUSSION

Future Research

The moderate to substantial relationship between the H-P Test and both CA and MA has been established in both the present and other independent investigations.

Thus, it appears that the H-P Test may be a promising psychological instrument; and, as such, the need for future research is both apparent and imperative. One favorable aspect of the H-P Test is the large "spread" of Presentation Scores within any one school grade, for within a particular school grade the Ss are quite comparable with regard to both CA and MA. Here is a large degree of unaccounted for variation.

One course for future research would be the attempt to relate this unexplained variability to other pertinent variables. Following the lead of Worley (38) and Leslie (24), the relationship between H-P Test performance and some measure of psychological adjustment should be undertaken.

Stimulus Value of Test
Candy as Reward

One interesting aspect of the H-P Test was its high

stimulus value for almost all Ss. The apparatus itself was particularly fascinating for the younger Ss, and the problem-solving nature of the task intrigued the older Ss.

Prior to the experiment there was some uncertainty regarding the reward's suitability throughout the extended age range of six to nineteen years. Only two Ss initially rejected the pile of candy they accumulated. One sixth grader explained that her dentist had forbidden her to eat candy because of considerable difficulty with cavities. The other S was a twelfth grader. Questioning revealed that he felt guilty about taking so much candy from the E. Contrary to the High School principal's prediction, his students still were quite interested in candy, and the experimental Ss were seen walking along the corridors eating their candy throughout the course of the day.

Unreliability of IQ Criterion

The variability of the Ss' IQ scores was perplexing. The possibility of an individual S's score changing slightly upon re-testing is not only reasonable but empirically established. The magnitude of some of the changes was, however, unusual. The change in mean scores for grade groups was sometimes high. On occasion there was a relatively high rank correlation between two or three sets of scores which suggests the possibility of a constant error of

measurement. This constant error could result from the tests being given by teachers unsophisticated in the administration of psychological tests and uncomprehending of the necessity for rigid attention to standardized procedures. The variable errors which result in low or even negative rank correlations are, however, of a different nature; the cause of this fluctuation can not be even speculated upon. This suggests that under more optimal intelligence testing, the relationship between MA and H-P Test performance might be higher, for the unreliability of a criterion itself reduces any correlations with other variables.

Problem IIIa

As the data were being collected it became increasingly apparent that there was a disproportionate gap between Problem III, 60 second Delayed Reaction 3, and Problem IV, Single Alternation. Successful solution of the single alternation problem demands that an intermediate step must be realized; the S must become aware of the fact that the reward is always located behind one of the two most lateral doors. The gaining of this knowledge results in the S forming the "End Concept."

An attempt was made to fill in the gap between Problem III and IV with a new problem of intermediate difficulty. This new problem, Problem IIIa, gave the S an additional cue

to the location of the reward. For trial I a red poker chip was placed in front of the center door while the S watched. The S was then turned away from the apparatus, and the poker chip was placed in front of the same door as the reward was placed. The S was then allowed to search. Upon finding the reward he was again asked to turn around. The chip and reward were switched to the opposite end door in the single alternation pattern. This problem consisted of ten trials of two responses each. It was then followed by the standard single alternation. It was designed to give additional cues to the "End Concept" which so many Ss found difficult to conceptualize.

The Ss who originally failed single alternation were re-tested with Problem IIIa, and if they passed IIIa were given another chance at Problem IV. The addition of this new problem resulted in some Ss being able to go on and solve other more difficult problems. However, since they received two chances to solve Problem IV and since several months passed between these two testings, they are not comparable to the other Ss and this problem was not included in the analysis. It did, however, result in somewhat better discrimination of the test within the group of the poorest scores. Appendix E lists these differences.

The Reversal of The Eighth and Tenth Grades

In the sub-division of the experimental population by

school grades, grade eight had a better average Presentation Score than grade ten. Ordinarily such reversals are not infrequent in small sample research where the effect of chance sampling variations is enhanced, because a few very deviant cases exert considerable influence upon the measures of central tendency. However, since the data were orderly in other fashions, this reversal was investigated further. Inspection of the IQ scores found grade ten to have the lowest average score of all groups. The average MA for all grade groups was calculated. The following averages were obtained reading from second to twelfth grade respectively: 7.44, 10.18, 12.42, 14.21, 14.38, and 15.83. The average MA for grade ten was higher than grade eight by a factor of only two months. Since there should be -- and was -- an average difference of two years chronological age, it suggests that the obtained reversal was merely a sampling variation rather than a "real" difference of ability (which would occur again if the experiment were replicated) favoring children in the eighth grade generally speaking. This conclusion is further borne out by inspection of Figure 1. It depicts the relationship between MA and Presentation Score when the Ss are grouped by MA units rather than by school grade units. Here the progression of mean Presentation Scores is extremely orderly and without reversals.

Examination of A Deviant Case

An 11 year 11 months old, sixth grade experimental S received a Presentation Score of 44. This was well off the distribution of her age group; and, in addition, her score was bettered by only two twelfth graders.

It seemed most appropriate to secure some additional information about this unusual young lady. The E visited her home and jointly interviewed both the mother and daughter. This girl was 12 years and 5 months old at the date of the interview. In appearance she was an attractive, blonde, physically mature girl who could easily pass for being four or five years older. She was initially mildly self-conscious in the home interview, though this was not present in the testing at school, but she became more at ease as the interview progressed. She was friendly, verbally fluent, and emotionally responsive.

The mother reported that her daughter had been quite early in her development: she began walking at nine months, she said her first words at approximately the same time, at thirteen to fourteen months she was putting together three and four word sentences, and at eighteen months she was speaking with noteable fluency. She learned to jump rope, skate, and ride a bike considerably before her same age play-mates. Furthermore, she displayed a surprising amount of confidence and announced she could and would do things with unusual

determination.

She has had no serious illnesses or accidents and has always been in good health. She has had, in fact, only two of the usual childhood diseases.

She became interested in school and school activities about age four, engaging in school activities and fantasies with a girl friend. About this time she would scribble, imitating writing, and became annoyed when her mother couldn't read her scribbling, because her mother could read everyone else's writing. She was beginning to print prior to starting first grade, and she learned writing so fast that her teacher asked the mother if she had been receiving instruction. The mother, a former school teacher herself, maintained that she had not given her daughter any instruction; she realized that different teachers used different methods and she had feared causing confusion on this score.

Her academic record is excellent and her grade point average fluctuates from A- to B+. Last month, upon finishing sixth grade, she took the Stanford Achievement Test (form unknown) and scored at the eleventh grade level. She has received three IQ testings on the Otis at two year intervals, receiving scores of 137, 103, and 127 respectively. (This points up again the huge sampling fluctuation characteristic of the IQ scores.)

For the last several years, she has shown considerable leadership in the neighborhood gang, where her mother states her ability to persuade and "get-around people" is unusual.

When asked her interests she promptly stated cooking and sewing, although further questioning revealed these are not highly developed skills but are spontaneous activities. As for the future, she candidly states she wants to be a housewife and possibly teach on the side. She would like to go to the University, and admits uncertainty regarding a major course. After graduation she would like to work in an office for a short while.

She is currently taking dramatic lessons in the Junior Carousel but does not appear to be highly motivated. She dismisses acting as a career stating, "Frankly, I'm not the glamour girl type."

She is interested in sports, both as a spectator and participant. She watches school and TV football and baseball; she plays ping-pong, badminton, and especially likes swimming.

She appears well-entrenched in teen-age interests and activities though only twelve years old. Long telephone calls plague the mother. She "hoards" money to spend on clothes and "rock and roll" records. She is "mad about clothes" and this deep concern appeared throughout the interview. She participates socially and has a circle of close friends.

All in all, this girl appeared to be a socially, physically and mentally precocious young lady who appeared

happy and well-adjusted. She appeared to have good relationship with her mother and was quite solidly femininely identified. She did exceptionally well in school without having to study hard. She appeared to be well accepted by her peers and family and had adequate and realistic self-esteem.

CHAPTER VI

SUMMARY AND CONCLUSIONS

The present study investigated the ontogenetic relationship between the ability to perform on the Hunter-Pascal Concept Formation Test (H-P Test) and chronological age (CA) and Mental Age (MA). The H-P Test uses an apparatus consisting of a white, wooden wall with five black doors; the doors are exactly alike. The subject's task was to discover which door concealed a reward consisting of "corn candy." The pattern of the placement of the reward behind the doors varied through eight problems arranged in order of increasing difficulty. The beginning tasks are modifications of the delayed reaction and double alternation problems of Walter S. Hunter, and the more difficult problems are a serial combination of single and double alternation. The H-P Test was easy to administer, and it had high stimulus appeal for the subjects. It required little time, seldom more than a half-hour.

A sample, stratified with regard to CA, of ten "normal" children, five males and five females, was drawn from each of the even numbered school grades, second through twelfth. This resulted in a total of sixty subjects.

The H-P Test yielded two main dimensions of performance: the Presentation Score and the Error Index. The

differences between the scores obtained by the various grade groups on both measures was significant beyond the .01 level. These two measures were related to CA and MA and orderly graphic functions were obtained. Although the small sample size required caution in interpretation, the MA curve was slightly positively accelerated and the CA curve was slightly negatively accelerated.

The relationship between these two test dimensions and CA and MA was also investigated with correlational statistics. Presentation Score correlated +.76 with MA and +.58 with CA. A partial correlation was computed holding both variables constant, and when MA was partialed out the correlation of Presentation Score with CA dropped to -.01. This pointed out that the MA variable, with a remaining r of +.60, possessed the primary relationship. When correlation coefficients were computed for Presentation Score versus MA and CA within all six school grade groups, an interesting relationship was discovered. The CA and MA variables had a complementary relationship. For the younger subjects, CA was most highly related to success at the H-P Test, but, once a certain age was reached (about the sixth grade), the primary relationship was with MA.

The Error Index correlated +.56 with CA and +.44 with MA. When the effect of CA was partialed out the relationship between Error Index and MA dropped to +.02, leaving a remaining r of +.38 for CA. Thus the ability to learn the problems

appeared most dependent upon MA while being able to avoid errors was related most highly to CA.

Both test dimensions were also investigated regarding their relationship to academic achievement as revealed in school grade point averages. The Presentation Score correlated +.34 with contemporary grades and +.10 with past grades: the Error Index correlated +.06 with contemporary grades and +.24 with cumulative grades. The Presentation Score had a moderate relationship with Stanford Achievement Test Scores (+.42, +.50, and +.65) in the portion of the experimental population (one-third) for which this information was available.

Both H-P Test measures were lesser predictors of contemporary grades than IQ scores or past grades which correlated +.69 and +.88 respectively. Since both the IQ scores and past grades were known to the grading teachers, there is some question of the contamination of judgments because of the lack of independence.

Contrary to many developmental investigations which demonstrate earlier maturation by girls, they did not demonstrate any superiority of performance over that of the boys.

A direction for further research has been suggested in the area of the relationship between the H-P Test performance and psychological adjustment.

As a result of this and similar studies it has been demonstrated that a series of problems such as those of the

H-P Test yield orderly data which imply increased cortical efficiency as we ascend the phylogenetic scale, and this efficiency is on a continuum from the lower vertebrates to adult human organisms.

BIBLIOGRAPHY

BIBLIOGRAPHY

1. Adams, D. K. Experimental studies of adaptive behavior in cats. Comp. Psychol. Monogr., 1929, 6, No. 27, 1-166.

2. Boring, E. C. A history of experimental psychology. (2nd ed.) New York: Appleton-Century-Crofts, 1950.

3. Collins, Martha P. A study of the development of concept formation as related to the maturation of place response, delayed reaction, and single alternation in a selected group of children, ages thirty-four to sixty months. Unpublished masters thesis, Univer. of Tennessee, 1957.

4. Davis, H. C. A psychological study of chronic alcoholics. Unpublished doctoral dissertation, Univer. of Tennessee, 1959.

5. Gellerman, L. W. The double alternation problem I: The behavior of monkeys in a double alternation temporal maze. J. genet. Psychol., 1931, 39, 50-72.

6. Gellerman, L. W. The double alternation problem II: The behavior of children and human adults in a double alternation temporal maze. J. genet. Psychol., 1931, 39, 197-226.

7. Gellerman, L. W. The double alternation problem III: The behavior of monkeys in a double alternation box-apparatus. J. genet. Psychol., 1931, 39,359-392.

8. Gesell, A., & Ilg, Frances L. The child from five to ten. New York: Harper, 1946.

9. Heron, W. T. Complex learning processes. In F. A. Moss (Ed.) Comparative psychology. (Rev. ed.) New York: Prentice-Hall, 1946.

10. Hodges, A. A developmental study of concept behavior as measured by the double alternation problem. Unpublished doctoral dissertation, Univer. of Tennessee, 1953.

11. Hunter, W. S. The delayed reaction in animals and children, Behav. Monogr., 1913, 2, No. 1.

12. Hunter, W. S. The delayed reaction in a child. Psychol. Rev., 1917, 24, 75-87.

13. Hunter, W. S. Kinaesthetic sensory processes in the white rat. Psychol. Bull., 1918, 15, 36-37.

14. Hunter, W. S. The behavior of raccoons in a double alternation maze. J. genet. Psychol., 1928, 35, 374-388.

15. Hunter, W. S. The sensory control of the maze habit in the white rat. J. genet. Psychol., 1929, 36, 505-537.

16. Hunter, W. S. & Bartlett, Susan C. Double alternation behavior in young children. J. exp. Psychol., 1948, 38, 558-567.

17. Hunter, W. S. & Hall, Beatrice M. Double alternation behavior of the white rat in a spatial maze. J. comp. physiol. Psychol.,1941, 32, 253-266.

18. Hunter, W. S., & Nagge, J. W. The white rat and the double alternation temporal maze. J. genet. Psychol., 1931, 39, 303-319.

19. Jenkins, W. O. An investigation of double alternation behavior in the white rat in the Skinner box. Unpublished masters thesis, Brown Univer., 1938.

20. Jenkins, W. O. Quick and dirty statistics. Unpublished manuscript, Univer. of Tennessee, 1956.

21. Kaiman, B. A developmental study of concept formation in children. Unpublished doctoral dissertation, Univer. of Tennessee, 1958.

22. Karn, H. W. The behavior of cats on the double alternation problem in the temporal maze. J. comp. physiol. Psychol., 1938, 26, 201-208.

23. Karn, H. W. & Malamud, H. R. The behavior of dogs on the double alternation problem in the temporal maze. J. comp. physiol. Psychol., 1939, 27, 461-466.

24. Leslie, D. L. The relationship between performance on the Hunter-Pascal Concept Formation Test and the neurotic-psychotic dichotomy. Unpublished doctoral dissertation, Univer. of Tennessee, 1957.

25. McCord, F. The delayed reaction and memory in rats. I Length of delay. J. comp. physiol. Psychol., 1939, 27, 1-37.

26. MacCorquodale, K. An analysis of certain cues in the delayed response. J. comp. physiol. Psychol., 1947, 40, 239-253.

27. McNemar, Q. Psychological statistics. New York: Wiley, 1949.

28. Munn, N. L. Handbook of psychological research on the rat. Boston: Houghton Mifflin, 1950.

29. Munn, N. L. The evolution and growth of human behavior. Boston: Houghton Mifflin, 1955.

30. Pascal, G. R., & Jenkins, W. O. The Hunter-Pascal concept formation test. Unpublished manuscript, Univer. of Tennessee, 1957.

31. Pascal, G. R., & Zax, M. Double alternation performance as a measure of educability in cerebral palsied children. Amer. J. ment. Defic., 1955, 59, 658-665.

32. Pascal, G. R., Stolorow, L. M., Zabarenko, R. N., & Chambers, C. S. The delayed reaction in mental defectives. Amer. J. ment. Defic., 1951 , 56, 152-160.

33. Schlosberg, H., & Katz, A. Double alternation lever pressing in the white rat. Amer. J. Psychol., 1943, 56, 274-282.

34. Skalet, Magda. The significance of delayed reaction in young children. Comp. Psychol. Monogr., 1931, 7, No. 4, 1-82.

35. Tapp, Marian J. A study of the development of concept formation as related to the maturation of place response, delayed reaction, and single alternation in a selected group of children, ages five months to thirty-two months. Unpublished masters thesis, Univer. of Tennessee, 1957.

36. Tinklepaugh, O. L. An experimental study of representative factors in monkeys. J. comp. physiol. Psychol., 1928, 8, 197-236.

37. Watson, J. B. Psychology as a behaviorist views it. Psychol. Rev., 1913, 20, 158-177.

38. Worley, F. The double alternation card test as an indicator of severity of mental illness. Unpublished masters thesis, Univer. of Tennessee, 1955.

39. Yerkes, R. M. The mind of a gorilla. Genet. Psychol. Monogr., 1927, 2, 1-193.

APPENDICES

APPENDIX A

TABLE XIII

CHRONOLOGICAL AGE FOR
ALL SUBJECTS

Subjects	School Grade					
	2nd	4th	6th	8th	10th	12th
1	8-6-11	9-0-10	11-11-12	13-1-13	15-7-6	17-8-9
2	7-7-11	9-3-11	11-8-15	14-0-16	15-10-17	16-3-3
3	7-3-21	9-9-6	11-9-6	14-1-9	15-4-13	17-5-24
4	7-6-1	10-5-28	11-2-25	13-7-16	16-0-10	17-8-14
5	7-8-25	9-4-18	11-9-16	13-2-22	16-3-7	18-11-23
6	7-8-22	9-6-15	12-10-9	14-5-21	15-7-28	17-7-12
7	6-11-12	8-10-5	11-4-27	13-2-10	16-8-29	18-7-1
8	7-3-10	8-11-27	12-4-28	14-0-0	15-2-5	18-1-28
9	7-8-12	9-3-19	11-9-0	13-11-25	16-1-11	17-11-5
10	7-5-29	9-3-5	10-6-26	13-5-20	17-1-2	17-2-16

TABLE XIV

MENTAL AGE FOR
ALMOST ALL
SUBJECTS

Subject	School Grade					
	2nd	4th	6th	8th	10th	12th
1	7-11	9-6	14-8	17-8	17-1	16-6
2	8-1	9-5	12-0	16-7	15-9	19-0
3	6-7	12-0	13-7	15-4	15-5	17-3
4	6-0	11-6	14-5	14-10	11-11	17-1
5	7-6	9-3	13-7	14-0	15-7	13-2
6	7-2	8-9	11-5	13-5	14-7	17-1
7	8-4	9-11	12-4	14-9	11-8	12-5
8	7-9	----	8-7	11-2	16-0	----
9	8-3	11-3	12-2	12-0	14-8	12-9
10	6-10	10-0	11-6	12-4	11-1	17-3

APPENDIX B

TABLE XV

PRESENTATION SCORE RAW DATA

Problem	2nd Grade Subjects									
	1	2	3	4	5	6	7	8	9	10
I	0	0	0	0	0	0	0	0	0	0
II	5	0	1	1	0	0	0	0	0	0
III	1	2	2	1	0	0	0	0	1	4
IV	6	4	6	12	18	20	20	20	20	20
V	4	12	40	40	40					
VI	50	50								
VII										
VIII										
Presentations	66	68	49	54	58	20	20	20	21	24
Credited	140	140	190	190	190	230	230	230	230	230
Presentation Score	206	208	239	244	248	250	250	250	251	254

TABLE XV (continued)

PRESENTATION SCORE RAW DATA

Problem	4th Grade Subjects									
	1	2	3	4	5	6	7	8	9	10
I	0	0	1	0	0	0	0	0	1	2
II	0	0	0	0	0	0	0	0	0	0
III	0	1	0	0	0	0	0	0	0	3
IV	10	8	6	10	16	10	20	20	20	20
V	4	8	12	8	4	40				
VI	10	10	20	35	50					
VII	42	60	60	60						
VIII	80									
Presentations	146	87	99	113	70	50	20	20	21	25
Credited	---	80	80	80	140	190	230	230	230	230
Presentation Score	146	167	179	193	210	240	250	250	251	255

TABLE XV (continued)

PRESENTATION SCORE RAW DATA

Problem	6th Grade Subjects									
	1	2	3	4	5	6	7	8	9	10
I	0	0	0	0	0	0	1	0	0	1
II	0	0	0	0	0	0	0	0	0	0
III	0	0	0	0	0	0	0	0	0	0
IV	4	4	4	2	12	6	10	20	20	20
V	4	4	8	12	20	40	40			
VI	10	10	10	10	15					
VII	18	60	60	60	60					
VIII	8									
Presentations	44	78	82	84	107	46	51	20	20	21
Credited	--	80	80	80	80	190	190	230	230	230
Presentation Score	44	158	162	164	187	236	241	250	250	251

TABLE XV (continued)

PRESENTATION SCORE RAW DATA

Problem	8th Grade Subjects									
	1	2	3	4	5	6	7	8	9	10
I	0	0	0	0	0	0	0	0	0	1
II	0	0	0	0	0	0	0	0	0	0
III	0	0	0	0	0	0	0	0	0	0
IV	4	6	4	4	4	4	4	6	6	20
V	4	4	8	4	12	8	4	36	40	
VI	40	10	10	10	25	10	15	15		
VII	12	12	36	24	12	60	60	60		
VIII	16	48	24	80	80					
Presentations	76	80	82	122	133	82	83	117	46	21
Credited	--	--	--	--	--	80	80	80	190	230
Presentation Score	76	80	82	122	133	162	163	197	236	251

TABLE XV (continued)

PRESENTATION SCORE RAW DATA

Problem	10th Grade Subjects									
	1	2	3	4	5	6	7	8	9	10
I	0	0	0	0	0	0	0	0	0	0
II	0	0	0	0	0	0	0	1	0	0
III	0	0	0	0	0	0	0	0	0	0
IV	4	6	4	6	4	10	4	16	12	6
V	4	4	4	4	4	8	16	12	20	40
VI	10	10	10	20	30	50	50	50	50	
VII	12	18	12	48	60					
VIII	32	48	72	80						
Presentations	62	86	102	158	98	68	70	79	82	46
Credited	--	--	--	--	80	140	140	140	140	190
Presentation Score	62	86	102	158	178	208	210	219	222	236

TABLE XV (continued)

PRESENTATION SCORE RAW DATA

Problem	12th Grade Subjects									
	1	2	3	4	5	6	7	8	9	10
I	0	0	0	0	0	0	0	0	0	0
II	0	0	0	0	0	1	0	0	0	0
III	0	0	0	0	0	0	0	0	0	0
IV	4	6	4	4	4	4	4	4	2	4
V	4	4	4	4	4	4	4	8	4	16
VI	10	10	10	15	15	15	10	10	35	10
VII	12	6	12	24	12	54	60	60	48	60
VIII	8	16	24	32	80	80			80	
Presentations	38	42	54	79	115	158	78	82	169	90
Credited	--	--	--	--	--	--	80	80	--	80
Presentation Score	38	42	54	79	115	158	158	162	169	170

TABLE XVI

ERROR SCORE RAW DATA

Problem	2nd Grade Subjects									
	1	2	3	4	5	6	7	8	9	10
I	0	0	0	0	0	0	0	0	0	0
II	5	0	1	1	0	0	0	0	0	0
III	1	3	2	2	0	0	0	0	1	6
IV	9	11	20	37	28	47	48	53	42	24
V	2	10	35	23	38					
VI	26	23								
VII										
VIII										
Total Errors	43	47	58	63	66	47	48	53	43	30

TABLE XVI (continued)

ERROR SCORE RAW DATA

Problems	4th Grade Subjects									
	1	2	3	4	5	6	7	8	9	10
I	0	0	2	0	0	0	0	0	1	2
II	0	0	0	0	0	0	0	0	0	0
III	0	1	0	0	0	0	0	0	0	5
IV	14	10	5	23	32	21	26	23	34	38
V	3	9	11	11	6	35				
VI	11	5	20	22	34					
VII	14	21	27	30						
VIII	32									
Total Errors	74	46	65	86	72	56	26	23	35	45

TABLE XVI (continued)

ERROR SCORE RAW DATA

Problem	6th Grade Subjects									
	1	2	3	4	5	6	7	8	9	10
I	0	0	0	0	0	0	1	0	0	1
II	0	0	0	0	0	0	0	0	0	0
III	0	0	0	0	0	0	0	0	0	0
IV	9	11	9	4	24	8	19	38	40	48
V	6	8	10	13	29	25	25			
VI	7	10	7	4	11					
VII	7	18	30	30	31					
VIII	4									
Total Errors	33	47	56	51	95	33	45	38	40	49

TABLE XVI (continued)

ERROR SCORE RAW DATA

Problems	8th Grade Subjects									
	1	2	3	4	5	6	7	8	9	10
I	0	0	0	0	0	0	0	0	0	2
II	0	0	0	0	0	0	0	0	0	0
III	0	0	0	0	0	0	0	0	0	0
IV	9	12	11	5	7	4	4	14	12	40
V	9	8	9	11	6	8	10	13	22	
VI	13	20	2	13	8	1	4	7		
VII	11	12	8	17	4	22	15	16		
VIII	8	36	9	36	37					
Total Errors	50	88	39	82	62	35	33	50	34	42

TABLE XVI (continued)

ERROR SCORE RAW DATA

Problems	10th Grade Subjects									
	1	2	3	4	5	6	7	8	9	10
I	0	0	0	0	0	0	0	0	0	0
II	0	0	0	0	0	0	0	1	0	0
III	0	0	0	0	0	0	0	0	0	0
IV	9	16	5	12	10	11	10	36	17	14
V	6	8	6	11	1	8	8	9	12	31
VI	11	7	2	6	11	25	21	25	18	
VII	14	12	6	28	22					
VIII	6	23	14	38						
Total Errors	46	66	33	95	44	44	39	71	47	45

TABLE XVI (continued)

ERROR SCORE RAW DATA

Problem	12th Grade Subjects									
	1	2	3	4	5	6	7	8	9	10
I	0	0	0	0	0	0	0	0	0	0
II	0	0	0	0	0	1	0	0	0	0
III	0	0	0	0	0	0	0	0	0	0
IV	8	11	10	11	10	7	7	11	6	12
V	6	8	5	8	8	10	7	5	6	14
VI	7	12	12	7	3	3	8	1	6	1
VII	8	4	15	10	8	17	28	30	21	23
VIII	9	10	12	17	26	34			35	
Total Errors	38	45	54	53	55	72	50	47	74	50

APPENDIX C

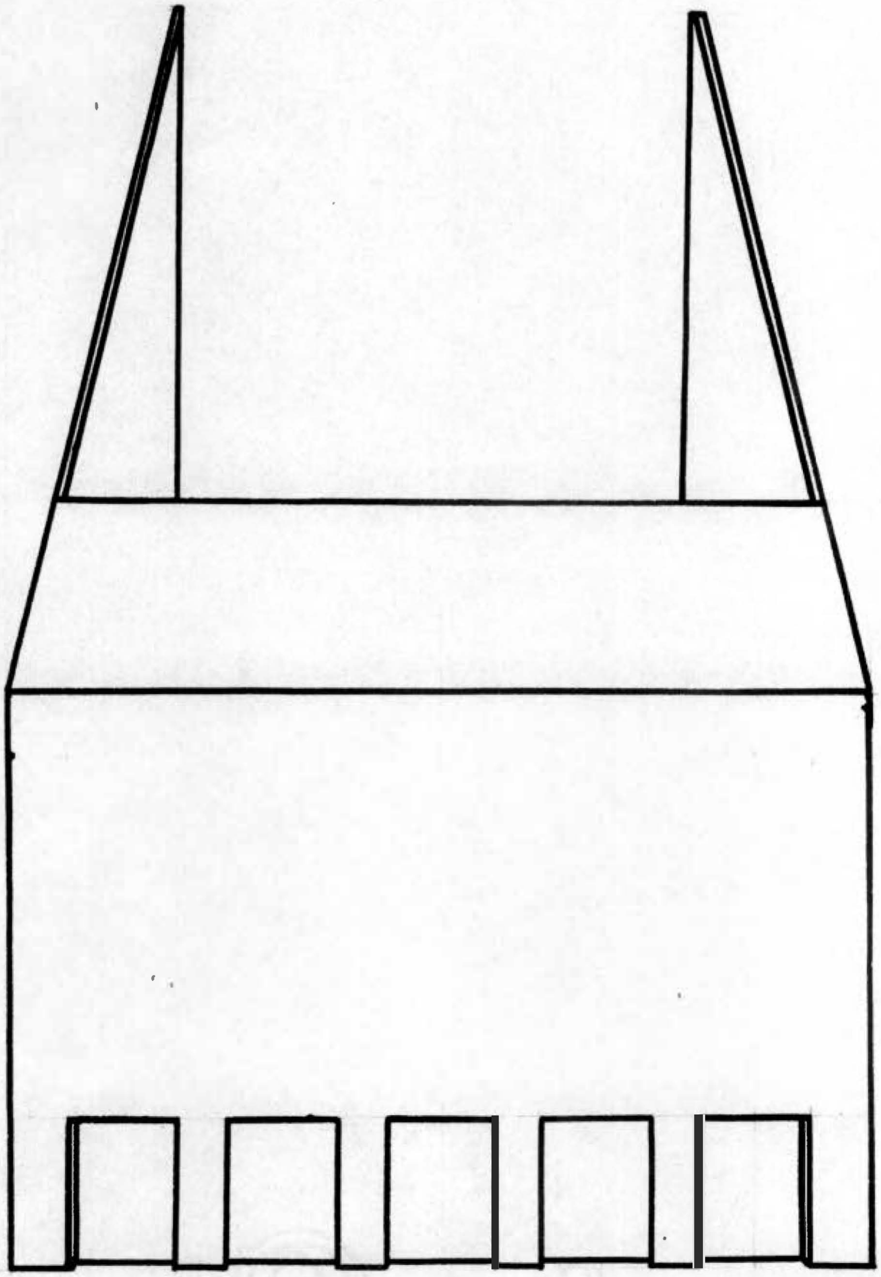

FIGURE 2

THE HUNTER-PASCAL TEST APPARATUS
SCALE OF $\frac{1}{4}$" = 1"

APPENDIX D

TABLE XVII

PAST GRADE POINT AVERAGE
FOR SUBJECTS GROUPED
BY SCHOOL GRADES

Subject	School Grade					
	2nd	4th	6th	8th	10th	12th
1	1.75	3.00	4.00	3.25	2.65	2.75
2	1.25	2.50	2.96	3.58	2.59	2.25
3	1.50	3.80	3.79	2.90	3.00	1.44
4	2.25	2.13	3.66	2.80	2.04	3.17
5	3.00	2.62	3.00	2.08	1.31	1.57
6	3.25	2.38	2.12	----	2.21	2.81
7	3.50	3.38	3.04	3.00	1.35	2.50
8	1.50	2.20	1.32	1.58	2.69	----
9	----	3.70	3.26	2.50	2.41	2.29
10	2.50	2.40	3.61	2.30	1.47	2.54
Mean	2.28	2.81	3.08	2.67	2.17	2.37
Median	2.25	1.67	2.68	2.00	1.69	1.73

TABLE XVIII

CONTEMPORARY GRADE POINT AVERAGE
FOR SUBJECTS GROUPED BY
SCHOOL GRADE

Subject	School Grade			
	2nd	4th	6th	8th
1	2.00	2.64	3.83	3.75
2	1.83	2.86	3.17	4.00
3	2.17	4.00	3.89	3.25
4	1.33	2.14	3.50	3.25
5	3.00	2.36	3.17	2.25
6	2.50	2.21	1.56	2.00
7	2.67	3.21	3.05	3.50
8	2.00	----	1.22	1.50
9	3.00	3.50	3.78	2.25
10	2.17	1.93	3.75	2.75
Mean	2.27	2.76	3.09	2.85
Range	1.67	2.07	2.67	2.50

APPENDIX E

TABLE XIX

RETESTING WITH PROBLEM IIIa

Problem	Grade: 2nd				4th			6th			8th
Subject Number	6	7	8	10	7	9	10	8	9	10	10
IIIa	0	0	1	2	0	0	0	0	0	0	0
IV	20	20	20	4	4	18	18	20	8	20	20
V				40	8	40	4		28		
VI					10		50		50		
VII					18						
VIII					80						
Presentations	20	20	21	46	120	58	72	20	86	20	20
Credited	230	230	230	190	0	190	140	230	140	230	230
Presentation Score	250	250	251	236	120	248	212	250	226	250	250
Original Presentation Score without IIIa	250	250	250	251	250	251	255	250	250	251	251

APPENDIX F

TABLE XX

MEAN OTIS IQ SCORES
FOR ALL SUBJECTS

Subject	School Grade					
	2nd	4th	6th	8th	10th	12th
1	93	106	123	135	116	110
2	107	102	102	121	106	127
3	90	123	116	112	106	115
4	80	110	128	111	79	114
5	97	98	115	106	104	88
6	92	92	89	96	99	114
7	121	112	108	113	78	83
8	107	---	69	82	111	---
9	108	121	104	89	98	85
10	91	108	109	93	74	115
Mean	98.6	108.0	106.3	105.8	97.1	105.7
Range	41	31	59	53	37	44

APPENDIX G

INSTRUCTIONS TO EXPERIMENTAL SUBJECTS

Human experimental Ss often display considerable variability with regard to their expectations toward and behavior in a psychological testing situation. In addition, the more sophisticated Ss have formulated certain attitudes about "psychologists." As a result of the E's adult status and his method of functioning in the school, he was, in a very real sense, an authority figure; with the older Ss some conflicts in this area were certain to have been present. Therefore, the pre-testing, rapport building talk was of paramount importance.

It was decided that a rigid, verbatim series of pre-test instructions could not be sufficiently flexible and would appear too formal. Also, considering the twelve year discrepancy between the youngest and oldest Ss, it was not possible to prepare a formalized talk which would be appropriate to all age levels. The points below constituted the outline of the E's instructions. If any general questions arose they were answered in more detail.

The E, like the Ss, was still in school. One of his requirements to graduate was to design and carry out an experiment. This experiment was extremely important to the E, and he would be very grateful to the Ss for their cooperation.

Since only a small number of students in each class

were being tested, it seemed only reasonable that the Ss might be curious, if not apprehensive, as to why they in particular had been chosen. This concern did not appear until the eighth grade Ss were reached, and, past this level, approximately one-fourth of the sample verbalized their concern spontaneously. The E routinely explained that all the students in their class had been arranged according to their age in years and months and individual students were selected only on this basis.

Some degree of looking for "gimmicks" and skepticism regarding the ease of the beginning problems was expected from the older Ss. It was explained that the series of problems were designed to be used by Ss ranging in age from several months through adulthood. Therefore, the initial problems would be quite easy, but the problems would become more difficult as the test progressed.

All Ss started with the first problem, and in a similar fashion, all Ss received the same token rewards of "corn candy."

The test was described as a series of different problems which the Ss were asked to solve. All questions regarding the nature or specific details of the test were put off with "You'll see how that works later."

Since any knowledge before testing of the pattern of reward placement would contaminate the results, a special

effort was made to secure the cooperation of the Ss in not talking about the testing experience. Questioning failed to reveal any breach of this confidence.

The following paragraphs give the administrative procedure and that portion of the instructions which were given verbatim. As the S enters the testing room, "Hello. You're S's name, aren't you?" The S is then seated at a table in front of the H-P Apparatus. "I guess you've come to play the games with me, haven't you? Take a look at this box over here! What do you think it is? What are those black things? Push one in! See how they work, they're doors aren't they? Try them all!" Then, holding up a piece of candy, "What's this? It's candy, isn't it? You and I are going to play a little game. Your job is to figure out what door the candy is behind. Watch, I'll show you how we do it."

The E then steps in front of the apparatus and places the candy behind the appropriate door for trial one of the first problem, Immediate Reaction. "Now, where is the candy?" Throughout the age range tested, the Ss immediately pointed to or opened the correct door with very few mistakes. The S was then told to reach in and take the candy out. "You get to eat the first piece now. All the rest of the candy you find you get to keep and take with you, but you have to save it to eat later." The E then goes on to

Problems II and III. Problem III necessitated a sixty second delay with some interpolated distraction. The testing was done in the school library, and the distracting stimuli were books appropriate to the child's grade level. There were few mistakes on the Delayed Problems, even at the second grade level. Since the testing was done in extremely confining quarters and to save time, Problem III was omitted for grades eight, ten, and twelve.

When Problem III was completed additional instructions were given. "Now we are going to play another game. I am going to put the candy behind the doors according to a certain plan, and your job is to try to figure out what the plan is so that you'll know what door the candy is behind. When I say 'Go' you open the doors and try to find out where the candy is. Open just one door at a time and keep looking until you find the candy. Remember, your job is to try to figure out what the plan is so that you'll know what door the candy is behind." At this point the E retired behind the box and began with Problem IV.

Whenever the S had completed the last trial of a problem, the E paused and came out from behind the apparatus. At this point he stated that the S had done well. Then, "Now we will play another game. This one is like the others we have played, but it is a little bit different."

The E continued on until the S experienced his first

failure. After the final presentation was over the E stepped aside and commented, "That one was difficult, wasn't it?" Unless the child made specific reference to or behaviorally appeared to be disappointed, no further comment was made. If some disappointment was apparent the E stated, "That problem was for older boys and girls. When you get to be older you will be able to solve it." The E then directed attention to all the candy the child had won and thanked the child for playing the game with him.

Specific questions which arose during the course of the test administration were handled in a non-commital manner consistent with the Stanford-Binet or Wechsler instructions: specific questions were deferred until later, the Ss were told they're doing "all right", "motivation" was maintained by "keep up the good work," and so forth. It was somewhat surprising that very few questions were asked.