1974

The development of concrete operational structures as a function of structural equilibration in normal and mentally retarded populations.

Kees J. Vanden Heuvel
University of Windsor

THE DEVELOPMENT OF CONCRETE OPERATIONAL STRUCTURES AS A FUNCTION OF STRUCTURAL EQUILIBRATION IN NORMAL AND MENTALLY RETARDED POPULATIONS

BY

KEES J. VANDEN HEUVEL

A Dissertation
Submitted to the Faculty of Graduate Studies through the
Department of Psychology in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy at the
University of Windsor

Windsor, Ontario

1974

ABSTRACT

The experiment tested the postulate that the cognitive structures of normal and mentally retarded individuals differ in the degree of equilibration (Inhelder, 1968). Three groups of subjects, 28 normal children (N), 28 mentally retarded children and adolescents (YMR), and 28 mentally retarded adults (OMR) were presumed to differ in equilibration. All the subjects were at the preoperational stage of cognitive development as was indicated by their performance on a pretest consisting of eight different tests of concrete operational functioning: a transitivity of substance test, a conservation of number test, five tests of substance conservation, and a conservation of weight test. Half the subjects in each of the age groups were exposed to a training procedure designed to inculcate concrete operational development; the remaining subjects served as controls. It was reasoned that the differences in equilibration should have been evident in the performance of the different groups of subjects in reaction to the training procedures. Transfer from the training task, which was a conservation of substance situation, was measured to concrete operational tests administered immediately after training (Posttest I) and again one to two weeks later (Posttest II). The posttests were verbatim replications of the Pretest.

It was found that the training procedures were effective in increasing the number of correct judgements on the induction task for all three experimental groups, relative to the control groups. There were no differences in the performance of the three experimental groups on the induction task. No transfer of training was evident on the tests of Posttest I, as judged by the fact that the number of correct judge-

ments made by the experimental groups was not significantly different than the number made by the control groups. On Posttest II the subjects of the three experimental groups made significantly more correct judgements than the subjects of the control groups on the conservation of substance tests and on the combined transitivity of substance and the conservation of number tests. There were no significant differences in the number of correct judgements made by the subjects of the three experimental groups on these two blocks of tests. No transfer was evident on the conservation of weight test of Posttest II. The training procedure did not have any significant effect on the number of concrete operational explanations given by the subjects of any of the experimental groups; the number of such explanations offered on the Posttests did not differ from the number offered on the Pretest.

The aforementioned results fail to support Inhelder's postulations concerning differences in equilibration between normal and retarded individuals. The results indicated that the three age groups, who were presumed to differ on equilibration, responded similarly to the training techniques. The implications of the present experiment are constrained by the fact that no changes were evident in the explanation dependent variable.

ACKNOWLEDGEMENTS

I wish to express my appreciation to the members of the Dissertation Committee for their many efforts on my behalf. The members of this committee were: Dr. M. Bunt (Chairman), Dr. A. Kobasigawa, Dr. B. Rourke and Dr. C.J. Brainerd.

For lending their assistance in the acquisition of subjects, I would like to thank Dr. J. McHugh, Mr. J. MacGregor, Mr. G. Soucy and Mr. D.H. Richardson - all at the Rideau Regional Centre, Smiths Falls; Mr. R. MacKenzie, Mr. E. Moran, Mr. R. Sharp, Mr. G. de Shield, Mr. D. Larsen - all at Prince Edward Heights, Picton; Mr. E. Lake of the Leeds & Grenville County R.C. Separate Board of Education; the staff of St. Francis de Sales School, Smiths Falls; the staff of St. John's Separate School, Perth; and Mrs. M. Thompson of Ottawa.

Special thanks goes to the children and adults who served as subjects in the experiment: without their cooperation, the whole project would have been impossible. Without their liveliness, the whole project would have been much less enjoyable.

I would also like to thank a group of people who may or may not have had specific inputs to the dissertation, but who have certainly had more pervasive inputs that led me to undertake it. Among them are my parents, who led me to appreciate the value of an education; my children, who have made the study of children more than an academic exercise; and Dr. C.J. Brainerd, Dr. V. Sermat and Dr. A. Kobasigawa who have collectively given me the idea of what psychology is all about - not through discourse, but through their behaviour.

Finally, I would like to thank Alfreda who served as wielder of the cattle prod that motivated me, typist, proofreader, editor, etc. It was in her role as second scorer that she made the ultimate sacrifice of love that a wife can make ... she agreed with her husband more than 95 percent of the time.

TABLE OF CONTENTS

TABLE OF CONTENTS (CONT'D)

LIST OF TABLES

LIST OF TABLES (CONT'D)

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Inhelder (1968) suggested that the cognitive development of the mentally retarded (MR), as contrasted with that of those with normal intelligence, involves a slower rate of progression through the same sequence of cognitive stages and a more limited upper stage of cognitive development (cf., Zigler, 1969). Inhelder attributed both of these differences to differences in the power of equilibration. The present experiment tested the validity of two deductions based on the aforementioned model: (a) That individuals equated on the level of cognitive development but differing in rate of development, will manifest differences in new structural learning. (b) That individuals presumed to have attained the upper level of their cognitive development will demonstrate no new structural learning. In the remainder of the present section of the paper, special emphasis shall be given to outlining the theoretical propositions of Inhelder's (1968) model, to reviewing the empirical research conducted to date which relates to the model and, to detail the logic involved in the derivation of the hypotheses.

The basic tenet of Inhelder's (1968) model is that a similar construction of logical operations occurs in both normal and MR individuals (Piaget and Inhelder, 1947). Specifically with regard to the development of concrete operational thinking in the MR, three lines of evidence support this commonality. First, when assessed on tasks of concrete operational functioning, many MR's demonstrate by the judgements they make the presence of concrete operational structures (Gruen and Vore, 1972; Lister 1970, 1972; McManis 1969a, 1969b, 1969d). Second, if the explanations proffered by the subject reflect the nature of the structure, as suggested by Brainerd (1972a), then the concrete operational structures of normal and retarded subjects are of a similar nature. Normal and MR subjects, matched on mental age (MA), apparently give the same type of explanations with the same frequency on a variety of concrete operational tests (Gruen and Vore, 1972). The finding that the explanations proffered by the MR, as well as those of normal children, reflect the operation of the reversibilities and quantitative identity suggests that Piaget's (1968, 1970a) mathematical descriptions of the groupings are descriptive of the cognitive structures of some normals and retardates (Gruen and Vore, 1972; Lister 1970, 1972; McManis 1969b). The third line of evidence bears on the succession of stages. The horizontal decalage occurring in the acquisition of the conservation of substance and weight, often reported with those of normal intelligence (e.g., Inhelder, 1968), also occurs in the retardate populations (Stephens, Miller and McLaughlin, 1969; McManis 1969a: Gruen and Vore, 1972: Lister 1970, 1972; Hunter and Lister, 1970). It appears, then, that normal children and the MR show some basic similarities in cognitive development. For both those of normal intelligence (e.g., Piaget 1970a) and for the MR (e.g., Inhelder, 1968) equilibration is cited as being the most prominent factor influencing cognitive development.

Inhelder (1968) differentiated normal and retarded development on the dimensions of (a) rate of cognitive structural development and (b) the level of structural development ultimately acquired. Comparisons of the chronological age (CA) at which normal and retarded subjects acquire competence in a given task support Inhelder's (1968) contention that the

MR's cognitive development occurs at a slower rate than is the case for normals. Designs matching normals and the MR on chronological age, invariably result in the MR lagging behind the normals in performance on Piagetian cognitive tasks (e.g., Gruen and Vore, 1972; Mitchell, Lovell, and Everett, 1962; Brown 1973). However, there is no support for Inhelder's suggestion that while the development of those with normal intelligence is characterized by an increasing power of equilibrium (as indicated by the mobility of successive structures), that of the MR's shows a decreasing power of equilibration. Support for such a suggestion must come from investigations involving repeated measurements over prolonged periods of time and such studies have not been completed.

The cognitive development of normal and retarded individuals differs not only in rate; the development of the MR is also characterized by an arrestation in structural development prior to the acquisition of formal operational reasoning (Inhelder, 1968). Inhelder suggested that structural development can be arrested at any stage of development. She based this conclusion on the observations that none of her retarded subjects gave any evidence of formal operational reasoning and that a large number of these also gave no evidence of concrete operational structures. Research conducted with the severely and profoundly retarded subjects supports Inhelder in this respect. Many severely and profoundly retarded adults are incapable of functioning at the sixth stage of sensorimotor intelligence (Woodward, 1959; Woodward and Stern, 1963). Similarly, large numbers of retarded adults fail to show indices of concrete operational reasoning (e.g., Woodward, 1961). While these data support Inhelder's (1968) suggestions as they relate to arrestation, to posit that further structural development is improbable in the MR after the age of 16, as she did, is a different matter. Inhelder based this suggestion on the observation that not one subject aged 16 years or over (the age by which those of normal intelligence manifest formal operational thought), benefitted from the experimenter's suggestions. However, since Inhelder's (1968) research design did not involve any systematic attempts at inducing learning nor any systematic measure of learning, the suggestion that further structural learning is improbable is mere conjecture.

Differences in rates of structural development between those of normal and retarded intelligence and the possibility of further structural development in retarded subjects over the age of 16 could be evaluated in "learning experiments" in which the subject is exposed to information specifically designed to induce further structural growth (cf., Inhelder and Sinclair 1969; Pinard and Laurendeau 1969). In one such experiment, Brison and Bereiter (1964) found no differences in the induction of conservation of substance responding among subjects of superior, normal and retarded intelligence. However, the Brison and Bereiter experiment was an inconclusive test of Inhelder's suggestion that normals and retardates are characterized by different rates of structural development in that the experiment failed in include a number of design aspects essential to an adequate training study (e.g., Inhelder and Sinclair, 1969). For example, Brison and Bereiter (1967) failed to (a) differentiate transitional and non-conserving subjects at the beginning of the experiment, (b) test for the persistence of the acquisition at a time after training, and (c) include tasks of concrete operational functioning other than substance conservation (which was the skill specifically trained). As yet, no studies have been reported which have attempted to induce further structural development in MR adults; hence Inhelder's postulate concerning the improbability of cognitive development in those over the age of 16 remains untested.

In examining the difference in the cognitive development of normals and the mentally retarded, Inhelder (1968) used equilibration as the fundamental explanatory construct. Differences in the power of equilibration were used to explain both the slower rate of development and the arrestation of development prior to the attainment of formal operational reasoning. In employing equilibration as the central explanatory construct, Inhelder was not necessarily negating the importance of other factors such as the genetic determination of intelligence, the importance of the individual's psychosocial history, etc. frequently cited by others attempting to explain the mental retardation phenomena (e.g., Zigler, 1969). In the Piagetian scheme, these factors are considered subordinate to equilibration (e.g., Piaget, 1973, chapter 1); equilibration is the supraordinate factor which governs the interaction of all these factors in their effects on cognitive structural development (Piaget 1960a). The role of equilibration in the development of cognitive structures is founded on probabilistic considerations in part determined by the gains and losses associated with each strategy (Piaget 1960a, 1967, 1970a). With respect to the difference in development in those of normal intelligence and the MR, given sequential strategies 1, 2......n which are essential for the development of a structure, the probability for the development of each successive strategy is less for the MR than for those of normal intelligence.

The arrestation purported to occur in the development of the MR was seen as being the outcome of a progressive decrease in the power of equilibration (Inhelder, 1968). Maturation and/or other factors closely associated with temporal events were implicated by Inhelder (1968) in a role subordinate to equilibration. She suggested that the arrestation in cognitive development may occur at approximately 14 to 16 years of age, since this is the age at which many of the changes associated with puberty are completed and this is also the age at which normal children have acquired formal operational structures. According to Inhelder (1968) arrestations in development result in intellectual structures characterized by a permanent state of disequilibrium or false equilibrium. Such an intellectual structure, while not in equilibrium with the realities of the situation, is incapable of assimilating the information or of accommodating to the information. Thus further structural development is improbable, if not impossible.

In summary, Inhelder's model of cognitive development of the mentally retarded contains the following propositions: (a) Retarded and non-retarded development are similar with respect to the process of structural development and the cognitive structures developed. Equilibration is proposed as being the fundamental factor in the structural development of both populations. (b) The development of cognitive structures occurs at a slower rate in retarded than normal individuals. The difference in the rate of development is attributed to differences in the power of equilibration of the structures in the retarded and non-retarded individuals. (c) The rate of development, as indicated by the mobility or power of equilibration of successive structures, progressively decreases in the MR and progressively increases in those of normal intelligence.

(d) The progressive decrease in the power of equilibration eventually results in arrestation of structural development of the retarded persons at about 14 to 16 years of age. Non-retarded individuals, however, proceed to develop until they have acquired formal operational thinking, which is characterized by true equilibration. As was evident in the literature cited, the empirical support for points (a) and (b) is quite extensive, while points (c) and (d) remain mere conjecture.

Derivation and Statement of Hypotheses

The present experiment investigated Inhelder's (1968) proposition that the differences in rate of structural development in those of normal and retarded intelligence result from differences in equilibration. Since the power of equilibration cannot be experimentally manipulated, it was necessary to select subjects presumed to differ on this dimension. This was accomplished by selecting subjects at the same stage of cognitive development (preoperational) but who were of different chronological ages. Since the subjects were matched on the stage of cognitive development but differed grossly in chronological age then these subjects in their past history likely had grossly different rates of structural development. Thus when equated on cognitive stage, chronological age should have related inversely to rate of structural development. The three samples of subjects were (a) normal children (N) aged 4 to 6 years, (b) young mentally retarded children and adolescents (YMR) aged 9 to 16 years and, (c) mentally retarded adults (OMR) aged 17 to 30 years. An ordinal scale in the rate of structural development such that N > YMR > OMR was postulated since the three groups varied on chronological age, and chronological age is at least scalable on an ordinal dimension.

According to Inhelder (1968), if the rate of structural development can be ordinated, then the power of equilibrium can also be ordinated such that N > YMR > OMR. Thus, the differences on the equilibration dimension were deduced from the different rates of structural learning demonstrated by the subjects prior to the experiment.

In the present experiment, the subjects were exposed to a training procedure designed to lead to correct judgements in the conservation of substance situation. Tests of concrete operational reasoning were administered immediately after the training session and again one to two weeks later. There were three groups of operational tests which differed in the way they related to the conservation of substance task on which the subjects were trained: (a) Five tests which measured the same conceptual skill with regard to the same object property (conservation of substance). (b) A test measuring a different conceptual skill with respect to the same object property (transitivity of substance) and a test measuring the same conceptual skill with respect to a different object property (conservation of number). In the course of normal development both of these skills are acquired prior to the conservation of substance. (c) A test measuring the same conceptual skill with respect to a different object property (conservation of weight). Conservation of weight is invariably acquired after the conservation of substance (Piaget, 1973). The hypotheses for the present experiment were deduced from Inhelder's (1968) model; specifically from the suggestion that the three age groups can be ordinated on the power of equilibration dimension so that N > YMR > OMR.

The present experiment was a test of Inhelder's model in that first, it avoided the circularity of Inhelder's thesis that differences in equilibration explain the different rates of structural development. Inhelder deduced from the slower rates of structural learning of her retarded subjects that they differed in equilibration. She then used the construct of power of equilibration to explain the observed differences in rates of development. In the present experiment, the differences in the power of equilibration were employed to predict differences in new structural learning. Second, the present experiment incorporated repeated measurements, so that if development did occur it would have been measured. Inhelder constructed the model on the basis of seeing each individual subject on a single occasion only, hence restricting the possibility of actually observing development. Third, the design of the present experiment incorporated measures of equilibrium of cognitive structures; hence group differences in equilibration could be assessed. The importance of equilibration in structural development was deduced by Inhelder from Piagetian theory and no data was systematically collected to support this contention. The hypotheses in the present experiment, deduced from Inhelder's model, are stated in the following paragraphs.

Hypothesis 1. The number of correct judgements made by the experimental groups during the induction phase should be capable of ordination so that N > YMR > OMR; however, the three control groups should not differ significantly on the number of correct judgements given during the induction phase. The differences in equilibration among the three age groups should lead to differences in structural learning, as measured by correct judgements. As the structural learning should be a function of the exposure to the induction procedures, the second order interaction of the age group x experimental treatment factors should be significant as is indicated in the hypothesis.

Hypothesis 2. The number of correct judgements and the number of appropriate explanations made by the subjects of the three experimental groups should be capable of ordination so that N>YMR>OMR on each block of tests on each of the posttests; however, there should not be any significant differences between the three control groups on the same measures. This hypothesis is also predicated on presumed differences in equilibration existent among the age groups at the beginning of the experiment. The second order interaction is predicted since the increase in performance indices should be due to the experimental induction techniques.

Hypothesis 3. If hypothesis 2 is confirmed, then it should also be found that the degree of transfer from the induction task, as measured by the number of correct judgements and appropriate explanations, should vary as a function of the block of tests. That is, the triple interaction of age group x experimental treatment x block of tests should be significant. This hypothesis is predicated on hypothesis 2 plus the postulate that the degree of transfer should be a function of relation of the block of tests to the induction task. The greatest amount of transfer should occur to the test used during the induction phase - the conservation of substance tests. The degree of transfer to the conservation of number and transitivity of substance tests should be less than that evident on the substance tests; and the transfer to the conservation of weight test should be less than evident on either the conservation of substance tests or the transitivity of substance or the conservation of number test. This latter prediction stems from the review of the literature where it was generally found that these two tests were acquired prior to the conservation of substance. The least amount of transfer should occur to the conservation of weight tests since, as was indicated in the introduction, the conservation of weight is almost invariably acquired after the conservation of substance.

Hypothesis 4. Within experimental group comparisons of the number of correct judgements and appropriate explanations on Posttest I and Posttest II should indicate that the N experimental subjects will show the least amount of loss from Posttest I to Posttest II; the amount of loss from Posttest I to Posttest II on the part of the YMR experimental subjects should be between that evidenced by the N and OMR subjects; and the OMR experimental subjects should show the greatest amount of loss. This hypothesis is also based on the presumed differences in equilibration existent between the different - aged subjects. One measure of equilibration is stability (Flavell, 1963); in the present experiment, stability is assessed by the retention of structural learning over the period of time intervening between Posttest I and Posttest II.

CHAPTER II

METHOD

General Design

There were five basic steps in the experiment. First, each subject was evaluated to ascertain comprehension of the relational terms "same, more, and less". Subjects failing to demonstrate comprehension of these terms were eliminated from participation in the rest of the experiment; hence those subjects whose non-conservation responding might have been due to an inability to comprehend the questions were precluded. (Griffiths, Shantz and Sigel, 1967). Second, the remaining subjects were pretested for concrete operational thinking via a test of transitivity, five tests of substance conservation, a test of number conservation and a test of weight conservation. Each conservation task, but not the transitivity task, consisted of two phases: (a) a prediction phase in which the subject was required to make conservation judgements prior to the perceptual deformation of one of the stimulus configurations and, (b) a transformation phase in which the subject was required to make conservation judgements after the perceptual deformation of one of the stimulus configurations. This procedure was designed to allow the differentiation of three stages in the acquisition of conservation: Stage I = total non-conservation, Stage II = conservation in the prediction phase only, and Stage III = conservation in both phases (Piaget, 1952; Brainerd, 1972b; Brainerd and Brainerd, 1972). Only those subjects performing at Stage I on all the conservation tasks and who did not make consistent transitive judgements on the transitivity task participated in the remainder of the experiment. Third, the remaining subjects were trained to acquire Stage III conservation responding on a liquid conservation task. The training occurred one to three days after administration of the pretest. Fourth, immediately after the completion of the training, each subject was posttested. Fifth, each subject was again posttested one to two weeks later.

The experiment is of a 3 (age) x 2 (experimental treatment) x 3 (test days) x 3 (groups of tests) factorial design. The first two factors were between subjects and the last two were within subject factors. Three age spans defined the levels of the age factor: 4 to 6 years, 9 to 16 years, and 17 to 30 years. The two treatments constituted the control and the experimental groups. During the training phase, the experimental subjects received training on Stage III conservation responding by receiving contingent reinforcement for their judgements (Brainerd 1972b, 1972c; Overbeck and Schwartz, 1970) and by acting out the object bound form of inversion reversibility (Brainerd and Allen, 1971). The control subjects were treated in an identical manner except that they were not exposed to either of the induction procedures. The three levels of the test days factor were the pretest and the two posttests administered to each subject. The three levels of the groups of tests factor were the conservation of substance tests, the transitivity of substance and the conservation of number tests, and the conservation of weight test.

Subjects

The relational terms pretest was administered to all students comprising the kindergarten classes of two elementary schools; one of the schools was located in a town with a population of 10,000 and the other in a town with a population of 5,000. Another 10 children were pretested from the subsequent years' kindergarten class in the first mentioned town; in this instance the children in the kindergarten class were pretested until the experimental and control groups of the N group were fully constituted. In all three of these kindergarten classes, approximately half of the children resided within the town limits; the remainder were bussed in from the rural environs. All children in the normal groups attended schools half days only.

Subjects in the younger mentally retarded (YMR) and the older mentally retarded (OMR) groups were drawn from two provincially administrated facilities for the mentally retarded. Both of the facilities are located in Eastern Ontario; one has a resident population of 1500 and the other of 400. All subjects in the YMR group attended academic school in their respective facilities, some on a full day basis and others on a half day basis. Approximately half the subjects pretested for the OMR group also attended school; the remainder were situated in job settings in their respective facilities.

All possible subjects for the OMR and YMR groups were initially discussed with their teachers and/or counsellors. The subject was not pretested if, in the opinion of the teacher or the counsellor, (a) the subject was lacking sufficient intellectual ability to comprehend any of the relational terms, (b) the subject's speech was incomprehensible or, (c) the subject was afflicted with gross sensory handicaps which would interfere with the administration of the tests. Inspection of

INSERT TABLE I ABOUT HERE

Table 1, which lists the number of subjects participating in each step

TABLE 1

Distribution, Age and IQ of Subjects

Level of Pretest Performance	Distribution of Subjects	Mean Age (years)[c]	IQ (median)[d]	IQ (range)[d]
		Normals		
No Quantification	7	–	–	–
Stage I	32a			
Experimental		–	–	–
Control		–	–	–
Stage II or III	43(16)[b]	–	–	–
		Young Mentally Retarded		
No Quantification	55	13.4	35[e]	22-70[e]
Stage I	29[a]			
Experimental		14.3	45	33-77
Control		14.4	49.5	31-72
Stage II or III	17(15)[b]	14.5	62	31-93
		Old Mentally Retarded		
No Quantification	37	20.8	33.5[f]	20-54[f]
Stage I	32[a]			
Experimental		21.1	54.5	32-79
Control		22.7	46.5	32-70
Stage II or III	45(29)[b]	19.2	59	36-82

a. Number in excess of 28 lost in the course of the experiment due to attrition.

b. Within parentheses is the number of subjects administered the complete pretest.

c. All normal subjects were 4.8 to 5.8 years.

d. IQ data on normal subjects not available. IQ's for the MR obtained from clinical files was based on one of four tests: Wechsler Adult Intelligence Test for Children, Stanford-Binet (L-M) or the Peabody Picture Vocabulary Test.

e. IQ data for three subjects not available.

f. IQ data for eight subjects not available.

of the experiment, indicates that the counsellors and the teachers recommended many who did not comprehend a number of the relational terms, thus suggesting that few subjects who comprehended the relational terms were excluded from the study. The age breakdowns for each of the groups is also indicated in Table 1. Of the 14 subjects constituting each age x treatment group, the number of males were: N experimental, 7; N control, 7; YMR experimental, 8; YMR control, 6; OMR experimental 9; and OMR control, 9.

Materials

In addition to the materials used in each test, which are described in the test's description, a Sony TC252 stereo tape recorder and audio tapes were used to record all of the subject's verbalizations during the pretest and posttests.

Procedure

The Relational Terms Pretest. This test served the purpose of (a) ensuring that non-conservation responding by the subjects in the later sections of the experiment was due to an inability to conserve (Griffiths, Shantz and Sigel, 1967) and (b) that all the subjects taking the tests of concrete operational reasoning were at least at the stage of gross quantification (Piaget, 1952). Materials. Two identical glasses (14 x 7 cm.) and a pitcher of coloured water were used. Procedure. The subject made three comparisons; in each comparison the two glasses, filled with varying amounts of coloured water, were placed on the table in front of the subject. In one comparison both of the vessels were half filled with water, in another one of the vessels was half filled while the other was three quarters filled, and in another one of the vessels was half filled while the other was one quarter filled. For each of the comparisons the following randomly ordered questions were asked: (a) Does one of the glasses have more water in it? Which one (if the answer was yes)? (b) Do the glasses have the same amount of water in them? (c) Does one of the glasses have less water in it? Which one (if the answer was yes)? Since all of the questions were asked for each of the comparisons, a total of nine questions was asked. In order to continue in the experiment, the subject was required to answer at least seven of the nine questions correctly.

Concrete Operational Reasoning Pretests

The pretest was comprised of eight tests of concrete operational thinking: conservation of number, four tests of conservation of liquid substance, conservation of solid substance, conservation of weight, and transitivity of substance. The procedure for each of the tests was patterned after the conservation of number tests employed by Brainerd and Brainerd (1972).

1. Conservation of number. Materials. A bowl containing 25 white and 25 red plastic poker chips (each 3.8 cm. in diameter) was used in the experiment. Procedure. Two parallel rows of seven chips were constructed, one row being comprised of white chips and the other of red chips; the two chips in each column were in one to one correspondence. The chips within a row were placed approximately three to four cm. apart from one another; the two corresponding chips in each column actually touched each other. The subject was first asked what the objects were called. In all further references to the chips, the experimenter used the name designated by the child. The subject was then asked if there

were the same number of chips in the two rows (or lines). Once the subject agreed that the two rows contained an equal number of chips, a two staged sequence of questions was asked: viz., a prediction phase and a transformation phase. Prediction. Leaving the chips intact, the experimenter asked the following randomly ordered questions: (a) If I were to push the chips in this row (designating one of the rows) very close together, would the two rows have the same number of chips in them? How do you know? (b) If I were to push the chips in this row (indicating the same row) very close together, would one of the rows have more chips in it? (If so, which one?) How do you know? (c) If I were to push the chips in this row (indicating) very close together, would one of the rows have less chips in it? (If so, which one?) How do you know? It was often found necessary to push the chips in the centre of the row together in order to demonstrate what was meant by "If I were to push the chips in this row very close together...". Transformation. The experimenter or the subject then pushed together the chips in the designated row, until the chips touched each other. The following, randomly ordered, questions were asked: (a) Do the two rows have the same number of chips in them? How do you know? (b) Does one of the rows have more chips in it? (If so, which one?) How do you know? (c) Does one of the rows have less chips in it? (If so, which one?) How do you know? After the subject gave the last explanation, the chips were returned to the bowl.

2. Conservation of liquid substance. Materials. Two "standard" glasses (14.0 x 7.0 cm.), one "thin" glass (17.0 x 6.0 cm.), one "wide" glass (8.0 x 8.5 cm.), one pyrex "pie plate" (4.5 x 25.0 cm.), one 100 ml.

"graduated cylinder" (25.0 x 3.0 cm.), and a pitcher with blue water were used in the liquid substance tests. Procedure. The two standard glasses, filled with equal amounts of water, were placed side by side on the table. When the water was to be poured into the wide or the tall glass, or into the pie plate, the two standard glasses contained approximately 200 ml. of water. When the water was to be poured into the graduated cylinder the two standard glasses contained approximately 100 ml. of water. After the subject agreed that the two vessels contained the same amount of water, the following two step assessment was executed. Prediction. The wide glass (or the tall glass, or the pie plate, or the graduated cylinder) was placed between the two standard glasses and the subject was asked the following, randomly ordered questions: (a) If I pour the water from this glass (randomly indicating one of the standard glasses) into this glass (indicating the glass set between the two standard glasses), would these two glasses (indicating the other standard glass and the glass into which the water is to be poured) have the same amount of water in them? How do you know? (b) If I pour the water from this glass (indicating), would one of these two glasses (indicating) have more water in it? (If so, which one?) How do you know? (c) If I pour the water from this glass (indicating) into this glass (indicating) would one of these two glasses (indicating) have less water in it? (If so, which one?) How do you know? Transformation. The experimenter or the subject then poured the water as indicated in the preceeding paragraph. After the pouring, the following randomly ordered questions were put to the subject: (a) Do these two glasses (indicating) have the same amount of water in them? How do you know? (b) Does one of these two glasses (indicating) have more water in it? (If so, which one?) How do you know? (c) Does one of these two glasses (indicating) have less water in it? (If so, which one?) How do you know? After completion of the questions, the water was returned to the pitcher. Each of the tests (i.e., the one with the wide glass, the tall glass, the pie plate, and the graduated cylinder) was considered to be separate from each other and given in random order among the other tests of concrete operational thinking.

3. The conservation of solid substance. Materials. Two identical balls of play-doh, approximately 5.0 cm. in diameter, were used in this test. Procedure. The two balls of play-doh were placed on the table and the subject was asked if the two had the same amount of play-doh in them. After the subject agreed to their equality, the following two step assessment was carried out. Prediction. The subject was asked the following randomly ordered questions: (a) If I were to roll this ball (indicating one of the two balls) into a "sausage" ("hot dog" or "wiener") would the two pieces have the same amount of play-doh in them? How do you know? (b) If I were to roll this ball (indicating) into a sausage, would one of the pieces have more play-doh in it? (If so, which one?) How do you know? (c) If I were to roll this ball (indicating) into a sausage would one of the pieces have less play-doh in it? (If so, which one?) How do you know? Transformation. The experimenter or the subject then rolled the designated ball into a sausage shape and the subject was asked the rollowing randomly ordered questions: (a) Do the two pieces have the same amount of play-doh in them? How do you know? (b) Does one of the pieces have more play-doh in it? (If so, which one?) How do you know? (c) Does one of the pieces have less play-doh in it? (If so, which one?) How do you know?

4. Conservation of weight. Materials. Two identical balls of

plasticine, approximately 4 cm. in diameter, and a Rexo-s* Photographic balance scale (Pelouze Manufacturing, Evanston, Ill.) were used in this test. Procedure. After placing the scales in the centre of the table, the experimenter asked several questions to ascertain if the subject knew how the scales worked. If it was found that the subject was unfamiliar with the balance scales, the principles were demonstrated using a variety of objects. The subject was then asked to place the pieces of plasticine on the balance scale to determine if they weighed the same. Once the subject agreed that the two pieces were of the same weight, the following two step assessment was carried out. Prediction. (a) If I were to flatten this ball of plasticine into a "pancake" (or "cookie"), would the two pieces weigh the same? How do you know? (b) If I were to flatten this ball of plasticine (indicating) into a pancake, would one of the pieces weigh more? (If so, which one?) How do you know? (c) If I were to flatten this ball of plasticine into a pancake, would one of the pieces weigh less? (If so, which one?) How do you know? Transformation. The designated ball was then flattened into a pancake shape and the subject was asked the following randomly ordered questions: (a) Do these two pieces of plasticine weigh the same? How do you know? (b) Does one of the pieces of plasticine weigh more? (If so, which one?) How do you know? (c) Does one of the pieces of plasticine weigh less? (If so, which one?) How do you know?

5. Transitivity of substance. Materials. Three identical glasses (10.0 x 5.5 cm.) were used in the test: glass A contained 100 ml. of yellow water, glass B contained 100 ml. of red water, and glass C contained 90 ml. of yellow water. The difference in water levels between

A and C was perceptible when they were placed next to each other, but not when they were any distance apart. Procedure. The glasses were compared in pairs (A and B, B and C); the order of the pair first compared was randomized. Vessels A and C were never perceptually compared to one another and they were always kept at least 1 metre apart. After making the comparisons, the experimenter held vessels A and C at arms length and, while jiggling both of the glasses, asked the following randomly ordered questions: (a) Do these two glasses have the same amount of water in them? How do you know? (b) Does one of the glasses have more water in it? (If so, which one?) How do you know? (c) Does one of the glasses have less water in it? (If so, which one?) How do you know?

Induction

Only those subjects performing at Stage I on all of the concrete operational tasks that comprised the pretest were included in the induction phase of the experiment. The two induction procedures employed met the dual criteria of being deducible from Piagetian theory and of having been successfully employed in previous research.

The object bound form of inversion reversibility involves having the subject return the transformed object to its original state by using the inverse of the action originally used to transform the object (Brainerd and Allen, 1971). Brainerd and Allen (1971) suggest that this procedure may provide the generalized experience required to promote advances in the cognitive operation of inversion reversibility that Piaget (1970c) maintains are the root of all operations and cognitive structures. The object bound form of inversion reversibility is not, by itself, a sufficient condition for the occurrence of operational inversion reversibility. The subject can reason, for example, that while one of the objects is in its transformed state, the two objects are quantitatively different, even though they will again be the same when the two are returned to perceptually identical states (Piaget, 1952b).

In contingent reinforcement, the second induction procedure used, the subject is told whether his conservation judgements are right or wrong. Telling the subject that he is wrong could lead to the cognitive disequilibrium that Piaget (1960a, 1960b) considers essential for further cognitive growth (cf., Langer, 1969). The power of contingent reinforcement as a technique for inducing further structural development is all the more convincing in that on subsequent posttests it has been found to result in superior explanations, even though only the conservation judgements were subjected to contingent reinforcement (Brainerd, 1972b).

The subjects at each age level were randomly divided into two groups: an experimental group and a control group. The training procedure was the same for both groups except that the control subjects were not exposed to the object bound form of inversion reversibility or contingent reinforcement. Materials. Two "standard" glasses (8.0 x 7.0 cm.) and a "tall" glass (10.0 x 5.5 cm.) — all of an opaque yellow colour —, and a pitcher of green water were used during the induction phase.

Procedure for the experimental group. After the subject agreed that the two "standard" glasses contained the same amount of water, the taller glass was introduced between the two "standard" glasses. The subject was then asked to pour the contents of one of the "standard" glasses into the taller glass. The following randomly ordered questions were then asked: (a) Do these two glasses (indicating the vessels with water in them) have the same amount of water in them? (b) Does one of the glasses (indicating) have more water in it? (c) Does one of the glasses (indicating) have less water in it? After each "yes-no" answer, the experimenter said "That's right" or "That's wrong", according to whether the judgement was correct or incorrect. After answering the questions, the subject poured the water from the taller glass back into the "standard" glass. The subject was then asked: Do these two glasses have the same amount of water in them? This procedure was repeated up to 12 trials or until the subject made three correct judgements on three consecutive trials.

Procedure for the control group. The subject was asked to agree that the two "standard" glasses contained the same amount of water before he poured the water from one of the "standard" glasses into the taller one. The following randomly ordered questions were asked: (a) Do these two glasses (indicating) have the same amount of water in them? (b) Does one of the glasses (indicating) have more water in it? (c) Does one of the glasses (indicating) have less water in it? After answering the questions, the subject poured the contents of the taller glass into the pitcher. The two standard glasses were again filled with equal amounts of water and the procedure was repeated. The number of trials was determined by the number of trials it took the matched experimental subject to reach criterion.

Immediate and Delayed Posttests

Immediately after the training trials were completed, and again one to two weeks later, both the experimental and control subjects participated in the posttests. Both of the posttests were verbatim replications of the pretest.

Randomization

Except in those instances that other activities interfered, the subjects were tested according to the alphabetical listing on the class or ward list. Assignment to the experimental or control groups was determined by the order in which the subjects were tested: in each age group, the first Stage I male and female subject were assigned to the experimental group, subsequent subjects were alternately assigned to the control and experimental groups. The matching of subjects in the control and experimental groups was done by matching the experimental subject with the subject of the same sex closest to the subject in chronological age.

On each occasion the subjects were tested, the eight tests of concrete operational functioning were administered in a random order as determined by a table of random numbers. On all of the conservation tests the prediction phase always preceeded the transformation phase but the relational questions asked within each phase, as well as the relational questions asked during the induction phase of the experiment, were asked in a random order.

Dependent Measures

The tests of concrete operational functioning yielded two sorts of responses: (a) the judgements to the "same-more-less" questions and (b) the explanations given in response to the "How do you know?" question. Either judgements or explanations are sufficient for the determination of the presence of operational structures, although the explanations provide a more conservative estimate in that they are more susceptible to Type 2 error (Brainerd, 1972a).

The judgements were employed in two ways, depending on the specific question which motivated the analysis. (a) The correct judgements were assigned "1s" and the incorrect judgements were assigned "0s", (b) On the conservation tests, subjects making two or less correct judgements in both the prediction and the transformation phases were assigned to Stage I; subjects making three correct judgements on either the prediction or transformation phase, but not both, were assigned to Stage II; subjects making three correct judgements on the prediction and the transformation phases were assigned to Stage III. As the transitivity of substance test did not involve a prediction phase, subjects answering two or less questions correctly were assigned to Stage I and subjects answering the three questions correctly were assigned to Stage III.

It should be noted that Stage II performance is defined somewhat differently than is the case in the published literature (e.g., Piaget, 1952; Brainerd, 1972b). The analysis of the pretest protocols of subjects showing a discrepant performance on the prediction and transformation phases (i.e., the subject answered the three judgement questions correctly on either the prediction or transformation phase, but not both) showed that there were 56 such instances. In 30 cases the subject answered the three questions correctly on the prediction phase but in 26, the subject correctly answered all the transformation questions correctly. Of the 15 discrepancies evident in the protocols of the N subjects, 10 were in the predicted direction in that the prediction phase performance was superior to that of the transformation phase. Of the 16 discrepancies in the protocols of the YMR subjects, 6 were in the predicted direction. And of the 25 discrepancies in the protocols of the OMR subjects, 14 were in the predicted direction. Consequently, for the present sample there seemed to be little basis for assuming that correct answers on the prediction phase invariantly precede correct answers on the transformation phase.

The scoring of the explanations involved the following steps:

1. The explanations were transcribed verbatim from the audio tapes; any observations recorded during the experimental sessions, which clarified the subject's explanation were appended to the transcribed explanation.

2. The explanations of 30 MR and 30 normal subjects were inductively searched by two persons to determine the number of explanation categories used by the subjects. The explanation categories are named and defined in Table 2.

INSERT TABLE 2 ABOUT HERE

3. The explanation protocols of 60 subjects administered the total pretest were independently scored by two persons to determine the inter-scorer reliability. The 60 subjects were drawn randomly, 10 each from the Stage I and Stage III subjects for each age group. The two scorers agreed on the classification of 95.4 percent of the responses.

4. The extent that each explanation category was associated with correct judgements was calculated using the pretest protocols of the 84 subjects who participated in the remaining phases of the experiment and the 60 subjects who demonstrated Stage II or III responding on one of the operational tests. Table 3 lists the frequency of occurrence of

INSERT TABLE 3 ABOUT HERE

each explanation category on the transitivity of substance test and the combined conservation tests, as well as the percentage of time each category was associated with a correct judgement. An explanation category

Table 2

Definitions of Scoring Categories for Explanations

1. Comparison. On the transitivity of substance test the explanation included at least one of the comparisons with the comparison glass or else included the statement that the comparison was made.

2. Inversion Reversibility. (a) The transformation could be reversed or else the "standard" stimulus left intact could be similarly transformed. (b) The two "standard" stimuli were the same before the transformation. The reference to the two standard stimuli needed to be explicit from the subjects' verbal statement, the observations made during testing, or from the context.

3. Inversion Reversibility Negated. (a) If the transformation was reversed the two "standard" stimuli would be equivalent, but they were not quantitatively equivalent when one of the stimuli was in its "transformed" state. (b) If the transformation was reversed then the two "standards" would not be quantitatively equivalent. (c) The two "standard" glasses were not quantitatively equivalent before the transformation.

4. Logical Necessity. (a) On the conservation tests, statements to the effect that the two stimuli were the same quantitatively, except in response to the "same" judgement question in which case the response was scored as a "tautology". (b) On the transitivity of substance test, statements to the effect that the two stimuli were not the same, except on the "same" judgement question in which case the response was scored as a "tautology".

TABLE 2 (CONT'D)

5. Logical Necessity Negated. In response to the "more" and "less" judgement questions, the subject's response was merely that they were not the same.

6. Identity. (a) Some matter would have to be added or subtracted to result in a quantitative change. (b) The quantitative features would always be the same, no matter how they were transformed. (c) The physical transformation was irrelevant to the quantitative characteristic. (d) The perceptual features when one stimulus was transformed, were irrelevant.

7. Second Invariant. Reference is made to the invariance of another quantitative object property.

8. Second Invariant Negated. The object property in question has changed because of a change in another quantitative property of the object.

9. Perception. (a) Reference is made to only one of the relevant dimensions. For example, in the conservation of number, the relevant dimensions were length and density. (b) A perceptual comparison was made and used as the basis of the explanation. For example, on the conservation of weight, the subject picked up the two objects to weigh them in his hands. (c) Reference is made to features of the object which were not relevant (e.g., colour, shape, marks on the stimuli, etc.). (d) Explanation is based on perceptual changes resulting from the physical transformation of the stimulus. (e) A vague relational term without a specified referent was used (e.g., bigger, smaller, fuller, emptier). (f) The relational word from another question is used as the key word in the explanation. (g) One stimulus has changed but one remained as it was before.

Table 2 (Cont'd)

10. No Perceptual Deformation. That the physical transformation will not result in a change in the perceptual configuration or that the transformation did not result in a change in the perceptual configuration.

11. Tautology. (a) The relational word from the question is merely repeated in the explanation or else is merely negated. (b) A strict synonym for the relational term in the question is used in the explanation.

12. No Logical Information. (a) The subject's explanation made no sense. (b) The explanation question was not answered, responded to by "I don't know", etc. (c) Magical explanations. (d) Explanations suggesting the experimenter secretively took some action to change the quantitative property. (e) Subject stated he would have to make an empirical comparison. (f) Subject reported that he learned it from someone else, etc. (g) Reference term and real object were confused. (h) Subject reported he knew by looking at the stimuli, etc. (i) Subject reported counting the objects (but there was no indication he did so), or grossly incorrect counting.

13. Counting. In number conservation: (a) Subject reports a specific number of elements (7 $\pm$ 2). (b) Subject counts correctly in response to question (7 $\pm$ 2). (c) Subject reports he counted the elements and there is an indication he did so correctly (7 $\pm$ 2).

14. Transformation. (a) The transformation to be completed or already completed is cited. (b) Only one of the stimuli was physically transformed. (c) The physical transformation with the change in the perceptual configuration is cited.

15. Reciprocity Reversibility. (a) The two relevant dimensions

Table 2 (Cont'd)

were mentioned in relation to one of the stimuli or to different stimuli. (b) Mention of the two relevant dimensions with the idea of compensation or the idea of compensation alone. (c) An explanation of the perceptual distortion is given by use of 15 (a) or 15 (b).

16. <u>One to One Correspondence.</u> In number conservation, reference to the fact that the chips in the two rows were in one to one correspondence.

17. <u>One to One Correspondence Negated.</u> The explanation cited the fact that the one to one correspondence was destroyed and consequently the two sets were different quantitatively.

TABLE 3

Association of Explanation Categories with Correct Judgements on Pretest

Explanation Category	Transitivity of Substance		Conservation Tests	
	Frequency of Occurrence	Percent associated with correct judgement	Frequency of Occurrence	Percent associated with correct judgement
Comparison	54	89	-	-
Logical Necessity	8	100	85	84
Logical Necessity Negated	16	23	24	13
Inversion Reversibility	-	-	316	92
Inversion Reversibility Negated	1	100	14	7
Identity	4	0	161	94
Second Invariant	-	-	8	75
Second Invariant Negated	-	-	10	10
Perception	193	55	3035	10
Tautology	39	33	299	23
No Information	115	37	1143	28
Counting	-	-	117	70
Transformation	-	-	535	15
No Perceptual Deformation	1	0	49	78
Reciprocity	-	-	81	33
One to One Correspondence	-	-	14	64
One to One Correspondence Negated	-	-	13	10

Note. - Includes only those explanations scorable in one category.

was defined as "appropriate" if it was associated with correct responding 70 percent or more of its occurrence, unless its occurrence was very infrequent (less than 5). It can be seen that two explanation categories were "appropriate" on the transitivity of substance test: viz., the "comparison" and the "logical necessity" categories. Similarly, six explanation categories were associated with correct judgements viz., (a) "inversion reversibility", (b) "logical necessity", (c) "identity", (d) "second quantitative invariant", (e) "counting" and, (f) "no perceptual deformation". All but the last category were deemed as "appropriate" explanation categories; the "no perceptual deformation" category, while associated with correct judgements 78 percent of its occurrence, is based on the false premise that the transformation would not lead to a perceptual change. Consequently, the "no perceptual deformation" category could not reflect operational thought. (For more detailed data on the relationship of each category with the judgements, see appendix B). Subjects also gave explanations which reflected more than one of the explanation categories. Explanations which were combinations of two "appropriate" categories were associated with a correct judgement 90 percent of the time on the pretest; consequently if an explanation was scored as reflecting two categories, and if two were "appropriate", the whole explanation was deemed appropriate. Combinations of one appropriate and one "inappropriate" explanation were deemed "inappropriate"; on the pretest, these were associated with correct responding on 64 percent of their occurrences. Combinations of two inappropriate responses were also deemed inappropriate; on the pretest these were associated with correct judgements only 12 percent of their occurrence.

5. Explanations falling into the appropriate categories on each test were assigned a score of "1" while those falling into the inappropriate categories were assigned a score of "0".

6. Posttest explanations of all subjects were scored by a single person.

Dependent measures.

The concrete operational tests included in the pretest and the two posttests yielded two sorts of responses either of which is sufficient for the determination of concrete operational structures: viz., (a) the judgements to the "same-more-less" questions and, (b) the explanations given in response to the "How do you know?" questions (Brainerd, 1972a). Consequently both responses were employed as dependent measures.

The judgements and explanations obtained on the five conservation of substance tests assessed the same structure as was trained during the induction phase of the experiment; consequently, these tests yielded measures used to assess the learning of the conservation of substance structure. The remaining three tests on the pretest differed from the conservation of substance tests in that they tapped either the same quantitative aspect (substance) but a different conceptual skill (transitivity) or else they tapped the same conceptual skill (conservation) but with reference to another quantitative feature (number and weight). As such, they provided the dependent measures of the field of application measure of equilibrium. The field of application is "the ensemble of objects or object properties which the equilibrated action system accomodates to and assimilates"(Flavell, 1963, p. 242). However, as different object properties are not assimilated to the operational system with equal facility (Piaget, 1971; Hooper, Goldman, Storck and Burke, 1971), it would be expected that differences should occur in the degree of transfer observed from the conservation of substance tasks to the other three tasks. The normative data reported in the literature, as well as the theoretical writings of the Piagetians (e.g., Piaget, 1970a) provided a basis on which to ordinate the three concept tests relative to the conservation of substance tests.

For a given object property, transitivity is a logically necessary skill for the occurrence of conservation (Piaget, 1973). Brainerd (1972) demonstrated this relationship for the object properties of length and weight. The pretest data of the Stage III subjects reflected this relationship in that a comparison of the stage obtained on the transitivity and substance tasks demonstrated that in those cases where the stage obtained on the two tasks were different, transitivity obtained the higher rating. Comparing each substance conservation test with transitivity in turn results in 5 comparisons; all 5 comparisons supported the above stated difference at the $p < .001$ level of confidence (all comparisons were one tailed).

The conservation of number test also appears to be acquired prior to substance conservation, not withstanding Piaget's (1952b) suggestion that number and substance conservation develop synchronously. The earlier acquisition of number conservation has been demonstrated with both those of normal intelligence (Brainerd and Brainerd, 1972; Goldschmid and Bentler, 1968) and the MR (Gruen and Vore, 1972; Lister, 1970, 1972; Hunter and Lister, 1970). Again the data of the Stage III subjects obtained in the present experiment supported the sequence. When conservation of number rating was compared in turn with the rating of the substance conservation test, all 5 possible comparisons were significant at $p < .01$ level (all comparisons were one-tailed).

While the conservation of number and the transitivity of substance skills are acquired prior to the conservation of substance, the conservation of weight is usually acquired subsequent to the conservation of substance. This invariant sequence has been repeatedly demonstrated with those of normal intelligence (e.g., Hooper et al., 1971) and the MR (McManis, 1969a; Gruen and Vore, 1972; Lister, 1970, 1972; Hunter and Lister, 1970). The performance of the Stage III subjects in the present experiment also supported this sequence. Of the five comparisons of the stage obtained on each, 4 were significant at the $p < .1$ level (all comparisons were one-tailed). Hence little or no generalization should take place from the conservation of substance training to the subsequent tests of the conservation of weight.

The second dimension of equilibrium assessed was the stability of the structure (Flavell, 1963, p. 242). This dimension was assessed by comparing the explanations and judgements obtained on the two posttests which were separated by a one to two week interval (cf., Piaget, 1970a).

CHAPTER III

RESULTS

Pretest

Although all subjects participating in the induction phase of the experiment were equated on stage of operational functioning on the basis of their pretest performance, the following analyses test their equivalence on the pretest with respect to the number of correct judgements and the number of appropriate explanations. Due to heterogeneous within cell variances, the judgement scores were transformed to arcsin scores. The summary table for the age group x treatment x block of tests analysis of variance is reported in Table 4. Post-hoc comparisons

INSERT TABLE 4 ABOUT HERE

(Tukey (A) Method; Winer, 1962) indicated that the main effect of the age group factor was a consequence of the N group making fewer correct judgements on the Pretest than either the OMR or YMR subjects ($p < .05$); the number of correct judgements made by subjects of the two MR groups did not differ significantly. The main effect of the block of tests factor occurred because more correct judgements were made on the combined conservation of number and transitivity of substance (transitivity-number) tests than on the conservation of substance (substance) or the conservation of weight (weight) tests ($p < .05$). Furthermore, more correct judgements were made on the substance tests than on the weight test ($p < .05$). This sequence of test difficulty, as judged by correct judgements, parallels the sequence found in the Stage II and III subjects.

Further post-hoc comparisons of the interaction of age group and block of tests factors indicated that on two of the blocks (transitivity-

TABLE 4

Analysis of Variance

Effects of Age Group, Treatment Grouping, and Block of Tests on Pretest Judgement Scores.

Source	SS	df	MS	F
Between subjects				
Age Group (A)	6.2713	2	3.1357	12.25***
Treatment Group (B)	.0248	1	.0248	.10
AB	.5005	2	.2503	.98
Subjects within groups	19.9681	78	.2560	
Within subjects				
Blocks of Tests (C)	4.2056	2	2.1028	41.72***
AC	.7295	4	.1824	3.62**
BC	.1454	2	.0727	1.44
ABC	.3478	4	.0870	1.73
C x Subjects within group	7.8610	156	.0504	

Note. - Judgement scores transformed to arcsin.

** p .05

*** p .01

number and substance) the YMR group made more correct judgements than the N group ($p<.05$); also the OMR group made more correct judgements than the N group ($p<.05$); however, the performances of the two MR groups failed to differ significantly. On the weight tests, the YMR subjects made more correct judgements than subjects in either the N or the OMR groups ($p<.05$); in addition, the subjects in the OMR group made more correct judgements than the subjects in the N group ($p<.05$). Within group comparisons of task difficulty showed that both the N and YMR groups made more correct judgements on the transitivity - number tests than on the substance tests ($p<.05$) or the weight test ($p<.05$); the number of correct judgements made on these latter two blocks of tests failed to differ significantly. Subjects in the OMR group made more correct judgements on the transitivity - number tests than the weight tests ($p<.05$), on the substance tests than the weight test ($p<.05$), but their performance on the transitivity - number and the substance tests failed to differ significantly.

Parametric comparisons of the number of appropriate explanations offered by each age x treatment group on the pretest were precluded by heterogeneous within cell variances on the transitivity - number tests (F_{max} = 20.02, 17/13, $p<.01$), the substance tests (F_{max} = 398.15, 17/13, $p<.01$) and the weight test (F_{max} = 25.19, 17/13, $p<.01$). Consequently, the number of appropriate explanations offered by each age x treatment group were contrasted using the Kruskal-Wallis one-way analysis of variance for ordinal data[1] (Siegel, 1956). These comparisons indicated that there was no difference in this measure on the substance tests (H = 7.81, $p>.10$) or on the weight test (H = 3.76, $p>.50$). However, there was a significant difference among the groups on the transitivity - substance tests ($H = 12.38$, $p < .05$). Post-hoc comparisons using the Wilcoxon matched-pairs signed ranks test of the individual experimental - control groups within each age group did not differ for the N group ($z = .67$, $p > .45$), the YMR group ($z = 1.12$, $p > .30$), or the OMR group ($z = 1.13$, $p > .30$). Hence the experimental - control subjects were combined and differences between the age groups were tested using the Mann-Whitney μ test. These analyses showed that the YMR subjects gave more appropriate explanations than the N subjects ($z = 2.33$, $p = .02$); the OMR subjects gave more adequate explanations than the N subjects at a borderline level of significance ($z = 1.77$, $p = .07$); and the number of appropriate explanations offered by the two MR groups failed to differ significantly ($z = .86$, $p = .39$).

The foregoing analyses indicate that even though the subjects in the various age groups were equated on stage of operational functioning, they were not necessarily equivalent with respect to the dependent variables of correct judgements or appropriate explanations. Also, even though the subjects were all classed at Stage I on all of the operational tests, it is evident that more correct judgements were made on some of the tests than others. Significant differences were limited to between age groups comparisons - all the comparisons of the experimental and control subjects within each age group failed to attain significance.

Induction Phase

The hypothesis that N>YMR>OMR in terms of change resulting as a consequence of training (hypothesis 1) was tested by comparing the number of subjects in each group (a) answering all 3 judgement questions correctly on one or more induction trials and (b) answering all 3 judgement questions correctly on two or more induction trials. For the N, YMR, and OMR experimental groups, 11, 10 and 11 subjects respectively answered all questions correctly on at least one trial. While there was no difference in the number attaining this criterion in each of the experimental groups ($p > .10$, Fisher Probability Test, Siegel, 1956) when each experimental group was compared to its comparable aged control group on this cirterion, it was evident that the subjects in the experimental groups were superior ($p < .005$ in all three comparisons). When the criterion is set at two or more perfect trials in the induction phase, 8, 5, and 6 subjects from the N, YMR and OMR respectively attained the criterion. Again, there was no difference in the number of subjects meeting this criterion in the three experimental groups ($p > .10$) but the within age experimental-control group comparisons were all significant ($p < .05$ in each comparison). Consequently, there was no support for the hypothesis when the stage analysis based on judgements was the dependent variable.

In comparing the number of subjects who made 3 correct judgements on a number of trials, hypothesis 1 was tested by using non-parametric tests, for it was presumed that the datum reflected dichotomous stages which could be ordered. The validity of a stage analysis is questionable (Brainerd, 1974; personal communication); the most powerful analysis which would test the hypothesis would involve assigning each correct judgement a score of "1" and each incorrect judgement a score of "0". Differences in the mean number of correct judgements per trial were tested using parametric statistics. Figure 1 portrays the relevant data.

INSERT FIGURE 1 ABOUT HERE

Heterogeneous within cell variances (F_{max} = 8.2, 6/13, p<.01) precluded testing the overall effectiveness of training with a 3 (age groups) x 2 (treatments) analysis of variance. The effectiveness of the induction procedures in increasing correct judgements for subjects in the experimental groups, as opposed to those in the corresponding control groups, can be seen by reference to Table 5 which reports the results for tests for the simple effects of main factors.

INSERT TABLE 5 ABOUT HERE

A further test of hypothesis 1 was made by contrasting the slopes of the individual learning curves over the induction trials. The trials for each subject were divided into four consecutive blocks of trials. Consequently the number of trials per block of trials could vary between subjects: subjects needing 12 trials had 3 trials per block whereas subjects needing 8 trials had 2 trials per block, etc. Since a slope analysis of each individual learning curve is only justified if it can first be shown that the data is linear, the group curves of the three experimental groups were tested for linear, quadratic, cubic and higher order components. For the N experimental group, the linear component accounted for 70 percent of the variation (F = 38.54, 1/39, p<.01); the quadratic component accounted for 22 percent of the variation (F = 12.00, 1/39, p<.01); and the cubic component accounted for 8 percent of the variation (F = 4.46, 1/39, p<.05). For the YMR experimental group the linear component accounted for 97 percent of the variation (F = 39.46, 1/39, p<.01); neither the quadratic or the cubic components contributed

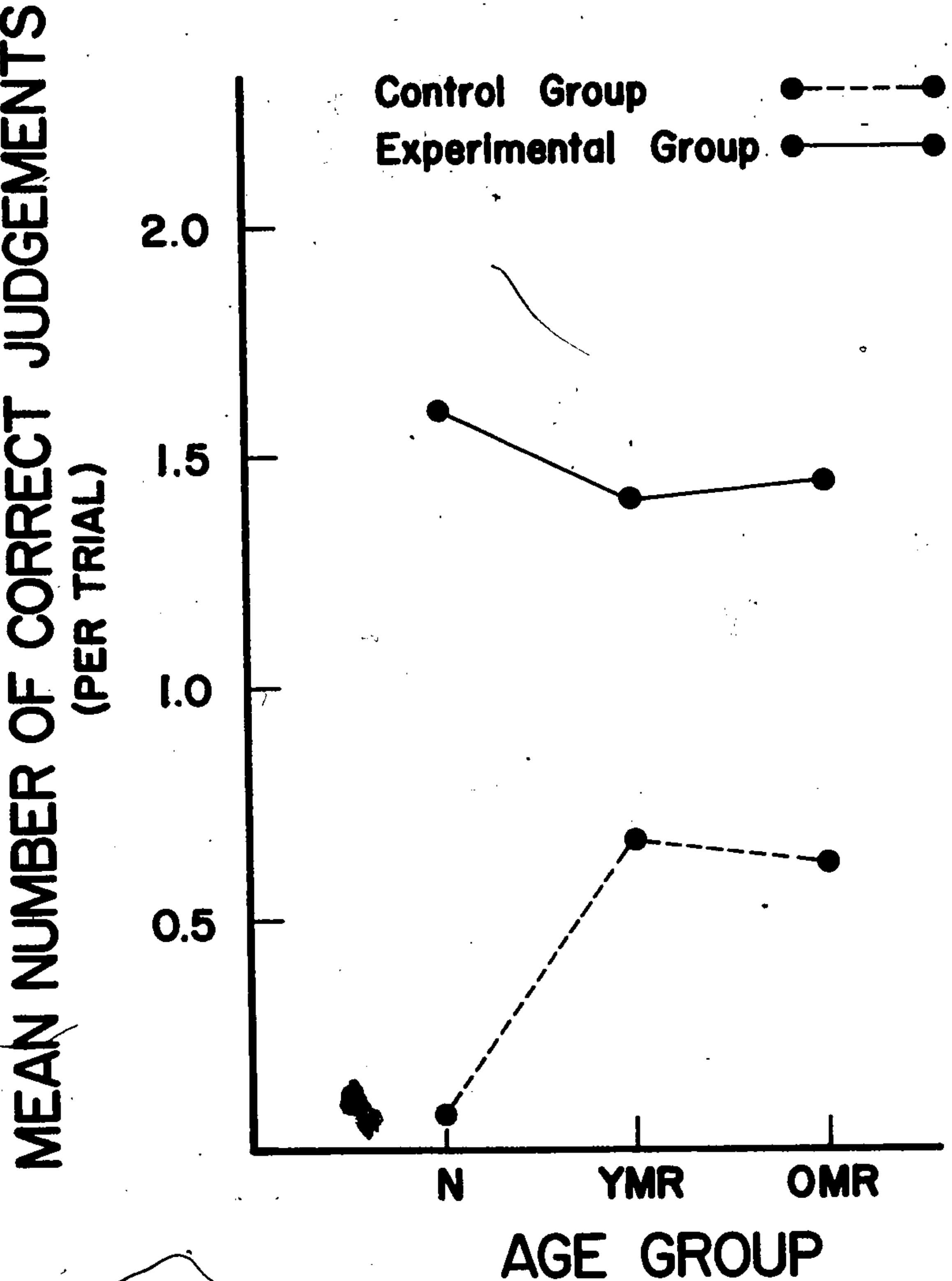

Figure 1. Mean number of correct judgements per trial during the induction phase

TABLE 5

Analysis of Variance:

Simple Effects of Treatment and Age Group

Source of Variation	df	MS	F
Simple Effect of Treatment:			
For Normals	1	21.30	34.35***
w.cell error	13	.62	
For YMR	1	5.76	11.75***
w.cell error	13	.49	
For OMR	1	6.17	9.49**
w.cell error	13	.65	
Simple Effects of Age Grouping:			
For Experimental Subjects	1	.78	.68
w.cell error	13	1.14	
For Control Subjects	1	2.64	4.25
w.cell error	13	.62	

Note. - All comparisons were one tailed except for the comparison of control groups in testing the simple effects of age grouping -- this comparison was two tailed.

* $p<.10$
** $p<.05$
*** $p<.01$

significantly to the remaining variation. Similarly, for the OMR group, the linear component accounted for 98 percent of the variation ($F = 46.76$ $1/39$, $p<.01$) while neither the quadratic or cubic components accounted for significant amounts. Higher order components did not account for a significant amount of the variation for any of the experimental groups.

Given that the linear component accounted for most of the variation in all three groups, the slope of each of the individual learning curves was determined. The mean slope for the N, YMR and OMR groups were .0712, .0898, and .0786 respectively. As was the case for the other analyses testing hypothesis 1, the analysis of the slopes of the individual learning curves failed to indicate any significant differences attributable to the age group factor ($F = .28$, $2/39$, $p>.25$).

Posttest I

Hypothesis 2 - that the three experimental groups would be ordinated in their performance on Posttest I - was tested separately with reference to (a) the conservation of substance tests, (b) the transitivity of substance and conservation of number tests and (c) the conservation of weight test. For each block of tests, three different sets of analyses were conducted. First, the judgement data was used to determine stage of operational functioning and the stages attained by subjects in each group were contrasted. Second, each correct judgement was given a score of "1" while each incorrect judgement was given a score of "0" and relevant parametric analyses were conducted. This set of analyses should be more sensitive to changes resulting from the induction procedure than the stage analyses. And third, the number of appropriate explanations given by each group were contrasted.

Stage Analyses. The judgement determined stage analyses contrasted the number of subjects in each treatment x age group attaining stage II or III. For the analyses on all three blocks of tests it was necessary to combine the stage II and III subjects because of low frequencies. Table 6 reports the relevant proportions.

On the substance tests all between group comparisons proved to be

INSERT TABLE 6 ABOUT HERE

non-significant (Fisher Exact Probability Test, $\alpha = .10$, one tailed comparisons). In addition, not one of the proportions is significantly greater than zero (Binomial Test, $\alpha = .10$, one tailed comparisons); thus no group showed a significantly greater number of stage II or III subjects on Posttest I than on the Pretest.

With reference to the transitivity - number section of Table 6, there were significantly more N experimental subjects attaining stage II or III than YMR experimental ($p<.10$) or OMR experimental ($p<.05$) subjects; the difference between the number of OMR and YMR experimental subjects attaining stage II or III was not significant ($p>.10$). Within age group comparisons of the experimental and control subjects also using the Fisher Exact Probability Test ($\alpha = .10$) indicated that the number of subjects attaining stage II or III in the N experimental group was greater than was the case for the N control group at the borderline level of significance ($p<.10$); however, there were no significant differences between the experimental and control subjects in the two MR groups. Comparisons of the proportion of subjects attaining stage II or III on Posttest I, relative to the proportion on the Pretest, indicated that only the N experimental group's performance had improved ($p<.05$, Binomial Test).

TABLE 6

Proportion of Subjects Attaining Stage II or III on Posttest I

Group	Block of Tests		
	Substance	Transitivity - Number	Weight
Normal			
Experimental	.14	.43	.14
Control	.00	.14	.00
YMR			
Experimental	.07	.14	.00
Control	.00	.14	.00
OMR			
Experimental	.07	.07	.00
Control	.00	.07	.00

Hence the stage analyses of the transitivity-number tests on Posttest I would support at least segments of hypothesis 2 in that when the performance of the experimental groups was contrasted N>YMR ($p<.10$) and N>OMR ($p<.05$).

With reference to the weight section of Table 6, all comparisons indicated that there were no significant differences in the number of subjects from each age x treatment group attaining Stage II or III on Posttest I (Fisher Exact Probability Test, $\alpha = .10$). It was also evident that the proportion of subjects attaining Stage II or III on Posttest I failed to differ significantly from the proportion at these stages on the Pretest (Binomial Test, $\alpha = .10$).

Correct Judgements. As was noted in the analyses of the number of correct judgements made on the Pretest, differences existed between at least two of the age groups on each block of tests. In evaluating the effect of the experimental treatment and age group factors on subsequent posttests, an adjustment was made for the effect of the variation due to the initial differences in the number of correct judgements. Analyses of covariance were used to analyze the posttest data, with the pretest data providing the covariate. Thus, for example, the number of correct judgements made by each subject on the substance tests of the Pretest provided the covariate for the analysis of the number of correct judgements made on the substance tests of Posttest I. Table 7 presents the summary table for the analyses of covariance. Not one of the F ratios

INSERT TABLE ABOUT HERE

attained $\alpha = .10$ level of significance; hence, when adjustment was made for differences on the Pretest, no significant effects could be attri-

TABLE 7

Analyses of Covariance of Correct Judgements on Posttest I

Source	SS	df	MS	F
Substance				
Experimental Treatment (A)	.258	1	.258	2.13
Age Group (B)	.513	2	.256	2.11
AB	.151	2	.076	.62
Error	9.336	77	.121	
Transitivity - Number				
Experimental Treatment (A)	.387	1	.387	2.28
Age Group (B)	.281	2	.140	.82
AB	.719	2	.360	2.11
Error	13.103	77	.170	
Weight				
Experimental Treatment (A)	.629	1	.629	2.63
Age Group (B)	.100	2	.050	.21
AB	.803	2	.402	1.68
Error	18.398	77	.239	

Note. - Judgement scores transformed to arcsin.

buted to either the age group or the experimental treatment factors on any block of tests.

Appropriate explanations. The explanations given on the tests of Posttest I provide another dependent measure with which to test hypothesis 2. The total number of appropriate explanations offered by the subjects of each age x treatment group on the five conservation of substance tests were contrasted using the Kruskal-Wallis analysis of variance for ordinal data. The overall analysis indicated a difference between the groups at the borderline level of significance (H = 9.62, .10<p>.05). The within age group experimental - control comparisons using the Wilcoxon matched-pairs signed-ranks test indicated that within the YMR group the experimental subjects offered more appropriate explanations than the control subjects (z = 1.86, p = .03), but that no significant differences existed between the two treatment groups within the N (insufficient differences to make statistical comparisons) and OMR (z = 1.24, p = .11) age groups. As there was no difference in the experimental and control groups of the OMR and N age groups, the experimental and control groups in these two age groups were combined for further post-hoc comparisons using the Mann Whitney μ test. On the substance test of Posttest I, the OMR subjects offered significantly more appropriate explanations than the N subjects (z = 2.19, p = .03). The performance of the YMR experimental group was superior to that of the N group at the borderline level of significance (z = 1.73, p = .08) but not significantly different from that of the OMR group (z = .01, p = .99). The performance of the YMR control group was not significantly different from that of either the OMR group (z = 1.36, p = .17) or the N group (z = .63, p = .53).

There was no significant difference in the total number of appropriate explanations offered by the subjects of each age x treatment group on the transitivity - number tests (H = 5.38, $p>.30$) or the weight test (H = 4.67, $p>.30$) of Posttest I.

Posttest II

The same sets of analyses used to test hypothesis 2 on Posttest I were also used to test the hypothesis as it related to Posttest II performances.

Stage Analyses. Table 8 reports the proportion of subjects attaining Stage II or III on the various blocks of tests of Posttest II. On the

INSERT TABLE 8 ABOUT HERE

substance tests, all between group comparisons proved to be non-significant (Fisher Probability Test, $\alpha = .10$, one tailed comparisons). Furthermore, since the proportion of subjects demonstrating Stage II or III functioning failed to differ significantly from zero (Binomial Test, $\alpha = .10$, one tailed comparisons), it is evident that no group showed a significantly greater number of Stage II or III subjects on the substance tests of Posttest II than on the Pretest.

With reference to the transitivity - number section of Table 8, between group comparisons using the Fisher Exact Probability Test indicated that more subjects from the N experimental group attained Stage II or III than subjects from the OMR experimental group ($p<.05$), the OMR control group ($p<.05$), and the YMR control group ($p<.05$). However, the proportion of Stage II or III subjects in the N experimental did not differ from that in the N control or the YMR experimental groups on Posttest II ($p>.10$). The proportion of N control subjects demonstrating

TABLE 8

Proportion of Subjects Attaining Stage II or III on Posttest II

Group	Block of Tests: Substance	Block of Tests: Transitivity - Number	Block of Tests: Weight
Normal			
Experimental	.07	.43	.07
Control	.14	.29	.07
YMR			
Experimental	.21	.21	.00
Control	.00	.00	.00
OMR			
Experimental	.00	.00	.00
Control	.00	.00	.00

Stage II or III function was greater than the proportion of subjects from the OMR control, the OMR experimental and the YMR control groups ($p<.10$). There were no significant differences between any of the MR age x treatment groups. Only the proportion of subjects from the N experimental ($p<.05$) and N control ($p<.10$) groups differed significantly from zero (Binomial Tests, $\alpha = .10$).

On the weight test of Posttest II (refer to Table 8), all comparisons between the age x treatment groups failed to indicate any significant difference in the number of subjects attaining stage II or III (Fisher Exact Probability Test, $\alpha = .10$, one tailed comparisons). Furthermore, as not one of the proportions is significantly greater than zero (Binomial Test, $\alpha = .10$, one tailed comparisons), it is evident that the number of Stage II or III subjects on the weight test of Posttest II failed to differ from the number of Stage II or III subjects on the weight test of the Pretest.

Correct Judgements. In analyzing group differences in the number of correct judgements made on the various blocks of tests of Posttest II, the covariate was the number of correct judgements made on the same block of tests on the pretest. The judgement scores had been subjected to an arcsin transformation because of heterogeneity of within cell variances. Table 9 presents the summary tables for the analyses of covariance for each block of tests.

With reference to the substance section of Table 9, the main effect

INSERT TABLE 9 ABOUT HERE

of the experimental treatment factor was due to the superiority of the experimental subjects over the control subjects ($p<.05$). The adjusted

TABLE 9

Analysis of Covariance of Correct Judgements on Posttest II.

Source	SS	df	MS	F
Substance				
Experimental Treatment (A)	.927	1	.927	6.07**
Age Group (B)	.259	2	.130	.85
AB	.508	2	.254	1.70
Error	11.766	77	.153	
Transitivity - Number				
Experimental Treatment (A)	.862	1	.862	4.37**
Age Group (B)	.432	2	.216	1.10
AB	.152	2	.076	.37
Error	15.190	77	.197	
Weight				
Experimental Treatment (A)	.514	1	.514	1.59
Age Group (B)	.088	2	.044	.14
AB	.089	2	.044	.14
Error	24.872	77	.323	

Note. - Judgement scores transformed to arcsin.

**p .05

means of the arcsin transformed scores were 1.00 and .81 for the experimental and control subjects respectively. The lack of a main effect for the age group factor and the non-significance of the interaction suggest that the experimental treatment was equally effective for subjects in all three experimental groups.

The same conclusions can be made concerning the performance of the subjects on the number - transitivity tests of Posttest II. The main effect of the experimental treatment indicates that the experimental subjects (adjusted $\bar{x}$ = 1.23) made more correct judgements than the control subjects (adjusted $\bar{x}$ = 1.07). The failure to obtain significant effects for the age group factor or for the interaction of age group x treatment condition would suggest that the experimental treatment was equally effective in increasing the number of correct judgements for all three experimental groups on the transitivity - number tests of Posttest II.

With reference to the weight section of Table 9, however, it can be seen that there was no significant difference in the number of correct judgements made by the experimental and control subjects on the weight test of Posttest II. Thus, when an adjustment was made for the initial differences in correct judgements on the Pretest, no significant effects could be attributed to either the age group or experimental treatment factors on the number of correct judgements on the weight test of Posttest II.

Appropriate Explanations. The number of appropriate explanations given by the subjects of each age x treatment group were contrasted using the Kruskal-Wallis analysis of variance for ordinal data. There were no significant differences in the number of appropriate explanations given on the substance tests ($H = 7.22$, $p>.20$), the transitivity - number tests ($H = 4.25$, $p>.50$) or the weight test ($H = 3.26$, $p>.50$) of Posttest II.

In summarizing the analyses of Posttest I and Posttest II performance, the judgement-based stage analyses indicated that the effects of the experimental induction procedure were limited to the transitivity - number tests of the two posttests. On both posttests it was found that the performance of the N experimental group was superior to that of the OMR experimental and control groups ($p<.05$). This would suggest that the age group factor was of importance in determining transitivity - number performances. The stage analyses also cast some doubt on whether the improvement in performance should be attributed to the experimental induction procedures: within age group comparisons of the experimental and control groups resulted in only one significant ($\alpha = .10$) comparison and that was between the N experimental and control groups on Posttest I. The difference in performance between these two groups was no longer significant on Posttest II.

In contrasting the number of correct judgements made on Posttest I, it was found that if adjustments were made for the initial differences in the number of correct judgements given on the Pretest, then significant effects ($\alpha = .10$) could not be attributed to either the experimental treatment or the age group factors. On the other hand, these analyses indicated that on Posttest II, significant effects could be attributed to the experimental treatment factor on the substance and transitivity - number tests, but not on the weight test.

The stage analyses and the correct judgements analyses, then, were somewhat contradictory. First, the stage analyses suggested that the effects were limited to the transitivity - number tests on both posttests whereas the analyses of correct judgements suggested that the effects were limited to the transitivity - number and substance tests of Posttest II. Second, the stage analyses tended to suggest that the age group factor was relevant whereas the analyses of correct judgements suggested that the experimental treatment condition was the relevant factor, at least on Posttest II.

The comparisons of the number of appropriate explanations given by the various groups on Posttest I and Posttest II indicated that the only significant differences in performance occurred on the substance tests of Posttest I. It was found that the OMR subjects gave more appropriate explanations than the N subjects (p = .02) and that the YMR experimental subjects gave more appropriate explanations than the N subjects (p = .08). When the comparisons were limited to the experimental groups - these are the groups that hypothesis 2 predicts would be ordinated - it was found that the only comparison to attain significance (α = .10) was the one indicating the superiority of the OMR group over the N group (z = 2.08, p = .04).

The superiority of the OMR experimental group may well be a continuation of the superiority, albeit not quite a significant one, of this group on the Pretest. Examination of Figure 2, which graphically presents the proportion of subjects in each experimental group giving one or more appropriate explanations on each of the test days for each block of tests, supports this interpretation. The aforegoing interpretation is also supported by the finding that [illegible] case did the subjects of the experimental groups show a signific[illegible]ease in the number of appropriate

INSERT FIGURE 2 ABOUT HERE

explanations offered over the course of the experiment on any block of tests. Friedman's two way analysis of variance were used to contrast the number of appropriate explanations given on the Pretest, Posttest I and Posttest II. On the substance tests there was no significant improvement for the N ($\chi_r^2 = .18$, $p > .90$), the YMR ($\chi_r^2 = .36$, $p > .80$) or the OMR ($\chi_r^2 = .36$, $p > .80$) group. On the transitivity - number test there was no significant improvement for the N ($\chi_r^2 = 1.68$, $p > .30$), the YMR ($\chi_r^2 = .18$, $p > .90$) or the OMR ($\chi_r^2 = .03$, $p > .98$) group. On the weight test there was no significant improvement for the N ($\chi_r^2 = 4.40$, $p > .10$), the YMR ($\chi_r^2 = 2.27$, $p > .30$) or the OMR ($\chi_r^2 = .61$, $p > .70$) group. It appears that the experimental induction procedures, then, did not result in a significant improvement in appropriate explanations for any experimental group on any block of tests.

Relation of Posttest I and Posttest II performances

Hypothesis 4 suggests that the structural learning that occurred would be more stable for the subjects in the N experimental group than for the subjects in the YMR and OMR experimental groups, and that it would be more stable for subjects in the YMR experimental group than subjects in the OMR experimental group. For the judgement related stage analysis subjects failing to obtain a stage II or III response on both of the posttests were eliminated from the analysis. Table 10 reports the number of subjects (a) showing no change or improving on their performance from Posttest I to Posttest II and (b) showing a decrease

INSERT TABLE 10 ABOUT HERE

from Posttest I to Posttest II on the substance, transitivity - number and weight tests. In the judgement based analysis the number of Stage II

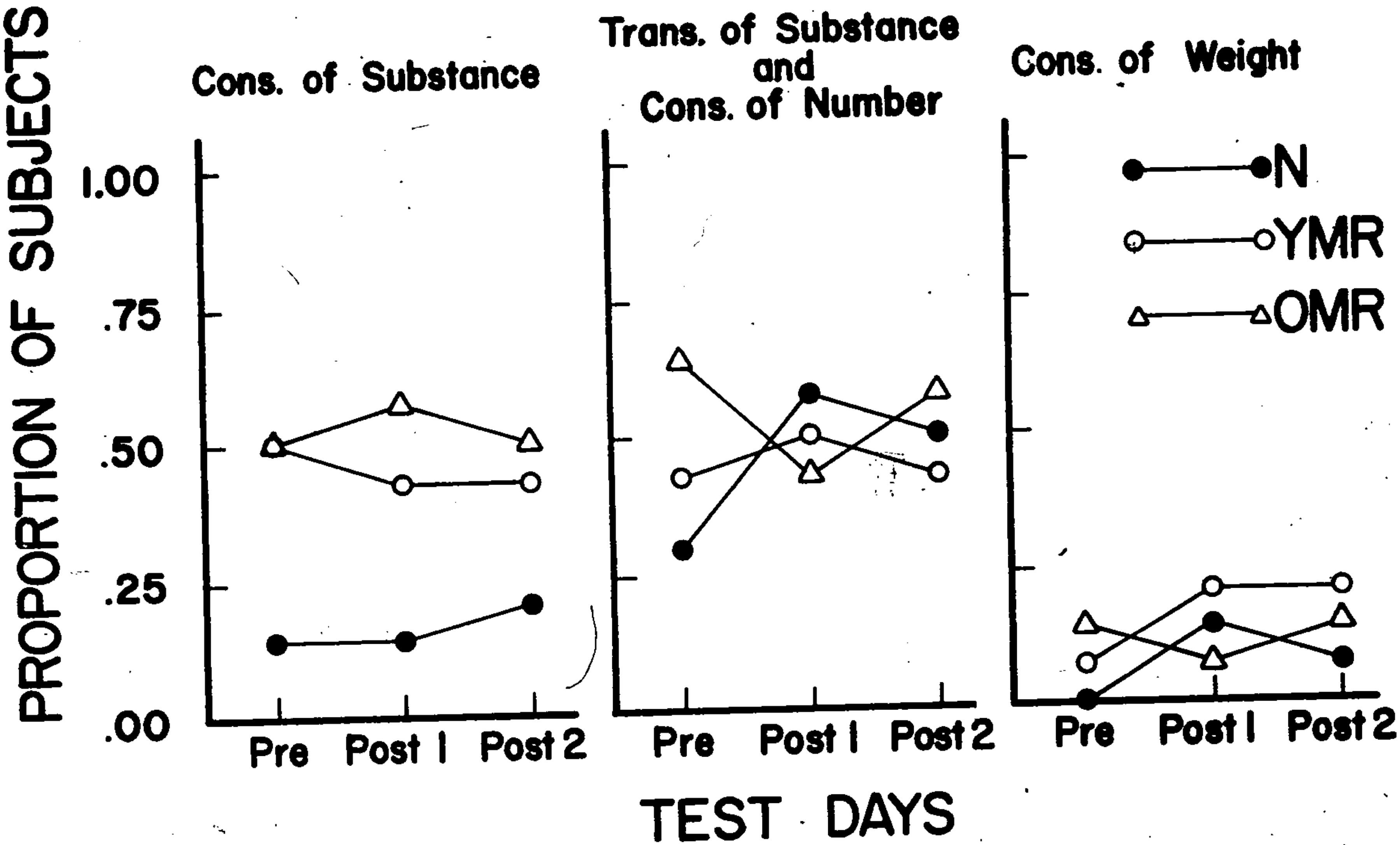

Figure 2. Proportion of subjects in each experimental group offering one or more "appropriate" explanations

TABLE 10

Number of Subjects Showing No Change or Decrease on Judgements and Explanations Between Posttests

Group	Substance		Transitivity - Number		Weight	
	No Change	Decrease	No Change	Decrease	No Change	Decrease
	Judgement Determined Stages					
N	1	1	6	1	1	1
YMR	4	0	3	1	0	0
OMR	0	1	0	1	1	1
	Explanations					
N	1	2	6	3	1	1
YMR	5	2	6	2	2	1
OMR	4	4	5	4	1	3

Note. - No change includes subjects performing at same level and those subjects improving their performance from Posttest I to Posttest II. Decrease category includes subjects performing at a lower level on Posttest II than on Posttest I. The analysis was limited to subjects from the experimental condition.

and Stage III performances on the combined test were counted. Stage II performance counted as 1 point, Stage III as 2 points. Within group comparisons contrasting the number of subjects maintaining a stable level or improving versus those showing a decrease from Posttest I to Posttest II all proved non significant (Sign Test, $\chi = .10$, two tailed comparisons). Between group comparisons of the experimental groups taken two at a time also failed to reach significance in each instance (Fisher Exact Probability Test, $\alpha = .10$, one tailed comparisons). Thus all comparisons based on the judgement data indicated that there is no difference in the stability of the structural learning that did occur.

The data based on the explanations given by the subjects is also reported in Table 10. Again subjects failing to provide an appropriate explanation on either of the posttests were eliminated from the analyses. The subjects were dichotomized into those (a) providing the same number of appropriate explanations on the two posttests and those showing an increase in appropriate explanations from Posttest I to Posttest II or (b) those showing a decrease in appropriate responses from Posttest I to Posttest II. As with the comparisons on the judgement data, all within cell comparisons proved to be non significant (Sign test, $\alpha = .10$, two tailed comparisons) and all comparisons of the experimental groups proved to be non significant (Fisher Exact Probability Test, $\alpha = .10$, one tailed comparisons). Both the explanation and stage analyses, then, failed to show any differences between the experimental groups on the stability dimension.

CHAPTER IV

DISCUSSION

The results from the present experiment failed to provide support for any of the hypotheses derived from Inhelder's thesis that mental retardation is characterized by, if not due to, deficiencies in equilibration. It was predicted that on the induction task the performance of the three experimental groups would be ordinated so that N>YMR>OMR (Hypothesis 1). While it was clear that the induction procedures were effective in increasing the number of correct judgements on the conservation of substance task used in the induction phase, three separate sets of analyses failed to provide any support for the hypothesis. It was evident that while the N, YMR, and OMR groups increased the number of correct judgements made on the induction task relative to the control groups, there were no significant differences among the three experimental groups.

Hypothesis 2 also failed to receive much support; this hypothesis predicted that because of the postulated differences in equilibration, what structural learning did take place during the induction phase would transfer differentially to the blocks of tests on Posttest I and II: in each instance the ordination of the experimental groups should have been N>YMR>OMR. The stage analyses suggested that the effect of the age factor was significant on the transitivity - substance tests of the two Posttests; whereas the analyses of correct judgements led to the conclusion that the effect of the age factor was not significant on any of the Posttests, but that on the transitivity - number tests and substance tests of Posttest II the experimental induction procedures led to more correct judgements. The apparent support of hypothesis 2 by the stage analyses must be considered in light of the performance of the three control groups, and also in light of the analyses based on the number of correct judgements.

With regard to the performance of the control groups, as compared to the experimental groups, the stage analysis indicated that the superiority of the N experimental group over the N control group was evident only on the transitivity - number tests of Posttest I ($p<.10$). By the time the Posttest II tests were administered, the difference between the two groups was no longer evident. In fact, the performance of the N control group was superior to that of the other two control groups and the OMR experimental group on the transitivity - number tests of Posttest II ($p<.10$). This suggests that the experimental procedure may not have been solely responsible for the superior performance of the N experimental subjects. The repeated testing, which resulted in each subject making 147 to 171 judgements and explaining 135 of these, may alone have been a sufficient condition to initiate disequilibration in the N subjects.

However, as indicated previously, the analyses of correct judgements suggested that the age group factor did not significantly contribute to an increase in correct judgements. In weighing the relative validity of the two sets of analyses (stage analyses and correct judgement analyses), several factors should be considered. First, the judgement analyses are more powerful statistically in that the stage analyses used dichotomous stages whereas the judgement analyses allowed for considerably more variation. Second, the judgement analyses took into account differences in pretest levels whereas the stage analysis did not. Third, the stage analyses actually consisted of a large number of individual group comparisons (two groups at a time); this procedure is very susceptible to fallacious conclusions in that the possibility of Type I error is inflated over the $\alpha = .10$ level (Siegel 1956, pp. 159 - 161). In general, then, the analyses of correct judgements is probably a much more valid test of the hypothesis. Consequently there is little support for hypothesis 2 in that the analyses of correct judgements indicated that all three experimental groups benefitted equally from the induction techniques, as was measured by the increase of correct judgements on Posttest II.

The analyses conducted on the number of correct explanations similarly failed to provide any support for hypothesis 2. Not a single analysis conducted with the explanation measures proved to be significant in the direction predicted by hypothesis 2; in fact, the within group analyses of the number of appropriate explanations made by each treatment x age group indicated there was no significant change with reference to this measure during the course of the experiment.

With reference to hypothesis 3, which predicted that the degree of transfer from the induction task would vary as a function of the age group and block of test factors, there was no evidence to support this hypothesis. It will be recalled that in order for hypothesis 3 to be adequately tested, hypothesis 2 needed to be supported. Since hypothesis 2 failed to be supported, hypothesis 3 cannot be supported by the results of the experiment.

Hypothesis 4, which predicted that the stability of performance from Posttest I to Posttest II should vary inversely with the age of the subjects, also failed to receive any support. Both the stage analyses and the explanation analyses indicated there were no differences in performance on Posttest I as compared to Posttest II for any of the experimental groups. Furthermore, the judgement analyses indicated that while the number of correct judgements made on Posttest I did not differ from the number made on the Pretest, more correct judgements were made on Posttest II than on the Pretest ($p<.05$). This would indicate that there was actually an improvement in the number of correct judgements made on Posttest II relative to Posttest I. Since the age group factor was not significant, but the experimental treatment factor was significant on the analyses of the substance and transitivity-number tasks, it is suggested that all experimental groups demonstrated similar gains. Hence, the results failed to support hypothesis 4.

The implications of the present research for Inhelder's (1968) formulations are constrained by two aspects of the results. First, the training procedures, while leading to a significant increase in the correct judgements for the three experimental groups, were not of sufficient power to change the responding of all of the subjects in each of the groups. This, in effect, limited the transfer to the posttests. The training procedure was not as effective for the normal group as may have been expected from the literature (e.g., Brainerd, 1972c). Although Brainerd (1972c) does not report the exact number of subjects demonstrating correct judgements during the training phase, 75 percent of the 20 subjects trained gave correct judgements on a substance conservation posttest after 4 training trials. By way of contrast, only 57 percent of the subjects in the *N* experimental group attained the criterion of two trials of 3 correct judgements during the induction phase after up to 12 training trials. Several differences in the two samples may have been responsible for the less effective training in the present experiment: (a) The normal children in the present experiment were drawn from kindergarten classes whereas Brainerd's (1972c) sample was comprised of grade one children. (b) In the present experiment, subjects were precluded from participation in training if they demonstrated concrete-operational thinking on a transitivity of substance or number conservation test. Assuming that transitivity is the earliest indication of concrete operational functioning (Piaget, 1973), the subjects in the present sample were clearly preoperational. Brainerd (1972c) pretested only for conservation of substance; consequently, some of his subjects may have already attained the stage of concrete operational thought.

Second, the finding that there was no significant improvement in appropriate explanations given over the course of the experiment, places some severe constraints on the implications of the present experiment for the theory - at least from a Genevan perspective. The Genevans (e.g., Inhelder, 1968) place considerable significance on using explanations as the dependent measure of cognitive structures, hence suggesting that the present experiment failed to support the hypotheses because no structural learning took place. However, the finding that some of the judgement analyses were significant, but not the explanation based analyses, is interpretable when one considers the posited relationship of cognitive structures with the two dependent measures (Brainerd, 1972a). Brainerd notes that according to Piagetian theory, operational structures antedate linguistic structures and that the latter develop from the former (e.g., Piaget, 1970b). Consequently, on a temporal basis, this means that a judgement based criterion will indicate the presence of an operational structure prior to an explanation based criterion. Support for this comes from demonstrations that explanations provide a much more difficult criterion for cognitive structures with both normal (Brainer and Brainerd, 1972) and retarded (Woodward, 1961; Boehm, 1967) samples.

Within the limitations of the design, the present experiment failed to provide support for the postulate that normal and mentally retarded individuals differ on the dimension of equilibration of cognitive structures. The results of the present experiment are in accord with those obtained by Brison and Bereiter (1967). Brison and Bereiter failed to find a difference in the proportion of subjects trained from groups of gifted, normal or retarded subjects in response to a conservation of substance training task. It is evident, however, that the training procedures used by Brison and Bereiter were much more effective in changing responding. This is probably because, relative to the training procedures in the present experiment, those of Brison and Bereiter were much more complex and lengthy. The Brison and Bereiter (1967) training task included, but was not limited to, (a) the object bound form of inversion reversibility, (b) models of correct responding, (c) models of appropriate explanations, and (d) either explicit or implicit information concerning the correctness of responding during training.

In summary, Inhelder (1968) posited that differences in (a) rates of cognitive structural development and (b) the upper level of cognitive structure attained between normal and mentally retarded individuals could be attributed to differences in the power of equilibration. It was reasoned that the three groups in the present experiment varied on the equilibration dimension so that N>YMR>OMR. Furthermore, as the OMR group was of an age that Inhelder (1968) precluded the possibility of further structural development, it was reasoned that the power of equilibration in the OMR group was such that they would give no evidence of structural learning.

Contrary to the predictions formalized in the hypotheses, it was found that: (a) the experimental induction procedures resulted in similar levels or correct judgements on the induction task on the part of the subjects of the three experimental groups and, (b) the degree of transfer, as measured by the number of correct judgements, was similar for all three experimental groups. Hence the present experiment did not support the postulate that the N, YMR and OMR groups were characterized by differences in either the mobility (field of application) or stability measures of equilibration.

Footnote

1. In comparisons using the Kruskal-Wallis one way analysis of variance the experimenter may, at his discretion, include a correction procedure for ties (Siegel, 1956). The effect of the correction for ties is to increase the value of H, consequently making it a less stringent test. In the present study it was decided not to use the correction procedure for two reasons: (a) The level of significance required was set at $\alpha = .10$, thus the probability of Type I error was quite high to start with. (b) The Kruskal-Wallis procedure was used in comparing differences in explanations and, except in a few instances, more than 50% of the subjects were tied at one rank (zero appropriate explanations). The correction is

$$1 - \frac{\Sigma T}{N^3-N}$$

where $T = t^3-t$ (when t is the number of tied observations in a tied group of scores).

N = number of all observations.

When t approaches N, the expression $\frac{\Sigma T}{N^3-N}$ approaches unity, thus resulting in a small denominator for the corrected value of H. For example, in comparing the number of appropriate explanations offered by the subjects in the 6 treatment x age groups on the conservation of weight test, the uncorrected value of $H = 3.76$ ($p>.50$, one tailed comparison). However, 78 of the 84 subjects offered no appropriate explanations; hence the correction factor $1 - \frac{\Sigma T}{N^3-N}$ was .1993. The corrected value of $H = 18.87$ ($p<.01$, one tailed comparison).

APPENDIX A

REVIEW OF THE LITERATURE

As Inhelder's (1968) work on cognition in the mentally retarded was the primary impetus for the present study, the present section is directed at examining (a) the relevant empirical literature, (b) some of the theoretical notions relating to the model and (c) some of the more important methodological issues which bear on the experimental validation of the model. The review is limited to studies involving at least one sample of MR subjects and those studies in which conrete operational functioning was the period of concern. In addition, only studies pertaining to cognitive structures (and the development of these structures) will be reviewed.

As is the case with Piagetian theory in general, the primary concern in Inhelder's model is the ontogenesis of cognitive structures (Piaget, 1970a, 1970b; Brainerd, 1972a). These structures (henceforth the term structure will have reference to cognitive structure, unless otherwise stated) are interposed between the functional invariants of organization and adaptation on the one hand, and the behavioural contents on the other. As such, structures are the main determinants of behaviour - as Flavell (1963, p. 169) states:

> "...! (structure) is that with which and into which he incorporates the data of the concrete problem before him."

Since structures are the main determinants of behaviour, given the appropriate stimulus situation one can determine the presence or absence of a structure from an examination of the subject's behaviour.

The literature pertaining to Inhelder's model is reviewed along the following dimensions: (a) evidence for concrete operational structures in the MR; (b) evidence bearing on the nature of these structures relative to both the descriptions of operational structures (eg., Piaget, 1970) and as compared to those of normals in experimental investigations; (c) evidence for a similarity in sequences of acquisition in normals and the MR; (d) evidence for a slower rate in the MR; and (e) evidence for arrestation in cognitive development on the part of the MR prior to the acquisiton of formal operational reasoning.

Presence or absence of cognitive structures

The behavioural data collected on the Piagetian inspired tests which can be used to infer the presence or absence of operational structures vary. On some tasks the subject's behaviour is sufficient, on others the subject is required to answer verbally posed questions, and on others the subject is required to answer questions concerning specific properties of materials in front of him. On the tests of conservation and transitive reasoning used in the present experiment, the subject was required to answer questions about various quantitative properties concerning sets of objects and then he was asked to justify his answer with an explanation. There is some controversy as to what the appropriate criterion should be for the inference of the presence of operational structures. The Genevans use the criterion of correct judgements and explanations reflecting the operation of one of the following cognitive operations: reciprocity reversibility, inversion reversibility, or quantitative identity (Piaget, 1968, pp. 36-37). Other experimenters often use correct judgements as the sole criterion (eg., McManis, 1969d, 1969c). In an attempt to solve the controversy over the appropriate criterion for inferring the presence of a cognitive structure, Brainerd (1972a) suggested that the criterion should be deduced from Piagetian theory as it relates to cognitive structures. On this basis Brainerd (1972a) contends that the inclusion of explanations as part of the criterion results in systematic error because it is at odds with the Genevan theoretical notion of the relationship existent between operational and linguistic structures. Piaget (1970a, 1970b, 1970c) marshalls considerable evidence to support the argument that intellectual operations give rise to linguistic progress and not vice versa. As Brainerd notes, from the perspective of Piagetian theory, the use of explanations as the criterion for the presence of operational structures introduces Type I error into the assessment situation. That is, if operational structures antedate linguistic structures and if we include linguistic structures as part of the criterion for the inference of cognitive structures, then some children who do, in fact, have the cognitive structure may not be so identified. Brainerd, then, is suggesting that correct judgements are the <u>minimum necessary</u> evidence required for the inference of the presence of operational structures. Presumably the Genevans require evidence of reversibility to avoid making Type II errors; employing judgements as the only criteria would increase the probability of making such an error. Type II errors, however, can also be reduced by judicious control of the stimulus situation (Brainerd, 1972a). It is the position of the present author that if the stimulus situation is well designed, and if there are no theoretical reasons to argue the contrary, then correct judgements are sufficient evidence for inferring the presence of a cognitive structure.

The nature of the Piagetian task in which we collect the behavioural data determines the level of cognitive structures attributed to the child. For example, successful performance on one of the conservation tasks involving a first order quantitative invariant such as substance and weight allows for the inference of concrete operational cognitive structures; success on a task involving a second order quantitative invariant such as volume, allows for the inference of formal operational reasoning.

Numerous studies have reported that a proportion of the population classified as being retarded by some criteria (eg., performance on a standardized test of intelligence or enrollment in a school for the mentally retarded child) are capable of making correct judgements in the Piagetian situations. Table A presents in summary form the results of these studies. In relation to this table, it will be noted that under criteria there are

INSERT TABLE A ABOUT HERE

three sub-sections. The "judgements" category includes those studies in which the S merely had to respond to one or more dichotomous questions. The "judgements and explanations" includes those studies where an appropriate explanation had to accompany the correct judgement. While this criterion is not the minimum necessary criterion, it is a sufficient one to make the inference of a cognitive structure. In the "other" category fall a variety of tests in which the Ss performance was taken to be sufficient to infer concrete-operational strucutres. An example of this would be the "multiple sorting" task used by Lovell, Mitchell and Everett (1962) in which the S needed to sort a set of materials three times - each time using a different dimension for the sort. Inspection of Table A reveals that, with few exceptions, a large proportion of retardates can meet the criteria for

TABLE A

Studies of Operational Functioning with the MR.

Study	Description of Subjects: CA (years)	MA (years)	IQ	Test	Percent of Subjects with Operational Thought
				Tests of Conservation	
Woodward 1961	-	-	25-79	substance a	11
Keasey & Charles 1967	13-29	5-11	-	substance b substance a	63 43
Stearns & Borkowski 1967	12-27	4-11	-	substance b substance b	42 37
McManis 1969a	7-21	5-11	47-73	substance a	51
McManis 1969d	9-18	5-8	$\bar{x}$ = 48.5	substance b	51
McManis 1969e	10-15	5-8	$\bar{x}$ = 56	substance b substance b substance b	77 40 42
Gruen & Vore 1972	8-12	5-9	60-78	substance b substance a	44 23
Lister 1970	8-16	-	46-80	substance a	60
Lister 1972	8-16	-	47-81	substance a	69
Langley et al. 1972	9-18	4-8	45-72	substance b	8
Brown 1973	$\bar{x}$ = 8	$\bar{x}$ = 6	-	substance b substance a	62 56
McManis 1969b	7-21	5-11	47-73	length a	60

TABLE A (CONT'D)

Study	Description of Subjects: CA (years)	MA (years)	IQ	Test	Percent of Subjects with Operational Thought
McManis 1970	$\bar{x} = 14$	$\bar{x} = 7$	$\bar{x} = 55$	length a	100
Lister 1972	8-16	-	46-80	length a	69
Woodward 1961	-	-	25-73	number b number a	23 16
Stearns & Borkowski 1969	12-18	4-11	-	number b	42
Gruen & Vore 1972	7-13	5-10	55-80	number b number a	57 43
Lister 1972	8-10	-	47-81	number b	87
Melnick 1973	7-15	4-10	59-100	number a	66
Brown 1973	$\bar{x} = 8$	$\bar{x} = 6$	-	number a	62
McManis 1969a	7-21	5-11	47-73	weight a	39
Lister 1969	13-16	-	42-79	weight a	48
Lister 1970	8-16	-	46-80	weight a	51
Lister 1972	8-16	-	47-81	weight a	50
Gruen & Vore 1972	7-13	5-10	55-80	weight b weight a	33 19

TABLE A (CONT'D)

Study	Description of Subjects: CA (years)	MA (years)	IQ	Test	Percent of Subjects with Operational Thought
Tests of Classification					
Lovell et al. 1962	9-15	-	ESN[c]	class inclusion class inclusion class inclusion class inclusion class inclusion	28 15 17 38 50
McManis 1968b	7-21	5-11	47-73	class inclusion	65
Prothro 1942	17-45	5-6	34-46	multiple classification	12-85
Lovell et al. 1962	9-15	-	ESN	bi-univocal multiplication of classes	30
Tests of Relational Thinking					
Prothro 1942	17-45	5-6	34-46	left-right test b	0
Lane & Kinder 1939	16	6-12	38-77	left-right test b	25
Lovell et al. 1962	9-15	-	ESN	seriation of length	52
Woodward 1961	-	-	25-73	seriation of length	10
McManis 1969b	7-21	5-11	47-73	transitivity of length	9
McManis 1969b	7-21	5-11	47-73	transitivity of weight	24
Lovell et al. 1962	9-15	-	ESN	bi-univocal multiplication of relations	30

TABLE A (CONT'D)

Study	Description of Subjects CA (years)	MA (years)	IQ	Test	Percent of Subjects with Operational Thought
				Tests of Spatial and Time Concepts	
Woodward 1962	-	-	25-73	copying spatial orders drawing reference points	28 1 6
Stearns & Borkowski 1969	-	4-11	-	reference points	3
Houssiadas & Brown 1967	8-15	-	55	mountain task	23
Montroy et al. 1971	7-16	5-11	53-79	time concepts	3-18
				Tests of Formal Operational Thought	
Jackson 1965	7-15		60-80	eight tasks of formal operational thinking a	0
Lister 1970	8-16	-	46-80	volume conservation a	34
Lister 1972	8-16	-	47-81	volume conservation a area conservation a	15 8
McManis 1969a	7-21	5-11	47-73	volume conservation a	9

a. criterion included judgements and explanations

b. criterion based on judgements only

c. ESN: educationally subnormal

the various tasks indicative of concrete operational functioning. The only exceptions were those instances in which the psychometric data indicated a very low level of functioning (eg., Woodward, 1961; Prothro, 1943).

The Nature of Cognitive Structures

Piaget (1970a, 1970b) has described the nature of a number of structures; at the concrete operational level they are described by the groupings or groupements (Piaget, 1970a). The specific nature of a cognitive structure can be discovered by analyzing the child's explanations of his judgements (Brainerd, 1972a). At least at the level of concrete operations, operational structure ontogenetically precedes linguistic structure - "the latter somehow growing out of the former to rely upon it subsequently" (Piaget, 1970b, p. 96). That is, Piaget sees language as being a dependent variable in relation to operational thinking, hence justifying the position that a study of the structural properties and semantic references of language can lead to the discovery of the properties of cognitive structures (Brainerd, 1972a; Flavell, 1963).

A problem arises, however, in that there seems to be little concensus on just what exactly constitutes an "adequate" explanation (Brainerd, 1972a). Piaget (1968) considers three categories of explanations to be adequate: reciprocity reversibility, inversion reversibility and quantitative identity. All three categories are consistent with the mathematical models Piaget describes as being isomorphic to the concrete operational cognitive structures. Many authors (eg., Lister, 1970, 1972; McManis, 1970) have determined on an a priori basis that only explanations is these categories will be considered adequate. There are several difficulties involved in deeming only these three categories of explanations as being adequate. First, a pragmatic reason, numerous authors mention many other categories as being adequate

Gruen and Vore, 1972; Keasey and Charles, 1967; Brainerd, 1972b). Second, language development in the retarded may be especially deficient as compared to cognitive development (Schiefelbusch, Copeland and Smith, 1967). Several investigators have, in fact, reported that verbal explanations appear to be inappropriate as indices of cognitive structures with the retarded (Woodward, 1961, Boehm, 1967). Given these difficulties in determining the adequacy of explanations, one solution would be to determine the adequacy of verbal explanations on an empirical basis as done by Brainerd (1972b) and Brainerd and Brainerd (1972). These authors inductively searched all the explanations given for mutually exclusive categories. The decision as to whether or not the category was an "adequate" one (i.e., was consistently related to the presence of a cognitive structure) was determined by noting what categories of explanations were employed by conservers (as indicated by a criterion only involving judgements).

Table B presents the results of three studies in which appropriate explanations were defined on an a priori basis to match the descriptions of operations set forth by Piaget (1968). The appropriate explanations were: reciprocity reversibility, inversion reversibility and identity. These studies, along with that of Inhelder (1968) give evidence that some MR's operational structures are isomorphic to the groupings described by Piaget.

INSERT TABLE B ABOUT HERE

Gruen and Vore (1972) compared the categories of explanations used by normal and mentally retarded conservers. The normal and retarded groups were matched on MA levels of 5, 7 and 9 years. Conservation status was determined with reference to a judgement based criterion; for the purpose

TABLE B

Studies with MR Subjects Employing Explanation Categories Acceptable to Piaget (1968).

Study	Conservation Test	Percent of Subjects Satisfying Criterion
Lister 1972a	Number	87
	Substance	69
	Length	66
	Weight	50
	Volume	16
	Area	8
Lister 1970b	Substance	58
	Weight	42
	Volume	34
McManis 1969b[c]	Length	58
	Weight	58

Note. - In all three studies correct judgements also were part of the criterion.

a. Reciprocity and inversion reversibility and identity explanations were defined as appropriate.

b. Reciprocity and inversion reversibility, identity and logical necessity were defined as appropriate. Logical necessity was scored when the subject explained conservation by having reference to the equivalence of the stimuli before the physical transformation.

c. Reciprocity and inversion reversibility were defined as appropriate.

of the present analysis a subject was judged to be a conserver if he made correct judgements on a conservation of number and/or substance and/or weight tests. Gruen and Vore compared the use of the normals and mentally retarded subjects on the following "logical" categories: (a) previous events, (b) logical necessity, (c) counting, (d) inversion reversibility, (e) reciprocity reversibility and (f) addition and subtraction (quantitative identity). When the two groups at each MA level were compared it was found that normals and retardates did not differ in the number of "logical" explanations used nor in the frequency with which each category was used. An analysis of the explanations used by the normal and mentally retarded conservers in the present experiment is presented in Appendix B. With reference to the nature of cognitive structures in normal and retarded populations, it appears that at least some normals and retardates use explanations reflecting the operations cited by Piaget (eg., 1970a) and that normals and retardates, when matched on MA, use similar explanation categories with the same frequency (Gruen and Vore, 1972).

Common Operational Construction

In the present section the evidence for an invariant sequence in the acquisition of substance, weight and volume concepts common to both normal and retarded populations will be considered. The purpose of the review is to examine the postulate that a similar operational construction occurs in those of normal and deficient intellectual functioning (Piaget and Inhelder, 1947; Inhelder, 1968). Only the relationship in the appearance of substance, weight and volume is considered because it is the only sequence sufficiently investigated with the retarded and/or because the other invariant sequences studied are of limited interest in that they are probably due to differences in the stimuli used to test the concepts (Flavell, 1970). Due to the numerous methodological difficulties encountered in demonstrating invariant sequences (cf., Flavell, 1970; Pinard and Laurendeau, 1969), the review will also be limited to studies in which the tests of interest were administered to the same subjects; this restriction is meant to minimize the variation due to procedural differences. Lastly, only those studies including at least one sample of mentally retarded subjects were included in the review. The studies to be reviewed are grouped in two categories: (a) studies in which the performance on the concept tests was compared on a group basis, and (b) studies in which each individual subject's test response pattern was analyzed. The latter analysis is more sensitive to investigating sequences in development since group averages may well obscure sequences occurring in the response patterns of the individuals and the group analysis is susceptible to floor (when all the tests are too difficult for the subjects) and ceiling (when all the tests are too easy for the subjects) effects.

Group performances. McManis (1969a) compared the performance of normal and MR subjects on tests of substance, weight and volume conservation. The subjects were matched at MA levels of 5, 6, 7, 8, 9, 10-11 years. The score on each tests was comprised of the subject's response to a judgement question before the physical transformation, a judgement question after the physical transformation and his explanation for the latter judgement. The overall analysis, collapsing across the various MA groups's supported the postulate of an invariant sequence in that significantly higher scores were obtained on the substance than weight test and on the weight than volume test by subjects in both the normal and MR samples. A detailed inspection of the performances of the groups at each MA level revealed the only evidence contrary to the predicted sequence occurred in both groups for subjects with a mental age of 5 years. These subjects obtained higher scores on the conservation of volume test than the substance or weight tests. McManis (1969a) explained this anomalous result by suggesting that the prediction response was artificially inflated because these subjects did not comprehend the verbal description of the physical transformation to be enacted. While McManis administered tests of all three concepts, Gruen and Vore (1972) administered only tests related to the substance and weight concepts. Their sample was comprised of MR and normal subjects at MA levels of 5, 7, and 9 years. The scores on the two tests were compared using two criteria: (a) judgements, and (b) judgements accompanied by "logical" explanations. When performance on the two tasks was compared using judgements only there were no significant differences at any MA level for either of the groups. When the performance on the two tasks was compared using the judgement plus explanation criterion, the only significant difference occurred at the MA 9 level; for both the MR and normal groups, the scores on the substance task were significantly superior to those obtained on the weight task. In an experiment comparing the performance of 24 educable mentally retarded subjects on the conservation of substance, weight and volume tests, Carlson and Michaelson (1973) found no significant difference attributable to the tests.

In general, the group data provided meager evidence to support the hypothesized invariant sequence in the acquisition of substance, weight and volume conservation. As mentioned, however, this may well have been a function of the samples selected in that floor and ceiling effects may have occurred.

<u>Individual performances.</u> Two studies (Hunter and Lister, 1970; Lister, 1972) provide an analysis of individual protocols of mentally retarded subjects administered tests of substance, weight and displacement volume conservation. The criterion for conservation in both studies involved correct judgements plus explanations. Hunter and Lister (1970) report that of 21 subjects who attained conservation status on one of the concepts, all did so for the conservation of substance task. Also, of the 40 subjects attaining conservation status on two of the tasks, all did so on the conservation of substance and weight tasks. Of the 61 subjects showing discrepant performances on the three tasks, then, all 61 supported the hypothesis that the three conservation tasks are ordinated in the predicted sequence. Lister (1972) found that of the 24 subjects showing discrepant performances of the substance, weight and displacement volume, the performance of 21 were in the predicted sequence. Of the three subjects failing to demonstrate the predicted sequence, one was a conserver of volume but not weight and substance, one was a conserver of substance and volume but now weight, and one was a conserver of weight and volume, but not substance.

In summary, while the three studies (McManis, 1969a; Gruen and Vore, 1972; Carlson and Michaelson, 1973) comparing the performances of groups failed to provide unequivocal support for the invariant sequence in the acquisition of substance, weight and volume, the two studies (Lister, 1972; Hunter and Lister, 1970) examining the individual response patterns did provide substantial support for the posited sequence in acquisition.

The evidence for a similar sequence in the acquisition of these concepts does not conclusively point to a similar mechanism of structural development - it merely makes such a position more plausible. Several alternative interpretations of the temporal relationship in acquisition are feasible (Flavell, 1970). First, the later acquired responses may only substitute for the earlier acquired responses. If such is the case then similarities in the social environments, rather than operational mechanisms, are implicated. Second, the structures requisite for the acquisition of the concept may be linked by similarities. Thus the later acquired concept may only reflect a perfected version of the structure necessary for the acquisition of the earlier concept. Or else, the structure for the earlier acquired concept may be integrated in the later structure for the later acquired concept. If the latter is the case, then this would constitute evidence for a similar operatory construction. Third, the earlier acquired concept may serve as a mediator for the later acquired concept; if such is the case, then the invariant sequence may or may not support similar operatory construction. Fourth, there is the possibility that the structures may have been acquired by different means even if the sequence in concept acquisition is the same (cf. Werner, 1957; Waddington, 1957; and Wolff, 1972). Fifth, a criticism directed specifically at the sequence examined, the decalage observed in the acquisition of the invariance of substance, weight and volume is of a "trivial" nature; consequently the foregoing analysis is not a suitable test for the postulate that normal and MR populations demonstrate a similar mechanism for operatory construction (Brainerd, 1974, personal communication).

Brainerd differentiates between decalages in acquisition which are of a "trivial" nature (such as the decalage occurring in the acquisition of substance, weight and volume) and those of a "non-trivial" nature (such as the decalage occurring in the transitivity of substance and the conservation of substance). Decalages are trivial when there is no logical alternative to the sequence of acquisition; for example, Brainerd suggests that substance must antedate weight because logically the acquisition of weight implies the acquisition of substance. Hence, according to Brainerd, only the examination of non-trivial decalages relate to the postulate of similar operatory construction in normal and MR populations. As noted previously, however, evidence relating to non-trivial sequences in both normal and MR populations is not available or else confounded by methodological differences. In summary, while there is evidence suggesting that there is an invariant sequence in the acquisition of the conservation of substance, weight and volume in both normal and MR populations, the deduction that this implicates similar operatory construction in the two populations is unwarranted.

Rate of Structural Development

The three previous sections have indicated that some proportion of the retardate samples assessed in various normative-descriptive studies have cognitive structures which are isomorphic to those of normals in that retardates (a) can make correct judgements in many Piagetian situations and (b) can employ explanations which fall in the three categories described as "adequate" by Piaget (1968), and (c) demonstrate invariance of the acquisition of substance, weight and volume conservation as do those of normal intelligence. In the present section we shall discuss a difference between normals and retardates, viz., the rate of structural succession.

Normative studies support Piaget's (1970a) contention that the ability to provide correct judgements and adequate explanations on Piagetian tasks is attained by most Ss within a restricted age range of three to four years (Inhelder, 1968). This lattitude in age range has been interpreted by some authors (eg., Brainerd, 1972b) to indicate the age of the S to be of little importance except as a "statistical associate of stage". While this may be the case for children of normal intelligence, Piaget does postulate that age (as an index of rate of development) is a relevant variable:

> "... For a specific subject the speed of transition from one stage to the following one has an optimal rate. That is, the stability and even the fruitfulness of a new organization (or structurization) depends on connections which cannot be instantaneous but cannot be indefinitely postponed either since they would lose their power of internal combination" (1970a, p. 713; cf. Piaget, 1973).

"The power of internal combination" Piaget attributes to structures in the above passage clearly has reference to the process of equilibration or the "self-regulating" characteristic of structures (cf., Piaget, 1970b).

In examining the differential rate of succession of structures in normals and retardates, some caution has to be exercised. The S's performance on a Piagetian task is influenced by several factors such as the specific materials used (Piaget, 1971; Inhelder, 1968), the criterion employed (Brainerd, 1972a; Gruen and Vore, 1972), and various aspects of the methodology (Rothenberg, 1969; Rothenberg and Courtney, 1969). For this reason, the review is limited to those studies in which retardates and normals were tested using the same materials, procedures, etc. Figure A graphically illustrates the delay in development associated with retardation as found in a paradigmatic study of rate of development. Gruen and Vore (1972) compared the number of

INSERT FIGURE A ABOUT HERE

correct judgements made on the conservation of number, substance and weight tasks by groups of normals and retardates matched on chronological

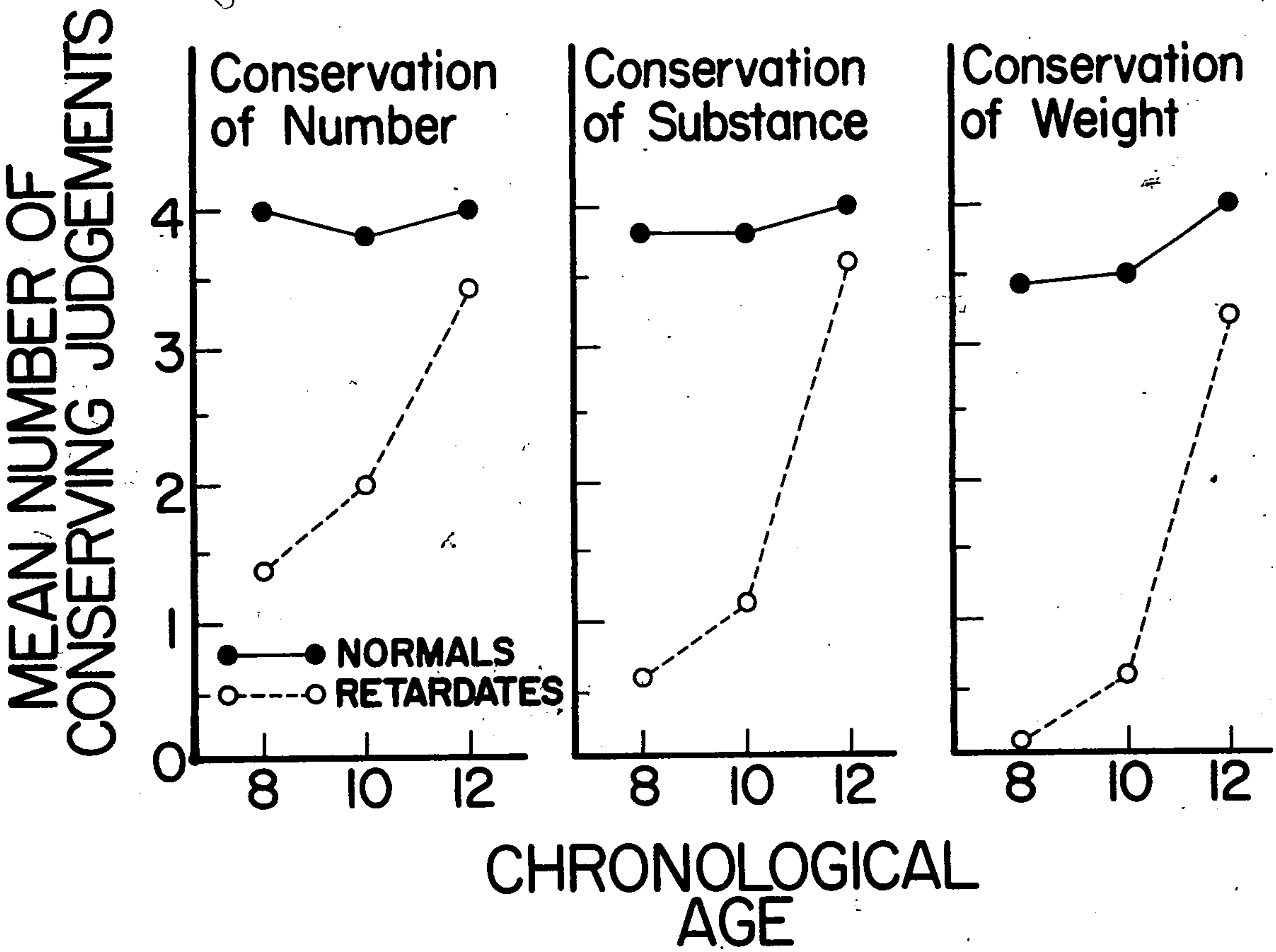

Figure A. Judgements on Conservation Tests as a Function of Intelligence Level and Chronological Age (After Gruen and Vore, 1972).

age. On all three conservation tasks the 8 and 10 year old normal group's performance was superior to that of the same aged retardate group; the two oldest age groups performed at the same level for all three tasks. The outcome of comparisons involving CA matched individuals of normal and deficient intellectual functioning is almost axiomatic when the comparison involves performance on a cognitive task of a developmental nature (Zigler, 1969). This is so much the case that few investigators include CA comparison groups when normal and MR subjects have been matched on CA; the performance of the former has been shown to be superior on tests of number conservation (Gruen and Vore, 1972; Brown, 1973), substance conservation (Gruen and Vore, 1972; Brown, 1973), weight conservation (Gruen and Vore, 1972) and various tests of classificational and relational thinking (Lovell, Mitchell and Everett, 1962).

Another manner of indicating the delay in rate of construction is to examine the normative data for studies including both normal and retardate samples. Table C reports the chronological age at which 70 to 75 percent of a normal and MR subject sample attained the criterion for concrete operational

INSERT TABLE C ABOUT HERE

thought. Again the evidence is in support of the observation that retardation involves a delay in the rate of structural development.

In summary, the evidence for a delay in structural development on the part of the MR, relative to similar chronological aged subjects of normal intellectual status, is conclusive.

Arrestation in cognitive development.

The second difference in structural development existing between those of normal and retarded intelligence is that the retarded become arrested in

TABLE C

Normative Age for the Acquisition of Concrete Operational Skills by Normal and MR Subjects

Study	Test	Chronological Age (in years) Normals	Retardates
Gruen and Vore 1972	Number Conservation a	9	12
	Substance Conservation a	9	12
	Weight Conservation a	9	12
	Number Conservation b	9	12
	Substance Conservation b	9	12
	Weight Conservation b	9	--
Lovell et al. 1962	Addition of Classes	8	15
	Multiplication of Classes	8	15
	Seriation	7	13
	Multiplication of Assymetrical Relations	8	
	Class Inclusion	9-11	-
	Multiple Classification	6-7	10-13
McManis 1969a	Substance Conservation	8	15
	Weight Conservation	10-11	17
McManis 1970	Length Conservation	6	14
	Seriation of Length	7	15
	Transitivity of Length	10-11	18

Note. - Chronological age level represents the age at which 70 percent of the subjects in the sample met the criterion for concrete operational functioning or the age at which 70 percent of the responses were of a concrete operational nature.

a. Criterion based on judgements only

b. Criterion based on judgements and explanations structural development prior to the acquisition of formal operational thought (Inhelder, 1968).

Theoretically, Piaget (1970a) has suggested that in the course of development, the various factors are so interrelated that acquisition of formal operational structures is the normal outcome. Only if this stage is attained can we talk about perfect equilibration for many cognitive activities. In the discussion on the rate of succession of structures, we also saw that Piaget considered as a plausible hypothesis that "new organization ... cannot be indefinitely postponed either since they would then lose their power of internal combination (Piaget, 1970a, p. 713). If retardation involves a delay in the succession of structures, then it would be expected that retardation may be characterized by a cessation in development before formal operational thinking was achieved and that, depending on the degree of retardation, other levels of structurization may also not be attained. Inhelder's (1968) classification system for mental retardation, in fact, is based on this concept of arrestation in development. Both Piaget and Inhelder have reference to the concept of equilibration to explain this arrestation in development. While Piaget talks about a loss in "the power of internal combination", Inhelder has reference to "false equilibration". In contra-distinction to Piaget, Inhelder (1968, p. 161) ties the arrestation in development closer to the chronological age of the organisms. She suggests that after the ages of 13 to 15, it is highly improbably that any development will occur but that it may well be possible before that age. It should be noted, however, that the 13 to 15 years age span may well be a function of Inhelder's selection procedures in that her sample was limited to preoperational and concrete operational children. Had Inhelder included profoundly and severely retarded children and adults in her sample, arrestation of development for these individuals may have been postulated to occur earlier in their life span.

The research literature is suggestive of the following plausible postulate: that arrestation of development is possible at any substage in any period of development. Woodward (1959) and Woodward and Stern (1963) report that many profoundly retarded children and adults are incapable of even the last stage of sensorimotor functioning. Similarly, Inhelder (1968) found a proportion of her sample of retardates to be incapable of concrete operational functioning. Further support for a position stating that some retardates are incapable of concrete operational thought comes from the studies employing substance and number conservation tasks reported in Table A. While the table reports the percentage of Ss who did meet the criterion of concrete-operational thought, it also indicates that a percentage of the subjects did not meet the criterion. The proposition that the development of the subjects included in the above studies had actually been arrested is difficult to test since further development may always have been forthcoming. Only if we accept the proposition that further development is unlikely after the age of 13 to 15 can we test the hypothesis; this would be done by selecting Ss considerably older than the designated age span. We can, however, empirically determine the validity of Inhelder's (1968) contention that "to be retarded means ... to be able to think by concrete operations but not by formal operations" (p. 294). In the empirical test of this statement, if one found retardates to be capable of formal operational thinking then the hypothesis would be demonstrated to be invalid. If, however, one found an absence of formal operational thought, then the validity of the statement would still be in doubt. As it turns out, the literature is equivocal: Jackson (1965) assessed the performance of 40 normals and 40 educable retardates on six tests selected from Inhelder and Piaget's (1958) book on the growth of formal operational structures. While 8 to 40% of the normal subjects demonstrated formal operational thinking, not a single subject in the retarded sample gave indications of this level of thought. The remaining studies assessing formal operational functioning in retardates all involved the use of the conservation of volume and/or area. McManis (1969a) assessed normal and retarded subjects in the mental age range of 5 to 11 years. McManis' procedure involved a prediction judgement, a transformation judgement and an explanation. In order to be credited with formal operational thinking, the subjects had to answer all three conservation of volume questions correctly. Of the 90 normal subjects, three percent answered all three questions correctly; of the 90 retarded subjects, nine percent answered all three questions correctly. Two problems exist in the interpretation of McManis' results: (a) he did not state what the criterion were for adequate explanations and (b) it is difficult to state whether the nine percent of retardates demonstrating volume conservation was significant relative to the number of subjects expected to answer the two judgements questions correctly just by chance. Lister (1970) tested 104 educable retardates on a conservation of both interior and displacement volume. The criterion for formal operational thought included both correct judgements and explanations involving reciprocity reversibility, inversion reversibility or quantitative identity. Thirty-four percent of the subjects met the criteria on both the interior and displacement volume tasks while another eight percent met the criteria on one or the other tasks. Lister (1972), employing the same criterion as in the previous study, reported that 16 percent of 115 educable retarded children showed formal operational thought on both of the volume tasks (another eight percent met the criterion on one or the other tasks) and that eight percent of the Ss met the criterion on complementary and interior area tasks (another eight percent met the criterion on one, but not both, area tasks). Thus, at the present time, the question of whether or not retardates are capable of formal operational thinking must be answered equivocally: the tests of formal operational reasoning used seems to be a major variable needing further investigation before this question can be answered.

One further question remains: Can formal operational thought be induced by exposing the individual to relevant experiences? Lister (1969) selected 12 educable retardates who gave non-conservation judgements on a weight task to participate in an induction of conservation of weight experiment. All subjects gave correct judgements and adequate explanations on conservation of weight posttests administered immediately after, two weeks after and four weeks after training. Eight months after the training, Lister assessed the eleven available subjects not only on conservation of weight but also on conservation of interior and displacement volume. All eleven subjects made correct judgements and gave adequate explanations on all tests. There is no assurance that the induction procedure was responsible for the performance on the conservation of volume tests in that (a) there was no control for increased task familiarity, (b) conservation of volume tests were not included in the pretest, or (c) other experiences in the eight month period may have been responsible; it is evident that some type of acquisition had taken place. The major problem, from the standpoint of interpreting the results, is to determine whether a true logico-mathematical learning had taken place or whether psuedo-learning had occurred. It may be that the induction techniques were too specific in providing correct answers and adequate explanations. The conservation of volume tests employed were very similar in nature to the specific situations included in the training procedures; hence they were actually inadequate to test for generalization for the purposes of distinguishing between pseudo-and logico-mathematical learning. The explanations offered by the subject to explain conservation of volume were primarily in the form of "it is still the same" plasticine, etc. This form of the identity explanation is qualitative, not quantitative (Piaget, 1968); thus, a large proportion of the children were giving preoperational explanations for tests at the formal operational level.

Lister (1970) induced volume conservation in 30 EMR children. On the pretest, all 30 subjects failed to make conservation judgements in both the interior and displacement volume situations. Of those 30 subjects, 16 failed to make conservation of substance and weight judgements, 6 made conservation of substance judgements but did not make conservation of weight judgements, and 8 made conservation of weight and substance judgements. The training procedures directed at inducing volume conservation included:

> "... a combination of manipulating reversibility, filling containers with and counting discontinuous material, measuring displacement and verbalizing reasons for conservation (including reversibility, identity, and compensation) and empirical findings (such as the distinction between appearance and reality, and the results of measurement)" (Lister, 1970, p. 59 - material within parentheses is part of the text).

Posttests were administered 1 week, 2 weeks, 6 weeks (on volume conservation only) and 6 months after the original training was completed. In addition to replicating the tests given in the pretest, the posttest also included identical tests which used different materials. Also, while the pretest situation involved making judgements of equality only, the posttests involved making judgements of inequality. The criteria for conservation in all three content areas required the subject to make the correct judgements and to be able to justify the judgement with an explanation reflecting either reversibility or identity. On the posttests assessing volume conservation, all subjects met the criteria for conservation. All but one subject met the criteria for substance and weight conservation on all posttests; the one subject, while making the correct judgements in both situations, failed to give acceptable explanations. Most of the explanations offered by the subjects reflected identity, some reflected inversion reversibility and fewer still reflected reciprocity reversibility.

Essentially the same comments can be made of Lister's (1970) study as were made of the other two training studies reviewed. The experimental design did not include a control for the effect of task familiarization. Also, it is questionable whether actual operative development took place. In relation to this last point, Lister (1970) presents three main lines of evidence intended to support the conclusion that operational thinking had been facilitated or initiated as a result of the induction procedures: (a) the posttest situations included tests with materials with which the S had not been specifically trained. (b) During the training period, judgements of equality had been emphasized while the posttests required both judgements of inequality and equality. (c) Although all the explanations reflected the operation of identity, inversion reversibility and reciprocity reversibility - the explanations verbalized by the E during training - the vocabulary used by the subjects in their explanations was often very different than that used by the experimenter; furthermore, some subjects occasionally demonstrated the explanation. All three of these lines of evidence, however, fall short of providing adequate proof. With respect to (a), even though the materials may have been somewhat different, the two situations were still essentially similar with regard to procedures, etc. As for (b), while it eliminated the possibility of a subject showing conservation as a function of a response set (always judging equality), it does not eliminate the possibility that the subject merely perseverated the initial judgement made by the experimenter in each situation. For example, in conservation of weight tests, the experimenter typically shows the subjects two stimuli and states whether or not they are equivalent. Lister's subjects merely needed to be aware of the experimenter's comments in relation to this initial comparison in order to be able to answer all the questions correctly. With regard to point (c), most of the subjects' explanations referred to the fact that it was still the same material. As previously noted, this would seem to be indicative of qualitative, as opposed to quantitative, identity; qualitative identity is present in the thought of the preoperational child (Piaget, 1968; DeVries, 1969).

Lister (1972) pretested 115 educable retardates on conservation tests relating to the concepts of number, substance, distance, length, weight, volume, and area. Fifty-one of these subjects, all of whom demonstrated non-conservation on at least two of the concepts, participated in the induction phase of the experiment. These subjects were divided into three groups; in each group two subjects were non-conservers on all of the pretests, two more demonstrated number, substance and length conservation and nine failed the area and volume tests only. One group was trained on conservation on all concepts, the second group received training on the conservation of area only, while the third group was given practice in reading. The induction technique for each concept trained included (a) the giving of examples of acceptable explanations, (b) verbalization of the reasons for conservation in conjunction with observation of materials, (c) the use of external criteria, and (d) the object-bound form of inversion reversibility. The posttests, administered one week and two months after training, included tests of each concept with new materials and tests of inequality for each concept in addition to the tests of the pretest. Considering the basic procedural similarity in all conservation tasks, however, it is doubtful whether these additions are sufficient to provide a true test of the generality of the acquisition (Inhelder and Sinclair, 1969). This question is of special importance, since the induction procedures were very detailed and explicit; hence raising the issue of whether psuedo-learning or logico-mathematical learning had taken place. Fifteen of the 17 subjects in both experimental groups demonstrated conservation responding in every situation on both posttests. The remaining 4 subjects in the two experimental groups failed to give any indication of conservation on any of the posttest items; these were the same 4 subjects who demonstrated total non-conservation on the pretests. Finally, not a single subject in the control group manifested an increment in conservation performance.

In conclusion, all three investigations attempting to induce further progress in operational development were successful in inducing conservation of volume (Lister, 1969, 1970, 1972). However, because the intervention procedures in all three studies included explicit training on correct judgements and adequate explanations, it is debatable whether or not formal operational thought had been inculcated. As was the case with our review of the normative-descriptive studies, the evidence demonstrating formal operational thinking in the MR is inconclusive. At the same time, however, the evidence for arrestation at periods more primitive than formal operational thinking was substantial.

APPENDIX B

SUPPLEMENTARY ANALYSES

In the present section four additional analyses will be presented: the first analysis is a further test of the experimental hypotheses; the last three analyses were provoked by the review of the literature and/or test the validity of some of the assumptions made in the design of the experiment.

Subjects trained during the induction phase

The analysis of the judgement data collected during the induction phase indicated that the training procedures were differentially effective in leading to correct judgements for subjects in the experimental groups. Only 8, 5 and 6 subjects of the N, YMR, and OMR experimental groups respectively gave correct judgements to all three questions on two or more of the training trials. It can be argued that generalization to the operational tests on the posttests would only be expected from those subjects demonstrating correct judgements during the induction phase. Consequently, the following analyses compare the posttest performance of the experimental subjects giving three correct judgements on at least two trials during the induction phase (henceforth these subjects will be called trained subjects). The proportion of trained subjects manifesting operational performance (Stage II or Stage III) on the posttests is reported in Table D. With reference to Table D, it can

INSERT TABLE D ABOUT HERE

be seen that only two comparisons attained significance.

The superiority of the YMR trained subjects over the OMR trained subjects ($p = .06$) on the substance conservation tests administered on posttest II was in the direction predicted by the hypothesis. However, since the performance of the N trained subjects was not significantly different from that of the YMR or OMR groups, the hypothesis that age would be inversely related to the power of equilibration is not confirmed. The second significant

TABLE D

Proportion of Experimental Subjects at Stage II or III on Posttests

Test Day

Conservation of Substance

Posttest I	N	YMR	OMR	Posttest II	N	YMR	OMR
N	.13			N	.13		
YMR		.20		YMR		.60	
OMR			.17	OMR		*	.00

Transititivity of Substance and Conservation of Number

Posttest I	N	YMR	OMR	Posttest II	N	YMR	OMR
N	.63			N	.75		
YMR		.40		YMR		.40	
OMR			.17	OMR	**		.00

Conservation of Weight

Posttest I	N	YMR	OMR	Posttest II	N	YMR	OMR
N	.13			N	.13		
YMR		.00		YMR		.00	
OMR			.00	OMR			.00

Note. - Along the diagonal of each matrix the proportion of subjects at Stage II or III in each group is reported at the intersection point of that group. If the comparison of two different age groups was significant in the predicted direction, the asterisks for significance were placed at the point of intersection of these two groups below the diagonal.

* p .10   *** p .01

comparison, the superiority of the N trained subjects over the OMR trained subjects on the combined conservation of number and transitivity of substance tasks of posttest II, again was in the predicted direction. As such, the results of the preceeding analysis with trained subjects supports the conclusions of the analysis based on the total group of experimental subjects which also indicated that the N group was superior to the OMR group on the combined transitivity of substance and number conservation tests of posttest II. The analysis of the performance by the total groups on these two tests also showed significant differences on posttest I: the N experimental group's performance was superior to the performance of both the YMR and OMR groups. The failure to find a similar pattern of significant differences in the analysis of trained subjects provides further support for the suggestion that the training procedures were not solely responsible for the gains made on the posttests.

In comparing the number of appropriate explanations offered by the trained subjects in each group it was found that no difference existed on any block of tests on any of the test days. In comparing the number of appropriate explanations offered on the substance conservation tests, there were no significant differences among the three groups on the pretest ($H = 3.35$, $p > .10$), on Posttest I ($H = 2.58$, $p > .20$) or on Posttest II ($H = 3.35$, $p > .10$). On the transitivity of substance and conservation of number tests, the performance of the three groups did not differ on the pretest ($H = 2.26$, $p > .30$), Posttest I ($H = 1.22$, $p > .50$) or on Posttest II ($H = 0.54$, $p > .70$). The finding that the three groups of trained subjects did not differ significantly in the explanations offered is congruent with the results of the comparisons involving the total group of subjects. As such, the same conclusion is valid, viz., that the training procedures did not differentially effect the explanations offered on the posttests.

Correspondence between correct judgements and explanation categories

As was mentioned in the review of the literature, difficulties are encountered in defining just what categories of explanations should be called appropriate. Since explanations are dependent variables of the nature of the cognitive structure (Brainerd, 1972a), the procedure used to define appropriate categories is of some importance. In the present experiment, appropriate categories of explanations were determined by assessing the degree of correspondence between the explanation category and correct judgements. The validity of this procedure is based on the postulate that correct judgements in the various Piagetian tests are sufficient evidence for the presence of an operational structure.

The data used in the definition of appropriate categories were the pretest judgements and explanations of all experimental and control group subjects and all subjects demonstrating stage II or stage III functioning who were administered the complete pretest. The following procedure was used in the determination of appropriate explanation categories: (a) The number of explanations accompanied by correct or incorrect judgements was calculated for each category of explanations. A separate analysis was conducted for each of the judgement questions on each concrete operational test. (b) The number of explanations accompanied by correct judgements was contrasted to the total number of explanations using the Binomial Test (Siegel, 1956).As the judgement questions were of a dichotomous nature, the value of P and Q was set at .50 in the expression:

$$p(x) = \frac{N!}{x!\ (N-x)!} \qquad P^x\ Q^{N-x}$$

where: p(x) is the probability of obtaining the observed number of correct judgements;

x is the number of explanations accompanied by correct judgements;

N is the total number (correct and incorrect) of explanations;

P is the proportion of explanations that would be associated with correct judgements by chance;

Q is the proportion of explanations that would be associated with incorrect judgements by chance.

Inspection of Table E, which reports the proportion of explanations

INSERT TABLE E ABOUT HERE

associated with correct judgements for each of the judgement questions, indicates that neither a logical criterion alone nor a statistical criterion alone can adequately define what categories of explanation should be considered appropriate. A statistical criterion would probably lead to the conclusion that the no perceptual deformation category should be considered as one of the appropriate categories. This category was associated with correct judgements 78 percent of the time it was used on the conservation tests; furthermore on the five occasions that the explanation was used, a sufficient number of times for a statistical comparison, three of the analyses indicated it was associated with correct responding above a chance level ($p < .10$, two failed comparisons). The definition of the category (refer to Table 2) suggests that the high correspondence with correct responding is due to the fact that the subject believes that the physical transformation will not lead to a perceptual change; this is supported by the fact that this explanation was used almost exclusively on the prediction phases of the tests, not on the transformation phases. Clearly this

TABLE E

Correspondence Between Correct Judgements and Categories of Explanations

Test	Proportion of Explanations Associated with Correct Judgements						Total	Proportion Correct
	Prediction Phase			Transformation Phase				
	same	more	less	same	more	less		
Comparison								
Transitivity				1.00***	.90***	.80***	54	.89
Logical Necessity								
Transitivity							8	1.00
Number		.86	.56		1.00*	1.00*	27	.82
Substance 1							5	1.00
Substance 2							6	1.00
Substance 3		.72					15	.73
Substance 4							9	.89
Substance 5							10	.90
Weight					.80		13	.77
Inversion Reversibility								
Transitivity								
Number	1.00**	1.00**	.50	1.00**			33	.85
Substance 1	1.00***	.88*	1.00**	.94***	1.00*	.67	52	.92
Substance 2	1.00***	1.00**	1.00***	1.00***			39	1.00
Substance 3	.89***	.75	.70	1.00***	1.00*	.89**	63	.87
Substance 4	1.00***	1.00**	1.00**	1.00***			35	1.00
Substance 5	1.00***	.89**	.86	1.00***	1.00**	.88*	62	.95
Weight	1.00***		.60	.91***			32	.84

TABLE E (CONT'D)

Correspondence Between Correct Judgements and Categories of Explanations

Test	Proportion of Explanations Associated with Correct Judgements						Total	Proportion Correct
	Prediction Phase			Transformation Phase				
	same	more	less	same	more	less		
Identity								
Transitivity							4	.00
Number	1.00***	.88*	.89**	.80	.80	.60	40	.85
Substance 1	1.00*						19	1.00
Substance 2		1.00*		1.00*			20	.95
Substance 3	1.00**			1.00*			19	.95
Substance 4							13	1.00
Substance 5	1.00***	1.00*	1.00*	.80	1.00*	1.00*	34	.97
Weight	1.00**						16	1.00
Second Invariant								
Transitivity							0	
Number							0	
Substance 1							1	1.00
Substance 2							2	1.00
Substance 3							0	
Substance 4							1	1.00
Substance 5							1	1.00
Weight							3	.33
Counting								
Number	.95***	.67	.60	.84***	.52	.55	117	.70

TABLE E (CONT'D)

Correspondence Between Correct Judgements and Categories of Explanations

Test	Proportion of Explanations Associated with Correct Judgements						Total	Proportion Correct
	Prediction Phase			Transformation Phase				
	same	more	less	same	more	less		
No Perceptual Deformation								
Transitivity							1	.00
Number							0	
Substance 1	.88*						15	.73
Substance 2	1.00***	.40	.75				19	.79
Substance 3							2	.50
Substance 4	.88*						12	.83
Substance 5							0	
Weight							1	1.00
Logical Necessity Negated								
Transitivity					.20	.17	16	.23
Number							1	.00
Substance 1							4	.25
Substance 2							5	.00
Substance 3							5	.20
Substance 4							3	.00
Substance 5							5	.20
Weight							1	.00

TABLE E (CONT'D)

Correspondence Between Correct Judgements and Categories of Explanations

Test	Prediction Phase same	Prediction Phase more	Prediction Phase less	Transformation Phase same	Transformation Phase more	Transformation Phase less	Total	Proportion Correct
Inversion Reversibility Negated								
Transitivity							1	1.00
Number							0	
Substance 1							1	1.00
Substance 2							6	.00
Substance 3							5	.00
Substance 4							0	
Substance 5							1	.00
Weight							1	.00
One to One Correspondence								
Number							14	.64
One to One Correspondence Destroyed								
Number							13	.08
Second Invariant Negated								
Trnaisitivity							0	
Number							0	
Substance 1							2	.00
Substance 2							2	.00
Substance 3							0	
Substance 4							0	
Substance 5							0	
Weight							6	.17

TABLE E (CONT'D)

Correspondence Between Correct Judgements and Categories of Explanations

Test	Proportion of Explanations Associated with Correct Judgements						Total	Proportion Correct
	Prediction Phase			Transformation Phase				
	same	more	less	same	more	less		
Reciprocity Reversibility								
Transitivity							0	
Number							0	
Substance 1							11	.46
Substance 2	.17				.43	.20	28	.25
Substance 3							13	.53
Substance 4				.40		.60	24	.29
Substance 5							4	.25
Weight							1	.00
Tautology								
Transitivity				.05***	.80	.60	39	.33
Number	.50	.00***		.57	.00***	.00*	50	.22
Substance 1	.86	.00**		.60	.00**		30	.30
Substance 2	.78	.00**	.00*	.83	.00*		34	.35
Substance 3	.71	.17**	.00**	.66	.00**	.00*	43	.26
Substance 4	.86	.00***			.00***		34	.18
Substance 5	.63	.00***		.50	.00***		41	.20
Weight	.75	.05***	.09***	.57	.00***		67	.18

TABLE E (CONT'D)

Correspondence Between Correct Judgements and Categories of Explanations

Test	Proportion of Explanations Associated with Correct Judgements						Total	Proportion Correct
	Prediction Phase			Transformation Phase				
	same	more	less	same	more	less		
				Perception				
Transitivity				.60	.60	.47	193	.55
Number	.27**	.13***	.13***	.24*	.05***	.00***	235	.12
Substance 1	.24***	.11***	.12***	.16***	.02***	.02***	511	.10
Substance 2	.28***	.16***	.05***	.16***	.03***	.01***	487	.10
Substance 3	.43***	.09***	.06***	.16***	.04***	.00***	462	.10
Substance 4	.24***	.03***	.04***	.12***	.04***	.04***	497	.08
Substance 5	.24***	.06***	.07***	.26***	.00***	.05***	400	.11
Weight	.13***	.08***	.06***	.15***	.02***	.06***	443	.08
				No Information				
Transitivity				.28***	.55	.26***	115	.37
Number	.72	.23*	.18***	.60	.23***	.14***	207	.35
Substance 1	.54	.21***	.12***	.45	.00***	.05***	143	.25
Substance 2	.63	.19***	.12***	.34	.22***	.23***	149	.28
Substance 3	.46	.16***	.07***	.43	.19***	.26*	158	.27
Substance 4	.61	.24***	.23**	.34	.16***	.10***	153	.28
Substance 5	.73**	.17***	.12***	.39	.06***	.13***	171	.26
Weight	.54	.11***	.16***	.46	.19***	.19***	162	.27

TABLE E (CONT'D)

Correspondence Between Correct Judgements and Categories of Explanations

Test	Proportion of Explanations Associated with Correct Judgements						Total	Proportion Correct
	Prediction Phase			Transformation Phase				
	same	more	less	same	more	less		
Transformation								
Transitivity							0	
Number	.48	.05***	.05***	.50	.00***	.00***	90	.18
Substance 1	.43	.16**	.18***		.00***	.17	57	.22
Substance 2	.31	.00***	.16**		.00**		48	.15
Substance 3	.34	.18***	.06***		.13*	.25	64	.19
Substance 4	.34	.08***	.08***	.17	.09***	.12**	61	.13
Substance 5	.42	.04***	.04***	.15	.07***	.11***	117	.11
Weight	.36	.09***	.05***		.00***	.00***	98	.10

Note. - The proportions reported under the individual judgement questions are based on 5 or more explanations. If the individual cell is left blank, the number of explanations was less than 5. The Binomial test required N = 5 in order to be able to attain the α = .10 level of significance as all comparisons were two tailed. The abbreviations used in the "test" column are: transitivity - transitivity of substance, number - conservation of number, substance 1 - conservation of liquid substance with wide glass, substance 2 - conservation of liquid substance with tall glass, substance 3 - conservation of liquid substance with pie plate, substance 4 - conservation of liquid substance with graduated cylinder, substance 5 - conservation of solid substance, weight - conservation of weight.

* $p < .10$ **$p < .05$ ***$p < .01$

explanation category reflects preoperational thinking.

On the other hand, the reciprocity category is logically sufficient to be included with the appropriate categories (Piaget, 1970a). In the present study, reciprocity reversibility was scored when the subject (a) mentioned both of the relevant dimensions or (b) mentioned one or two of the relevant dimensions plus the idea of compensation. This definition of reciprocity reversibility is in accord with that used by Brainerd (1972b). It is, however, stricter than that used by Brainerd (1972c); in that experiment reciprocity reversibility was scored when the subject had reference to only one of the dimensions (e.g., This one is wider than this one.). In the present experiment the explanations referring to one dimension only were scored as "perception". On the other hand, the scoring of reciprocity reversibility seems less strict than that typically used by Piaget (e.g., 1970a); Piaget typically requires the subject to include the implication of compensation between the two dimensions. With reference to Table E, it can be seen that reciprocity reversibility as scored in the present experiment did not relate significantly to correct judgements.

Piaget's (e.g., 1970a) analysis of the transitions occurring in conservation performance suggests that having reference to the two dimensions, without the idea of compensation, may only reflect a level of thought transitional between preoperational and operational thought. Thus the lack of association between correct judgements and reciprocity reversibility explanations may be due to the fact that the definition of the reciprocity reversibility did not necessarily require concrete operational thinking; if such is the case then the definition of reciprocity reversibility is not consistent with Piagetian theory. This possibility was investigated by examining the 84 explanations originally scored as reciprocity reversibility. It was found that this category could be further divided into three categories: (a) Empirical Prediction: explanations offered in the prediction phases of the conservation tests which are of the form "the water will go higher in that one (e.g., cylinder) because it is narrower than that one (standard glass). (b) Two dimensions: explanations given in either the prediction or transformation phase which had reference to the two dimensions (e.g., that one (cylinder) is narrower and taller than that one (standard glass); or the water is narrower and higher in that one (cylinder) than that one (standard glass). (c) Compensation: explanations given in either the transformation or prediction phase implying compensation between the two dimensions (e.g., that one (cylinder) is narrow and tall and makes the water go up high, but that one (standard glass) is wider and shorter); or explanations referring to the invariance of the amount and mentioning one or two dimensions (e.g., they are the same amount of water, its just that one (standard glass) is wider).

The empirical prediction explanation comprised 11 percent of the total number of reciprocity reversibility explanations; all empirical prediction explanations were associated with incorrect judgements. The association with incorrect judgements was significant ($p = .004$, Binomial Test, two tailed comparison). The two dimension explanations comprised 51 percent of the total reciprocity reversibility explanations. Of the total number of two dimension explanations given, 25 percent were associated with correct judgements. The two dimension explanation category was also significantly associated with incorrect judgements ($z = 3.17$, $p = .002$; Binomial Test, two tailed comparison). The compensation explanations comprised 38 percent of the total reciprocity reversibility explanations. Of the total number of compensation explanations offered, 58 percent were accompanied by correct judgements. The association of the compensation category with correct judgements was not significant ($z = 0.87$, $p = .38$; Binomial Test, two tailed comparison).

With regard to the reciprocity reversibility category of explanations, then, it would appear that even though this type of explanation may appear as a sufficient basis for conservation, the statistical relation between the use of this category and correct judgements does not bear this out. Even when the reciprocity reversibility category was more congruent with Piaget's (e.g., 1970a) definition of this category - i.e., as implying compensation - the statistical relation to correct judgements failed to exceed the chance level.

In the present study, categories of explanation were deemed appropriate or inappropriate on a joint logical and statistical bases. First, the explanations of a category deemed appropriate needed to be accompanied by correct judgements on approximately 70 percent of its occurrences. Second, given the literature on concrete operational thought, the category needed to be congruent with these writings to the extent that it did not reflect preoperational thinking. Two explanation categories appeared to meet these criteria for the transitivity of substance task (comparison and logical necessity) and five explanation categories appeared to meet these criteria for the conservation tests (logical necessity, inversion reversibility, identity, second invariant and counting).

Appropriate explanations of stage II and III subjects

The complete pretest was administered to 16 N, 15 YMR and 29 OMR subjects who proved to function at stage II or III on at least one of the tests. In the present section, the number and type of appropriate explanations used by these concrete operational subjects are contrasted to investigate the possibility that even though all had demonstrated concrete operational thinking (by a judgement criterion), the structures of the YMR, OMR, and N subjects may be of a different nature.

The Kruskal-Wallis one way analysis of variance test was used to contrast the number of appropriate explanations offered by the concrete operational subjects from each age group on (a) the transitivity of substance, (b) the conservation of number, (c) the five substance conservation and (d) the conservation of weight tests. Only the tests on which the subject attained the Stage II or III criterion were used in the analyses. The Stage II and III subjects were combined because of low frequencies. There was no significant difference in the number of appropriate explanations offered by the concrete operational subjects from each age group on the transitivity of substance test ($H = 2.92$, $p > .10$), the conservation of number test ($H = 0.51$, $p > .30$), the combined five conservation of substance tests ($H = 0.45$, $p > .80$) or the conservation of weight test ($H = 0.28$, $p > .10$). Thus there is no difference in the total number of appropriate explanations offered by the subjects manifesting operational thought in each of the age groups.

The possibility of differences in the categories of explanations used by concrete operational subjects from each age group was investigated by comparing the number of explanations offered in each of the categories defined as appropriate. The "H's" resulting from the Kruskal-Wallis one way analysis of variance are presented in Table F. With reference to Table F,

INSERT TABLE F ABOUT HERE

the only overall comparisons to attain significance were in the use of the

TABLE F

Value of Kruskal-Wallis' "H" in comparisons of Categories of Explanations Used by N, YMR, and OMR Concrete Operational Subjects

Test(s)	Comparison	Logical Necessity	Inversion Reversibility	Identity	Counting	Second Invariant
Transitivity of Substance	2.92	0.15				
Conservation of Number	-	6.60**	1.82	1.28	0.83	-
Conservation of Substance (5 tests)	-	5.00*	1.08	1.53	-	0.13
Conservation of Weight	-	0.95	0.55	2.84	-	0.60

Note. - Values of "H" are not corrected for ties. All comparisons were two tailed.

* p .10

** p .05

logical necessity category on the conservation of number and the conservation of substance tests. Paired comparisons of the N, YMR, and OMR groups on their use of the logical necessity category on the conservation of number test indicated that the OMR group used significantly more logical necessity explanations than the N group (z = 2.32, p = .02, two tailed comparison) and the YMR group (z = 1.71, p = .09, two tailed comparison). The N and YMR groups did not differ (z = 1.32, p = .19). The paired comparisons using the Mann-Whitney test on the conservation of substance tests indicated that while the OMR group used significantly more logical necessity explanations than the YMR group (z = 1.65, p = .10, two tailed comparison), there was no significant difference in the use of logical necessity explanations between the N and OMR subjects (z = 1.37, p = .17, two tailed comparison) or between N and YMR subjects (z = .03, p = .98, two tailed comparison).

Logical necessity was scored when the subject stated that the quantitative property of concern was the same in response to the "more" and "less" judgement questions. (If the subject responded with "they're the same" explanation to the "same" judgement question, the explanation was scored as a tautology). The logical explanation category probably required the least verbal effort as measured by the number of words needed to express the semantic content of the idea. Hence, the more frequent use of the logical necessity category may reflect a more limited verbal ability on the part of the OMR subjects. However, it should be noted that there was no significant difference in the use of the other explanation categories, even though they tended to require more words.

Sequences in Acquisition of Conceptual Skills

In the interpretation of the results of the present experiment, it was assumed that some conceptual skills were acquired before others. Specifically, it was suggested that transitivity of substance and conservation of number were acquired prior to substance conservation, which in turn was acquired prior to weight conservation. The present analysis is directed at examining this postulate.

The datum for the analysis was the stage classifications for each test (I, II or III) for the 16 N, 15 YMR, and 29 OMR subjects demonstrating stage II or III performance on at least one test of the pretest. The performance of each subject was ordered according to the stage attained on each of the eight tests. A Friedman two way analysis of variance conducted on the ordinal data indicated that the tests were of differential difficulty ($\chi^2_r = 59.13$, df = 7, $p < .001$).

Comparisons of the individual pairs of tests were conducted using the sign test (Siegel, 1956). In comparing pairs of tests using the sign test, only those pairs on which the subjects obtained different levels are included in the analysis; comparisons involving tied ranks are eliminated from the analysis (Siegel, 1956). Consequently the proportion presented in Table G are based on numbers of different magnitudes. The proportion reported in Table G were conducted by the formula:

$$\frac{\text{Number of subjects obtaining higher level on test reported along top axis}}{\text{Total number of subjects showing discrepant performance on the two tests}}$$

INSERT TABLE G ABOUT HERE

As can be seen with reference to Table G, (a) the performance on the transitivity of substance test was significantly superior to the performance on all the other tests. (b) Similarly, the performance on the conservation of number test was superior to that observed on all other tests, except for the transitivity of substance test, (c) There was no significant difference

TABLE G

Within Subject Comparisons of Test Difficulty

Comparison Test	Test Forming Numerator in Proportion							
	Trans. Sub.	Cons. No.	Cons. Sub.1	Cons. Sub.2	Cons. Sub.3	Cons. Sub.4	Cons. sub.5	Cons. Weight
Trans. Sub.	-							
Cons. No.	.67**	-						
Cons. Sub. 1	.83****	.84***	-					
Cons. Sub. 2	.81****	.89****	.64	-				
Cons. Sub. 3	.79****	.89****	.59	.57	-			
Cons. Sub. 4	.86****	.96****	.64	.56	.47	-		
Cons. Sub. 5	.82****	.87****	.65	.67	.67	.65	-	
Cons. Weight	.88****	.90****	.79**	.71*	.73*	.75**	.70	-

Note. - The proportion includes only those subjects showing discrepant stage performance on the two tests.

The proportion is defined by: Number of subjects obtaining higher stage on test in "numerator" row / Total number of subjects with discrepant performance on the two tests

Abbreviations: Trans. Sub. - transitivity of substance; Cons. No. - conservation of number; Cons. Sub. 1 - conservation of substance (solid); Cons. Sub. 2 - conservation of substance (wide); Cons. Sub. 3 - conservation of substance (tall); Cons. Sub. 4 - conservation of substance (pie plate); Cons. Sub. 5 - conservation of substance (graduated cylinder); Cons. weight - conservation of weight.

* p .10   ** p .05   *** p .01   **** p .001

in the performances on the five substance conservation tests; therefore they are probably equivalent in difficulty. And (d) the performance on four of the five conservation of substance tests was superior to that obtained on the conservation of weight test; the conservation of substance (cylinder) and conservation of weight comparison failed to reach significance ($p = .17$).

In summary, the stage data would support the postulated order of acquisition of transitivity of substance, conservation of number, conservation of substance, conservation of weight. However, since not one of the proportions is equal to unity there is little justification for assuming that the above stated sequence is an invariant one.

BIBLIOGRAPHY

Boehm, L. D. Conscience development in mentally retarded adolescents. Journal of Special Education, 1967, 2, 93-103.

Brainerd, C. J. Judgements and explanations as criteria for the presence of cognitive structures. Psychological Bulletin, 1972, 78, (a).

Brainerd, C. J. On the age-stage issue in developmental theory: An illustrative experiment. University of Alberta, 1972, (b).

Brainerd, C. J. Reinforcement and reversibility in quantity conservation acquisition. Psychonomic Science, 1972, 27, 114-116.

Brainerd, C. J. Order of acquisition of transitivity, conservation, and class-inclusion of length and weight. Unpublished manuscript, University of Alberta, 1972, (d).

Brainerd, C. J. & Allen, T. W. Experimental inductions of the conservation of "first-order" quantitative invariants. Psychological Bulletin, 1971, 75, 128-144, (a).

Brainerd, C. J. and Brainerd, S. H. Order of acquisition of number and quantity conservation, Child Development, 1972, 43, 1401-1406.

Brison, D. & Bereiter, C. Acquisition of conservation of substance in normal, retarded and gifted children. In D. Brison & E. Sullivan (Eds.), Recent research on the acquisition of conservation of substance. Toronto: Ontario Institute for Studies in Education, 1966.

Brown, A. L. Conservation of number and continuous quantity on normal, bright and retarded children. Child Development, 1973, 44, 376-379.

Carlson, J. S. and Michaelson, L. H. Methodological study of conservation in retarded adolescents. American Journal of Mental Deficiency, 1973, 78, 348-353.

Copeland, R. W. How children learn mathematics: teaching implications of Piaget's research.

DeVries, R. Constancy of generic identity in the years three to six. Monographs of the Society for Research in Child Development, 1969, 34, (3, Whole No. 127).

Flavell, J. H. The developmental psychology of Jean Piaget. Princeton: Van Nostrand, 1963.

Flavell, J. H. Concept development. In P. Mussen (Ed.), Carmichael's manual of child psychology. Vol. 1 (3rd Ed.), New York: John Wiley & Sons, 1970, pp. 983-1059.

Goldschmid, M. L. & Bentler, P. M. The dimensions and measurement of conservation. Child Development, 1968, 39, 787-802.

Griffiths, J. A.; Shantz, C. A. & Sigel, I. E. A methodological problem in conservation studies: The use of relational terms. Child Development, 1967, 38, 841-848.

Gruen, G. E. & Vore, D. A. Development of conservation in normal and retarded children. Developmental Psychology, 1972, 6, 146-157.

Hooper, F. H.; Goldman, J. A.; Storck, P. A. & Burke, A. M. Stage sequence and correspondence in Piagetian theory: A review of the middle childhood period. Research Relating to Children, Bulletin 28, Washington, D.C.: U.S. Government Printing Office, 1971.

Houssiadas, L. & Brown, L. B. The coordination of perspectives by mentally defective children. The Journal of Genetic Psychology, 1967, 110, 211-215.

Hunter, I. & Lister, C. How children understand quantity. Science Journal, 1970, 6, 35-39.

Inhelder, B. The diagnosis of reasoning in the mentally retarded. New York: John Day Company, 1968.

Inhelder, B. & Piaget, J. The growth of logical thinking from childhood to adolescence. New York, N.Y.: Basic Books, Inc., 1958.

Inhelder, B. & Sinclair, H. Learning cognitive structures. In P. Mussen, J. Langer, & M. Covington, (Eds.), Trends and issues in developmental psychology. New York: Holt, Rinehart & Winston, 1969.

Jackson, S. The growth of logical thinking in normal and subnormal children. British Journal of Educational Psychology, 1965, 35, 255-258.

Keasey, C. T. & Charles, D. C. Conservation of substance in normal and mentally retarded children. The Journal of Genetic Psychology, 1967, 111, 271-279.

Lane, E. B. & Kinder, E. F. Relativism in the thinking of subnormal subjects as measured by certain of Piaget's tests. The Journal of Genetic Psychology, 1939, 54, 107-118.

Langer, J. Disequilibrium as a source of development. In P. Mussen, J. Langer, & M. Covington (Eds.), Trends and issues in developmental psychology. New York, N.Y.: Holt, Rinehart and Winston, 1969, pp. 22-37.

Langley, J.; Drew, C. J. and Watson, C. M. Performance of retarded children on a liquid conservation task: Protocol objectivity and visual screening. American Journal of Mental Deficiency, 1972, 76, 729-732.

Lister, C. M. The development of a concept of weight conservation in E.S.N. children. British Journal of Educational Psychology, 1969, 39, 245-252.

Lister, C. M. The development of a concept of volume conservation in ESN children, British Journal of Educational Psychology, 1970, 40, 55-64.

Lister, C. The development of ESN children's understanding of conservation in a range of attribute situations. The British Journal of Educational Psychology, 1972, 42, 14-22.

Lovell, K.; Mitchell, B. & Everett, A. R. An experimental study of the growth of some logical structures. British Journal of Psychology, 1962, 53, 175-188.

McManis, D. L. Additive composition by normal and retarded children of matched MA. Psychonomic Science, 1968, 11, 207, (b).

McManis, D.L. Conservation of mass, weight and volume by normal and retarded children. American Journal of Mental Deficiency, 1969, 73, 762-767, (a).

McManis, D. L. Conservation and transitivity of weight and length by normals and retardates. Developmental Psychology, 1969, 1, 373-382m (b).

McManis, D. L. Comparisons of gross intensive, and extensive quantities by normals and retardates. Child Development, 1969, 40, 237-244, (c).

McManis, D. L. Comparison of gross, intensive and extensive quantities by retardates. The Journal of Genetic Psychology, 1969, 115, 229-236, (d).

McManis, D. L. Conservation of identity and equivalence of quantity by retardates. The Journal of Genetic Psychology, 1969, 115, 63-69, (e).

McManis, D. L. Conservation, seriation and transitivity performance by retardates and average individuals. American Journal of Mental Deficiency, 1970, 74, 784-791.

Melnick, G. I. A mechanism for transition of concrete to abstract thought. Child Development. 1973, 44, 599-605.

Montroy, P.; McManis, D. L. & Bell, D. Development of time concepts in normal and retarded children. Psychological Reports, 1971, 28, 895-902.

Overbeck, C. & Schwartz, M. Training in conservation of weight. Journal of Experimental Child Psychology, 1970, 9, 253-264.

Piaget, J. The child's conception of number. New York: W.W. Norton & Co., 1952.

Piaget, J. The general problems of the psychobiological development of the child. In J.M. Tanner and B. Inhelder (Eds.), Discussions on Child Development, (Vol. 4), New York, N.Y.: International Universities Press, Inc., 1960, pp 3-27, 77-83(a).

Piaget, J. Second Discussion: Equilibration and the development of logical structures. In J.M. Tanner and B. Inhelder (Eds.), Discussions on child development (Vol. 4). New York, N.Y.: International Universities Press, 1960, pp. 98-115 (b).

Piaget, J. Six psychological studies. New York: Vintage Books, 1967.

Piaget, J. On the development of memory and identity. Barre, Mass.: Clark University Press, 1968.

Piaget, J. Piaget's theory. In P. Mussen, (Ed.), Carmichael's manual of child psychology. Vol. 1 (3rd Ed.). New York: John Wiley & Sons, 1970, 703-732 (a).

Piaget, J. Structuralism. New York: Basic Books, 1970 (b).

Piaget, J. Genetic epistomology. New York: Columbia University Press, 1970, (c).

Piaget, J. The theory of stages in cognitive development. In D.R. Green, M.P. Ford, & G.B. Flamer, (Eds.), Measurement and Piaget. New York: McGraw-Hill, 1971, pp. 1-11.

Piaget, J. The child and reality: Problems of genetic psychology. New York: Grossman publishers, 1973.

Piaget, J. & Inhelder, B. Diagnosis of mental operations and theory of intelligence. American Journal of Mental Deficiency, 1947, 51, 401-406.

Pinard, A. & Laurendeau. M. "Stage" in Piaget's cognitive-developmental theory: Exegesis of a concept. In D. Elkind and J.H. Flavell (Eds.), Studies in cognitive development: essays in honor of Jean Piaget. New York: Oxford University Press, 1969, pp. 121-170.

Prothro, E. T. Egocentricity and abstraction in children and adult aments. American Journal of Psychology, 1943, 56, 66-77.

Rothenberg, B. B. Conservation of number among four-and five-year-old children: some methodological considerations. Child Development, 1969, 40, 383-406.

Rothenberg, B. B. & Courtney, R. C. Conservation of number in very small children. Developmental Psychology, 1969, 1, 493-502.

Schiefelbusch, R. L. The development of communication skills. In R.L. Schiefelbusch, R.H. Copeland and J.D. Smith (Eds.), Language and Mental Retardation: Empirical and conceptual considerations. New York: Holt Rinehart and Winston, 1967, pp. 92-109.

Siegel, S. Nonparametric statistics for the behavioral sciences. New York, N.Y.: McGraw-Hill Book Company, Inc., 1956.

Stearns, K. & Borkowski, J. G. The development of conservation and horizontal-vertical space perception in mental retardates. American Journal of Mental Deficiency, 1969, 73, 785-790.

Stephens, W. B.; Miller, C. K. & McLaughlin, J. A. The development of reasoning, moral judgement, and moral conduct in retardates and normals. Philadelphia: Temple University. Project No. RD-2382-P, 1969.

Waddington, C. H. The strategy of genes. London: George Allen & Unwin Ltd., 1957.

Werner, H. The comparative psychology of mental development. New York: International Universities Press, 1957.

Winer, B. J. Statistical principles in experimental design. New York: International Universities Press, 1957.

Wolff, P. H. Operational thought and social adaptation. In Piers, M. W. (Ed.) Play and Development. New York: W.W. Norton & Company, 1972, pp. 28-42.

Woodward, M. The behavior of idiots interpreted by Piaget's theory of sensori-motor development. British Journal of Educational Psychology, 1959, 29, 60-71.

Woodward, M. Concepts of number of the mentally subnormal. Journal of Child Psychology and Psychiatry and Allied Disciplines, 1961, 2 249-259.

Woodward, M. & Stern, D. J. Developmental patterns of severely sub-normal children. British Journal of Educational Psychology, 1963, 59 10-21.

Zigler, E. F. The developmental approach of Jean Piaget: Open discussion. In M. Garrison (Ed.), Cognitive models and development in mental retardation. Monograph supplement to American Journal of Mental Deficiency, 1966, 70, 84-105.

Zigler, D. Developmental versus difference theories of mental retardation and the problem of motivation. American Journal of Mental Deficiency, 1969, 73, 536-556.

VITA AUCTORIS

1963 Graduated from St. Patrick's High School, Sarnia, Ontario.

1967 Bachelor of Arts Degree (Honours) from the University of Western Ontario, London, Ontario.

1968 Master of Arts degree from the University of Toronto, Toronto, Ontario.