PSYCHOLOGICAL RELATIONSHIPS AMONG SOME PIAGETIAN OPERATIONS

A STUDY OF THE PSYCHOLOGICAL RELATIONSHIPS AMONG THE PIAGETIAN OPERATIONS OF TRANSITIVITY, SERIATION AND CLASSIFICATION

by

WENDY MARGARET BRENNAN, B.A., M.A.

A Thesis

Submitted to the Faculty of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Doctor of Philosophy

McMaster University

August 1969

DOCTOR OF PHILOSOPHY (1969)
(Psychology)

McMASTER UNIVERSITY
Hamilton, Ontario

TITLE: A Study of the Psychological Relationships among the Piagetian Operations of Transitivity, Seriation and Classification.

AUTHOR: Wendy Margaret Brennan, B.A. (University of Manchester)

M.A. (University of British Columbia)

SUPERVISOR: Dr. L. R. Brooks

NUMBER OF PAGES: xi, 187.

SCOPE AND CONTENTS: Five to eight year old children were subjects in this study which examined relationships among the Piagetian operations of class inclusion, transitivity and seriation. Behavioral and verbal measures were taken of the latter two and the stimuli varied along size and weight dimensions. Transitivity and seriation were related on the verbal level only. Behavioral measures showed less relationship across operations and dimensions than verbal measures and doubts were raised as to the adequacy of using behavioral measures alone in assessing operational understanding. Only one measure of class inclusion showed any relationship to transitivity and seriation. Independent tests revealed that while simultaneous combinativity may be a component ability of all these operations, successive combinativity definitely was not.

ACKNOWLEDGEMENTS

I would like to thank Dr. L. R. Brooks for his assistance throughout the course of the research and for his insightful criticisms during the preparation of the thesis. I would also like to extend my sincere gratitude to Dr. Linda Siegel for her generous advice and encouragement during the analysis of the data and the writing of this dissertation.

I am indebted to the Hamilton Separate School Board and especially to Sister Conception, Principal of St. Joseph's School, and her staff for allowing me to use their pupils and facilities in this study. Thanks are also due to Mrs. Bobbe Papernick for her clerical work on early drafts of the thesis, and to Mrs. Mary Patterson who prepared the final manuscript.

Finally, I am grateful to my husband, Tony, for his encouragement and understanding throughout my student career.

To Casey,

who made his presence felt throughout the preparation of this thesis.

Sir Hugh Evans. ...What is *lapis*, William?

William. A stone.

Sir Hugh Evans. And what is *a stone*, William?

William. A pebble.

Sir Hugh Evans. No, it is *lapis*: I pray you remember in your prain.

William. *Lapis*.

Sir Hugh Evans. That is good William.[1]

---

[1] Shakespeare, Wm., *The Merry Wives of Windsor*, Act IV, Scene i, ll. 33-40.

Table of Contents

Tables

Figures

CHAPTER ONE

Introduction

The child from five to eight years old has been the subject of much recent research since it has become apparent that many profound cognitive changes are occurring within this age period. These changes have been studied and noted by developmental psychologists of widely differing theoretical and experimental approaches (see White, 1965). For instance, American S-R psychologists such as the Kendlers (1962a and b) have found developmental changes between five and eight years of age in a child's ability and style of learning concepts and in his capacity to indulge in inferential behavior. The Russians, Razran (1933) and Vygotsky (1962) have noted marked developments in this age range, the former in the child's resistance to classical conditioning, the latter in speech and its subsequent role in the planning and representation of behavior. Dramatic perceptual changes have been noted by Bruner (1964). And, of course, in the theoretical system of Jean Piaget one of the most important developments in the cognitive structure of the child occurs at this age, namely the change from what Piaget calls the preoperational stage of development to the concrete operational stage.

The present study is set within Piaget's theoretical and methodological framework. It endeavours to test some of Piaget's notions regarding the structural nature of concrete intelligence as it is developing in the child between his fifth and eighth years.

PIAGET'S THEORY OF THE STRUCTURAL NATURE OF CONCRETE OPERATIONAL INTELLIGENCE

As a genetic epistemologist Piaget is interested in the growth of knowledge, the laws under which it develops and changes. His research seeks to analyze how superior levels of intelligence and scientific thinking come about from elementary forms of cognition. One aim of his research is to uncover the structuring capacity of the organism at each developmental level, to examine the inner structure which he insists underlies a knowing response at a particular developmental level. He is mainly interested in what is general or generalizable in the knowing structure of an individual and not in the unique aspect of a given behavior. For instance, Piaget examines a child's ability to order objects according to their sizes so that he may discover the general behavior of relating and ordering that is manifest in a variety of concrete situations. He is not very concerned about the procedural or material details of these situations.

For Piaget the development of intelligence is only understood through the study of the structure of intelligence. A reaction of an organism is a response of its underlying structure and not merely a response to outside stimulation. In order to explain a response one must examine the underlying structure that makes the response possible. For Piaget, behavior at all levels demonstrates aspects of structuring. But the capacity for structuring the environment differs with age; hence a theory of the stages of intellectual development is necessary. One of the most fundamental criteria of Piaget's concepts of a stage, his notion of "structures d'ensemble", is that "the typical actions or operations of a given level are not simply juxtaposed one with another in an additive fashion, but are organically interconnected by ties of implication and reciprocal dependence that unite and group them into total structures" (Pinard and Laurendeau, 1969, p136). In other words, there is an interrelatedness among the development of diverse behaviors conforming to the same underlying structure.

The models which Piaget uses to describe this structure are his groupings, a set of quasi-logical entities of his own creation possessing some of the properties of the group and the lattice, two logical structures familiar to mathematicians. The groupings define the formal properties of the reasoning process of the concrete-operational child and are intended as an ideal system of all possible cognitive operations. Piaget proposed the existence of a total of nine groupings which deal with the basic nature of the concepts of classes and relations and the possible operations one can perform on them. His approach in devising them was logical, not empirical: they were not derived from observations of children's behavior or thinking. In fact, there is no empirical evidence as yet for the existence of two of his groupings.

The entities which are organized into these structures are active, intellectual operations. Operations are actions at first carried out on objects and later carried out in thought. They are reversible in that they can return to the point of origin, unlike simple actions which are irreversible. According to Piaget the first operations appear around seven or eight years of age in a concrete form--i.e., they can only be carried out on actual objects and not on the level of abstract thought. All concrete operations--e.g., conservation, transitivity, relationships among classes, notions of serial order and measurement, concepts of number, etc.--belong to the logic of classes and relations as described by Piaget's nine groupings (Piaget, 1953). Operations originally "derive from sensory-motor actions by progressive interiorization and structuring of coordinating action schemes" (Furth, 1969, p62). Furth interprets this interiorization of actions to form operations to mean that thought becomes increasingly independent from specific stimuli situations.

Wohlwill (1963 and 1966a) interprets the notion of "structures d'ensemble" as requiring that the acquisition of a logical operation at a given stage implies simultaneous mastery of all the problems or tasks founded on this operation. The mastery of all tasks exemplifying one grouping should proceed in synchrony. He cites Piaget's claim that conservation of area, distance, and length all appear at the same time since all involve the operations of addition and subdivision of parts and coordination of positions; and also his claim that locating a point in two-dimensional space, linear measurement, and the spontaneous conceptualization of spatial coordinates all develop together since they are all based on the operation of logical multiplication. For Wohlwill, finding an interrelationship between responses to tasks which differ but which, according to Piaget's system, are based on the same intellectual operations, lies at the core of Piaget's conception of the stages of mental development. One of the purposes of the present study is to test Piaget's claim that the construction of asymmetrical transitive relations implies a system of serial relations, since both are based on Piaget's grouping V--the addition of asymmetrical relations (Piaget, 1953).

Wohlwill points out further (1966a) that not only does Piaget claim that mastery of all tasks exemplifying one grouping should proceed in synchrony but that he also often makes the same assumption for tasks based on different groupings

In fact one could interpret Piaget's stage theory as implying that all concrete operations should appear at the same time. Piaget's main collaborator Inhelder (1956) says that the structure of concrete operations implies the solution of all elementary problems of classification, seriation, and numerical conservation. Pinard and Laurendeau (1959) remark that Piaget and his associates in Geneva emphasize on many occasions the "structural isomorphism and genetic synchronism" of certain concepts that are quite dissimilar and even arise from different areas (i.e., logical and sublogical, the latter dealing mainly with concepts of space, time, or measurement). For instance, they posit a relationship between hierarchical inclusion of classes and certain spatial constructions of topographical wholes (Inhelder and Piaget, 1959); also relationships between the compositions of spatial wholes and collections of discontinuous objects (Piaget, Inhelder, and Szeminska, 1960). It is a second purpose of the present research to examine the developmental relationship postulated between the operations of class-inclusion and seriation. Each is a manifestation of a different grouping and they are two of the major operations whose presence define the concrete operational stage (Inhelder and Piaget, 1964).

PIAGET'S METHOD

Two pertinent questions concerning how to test the aspects of Piaget's theory reviewed above come to mind:

1) How does one test a theory which postulates the gradual approximation of thought to certain logical models? and 2) Given a method of diagnosing the presence of certain structures, how does one discover the interrelationships between structures?

As for Piaget's answer to the first question, thinking is said to approximate the model when a child _behaves as though_ he were operating according to a particular model. Examining this calls for a special approach to observation, "designing intellectual tasks which externalize thought into action in such a way that one is able to infer the logical assumptions on which the action was based and the nature of the logical system...from which these assumptions were derived" (Bruner, 1959, p363). Thus, Piaget uses a clinical method of observation "for the object of clinical diagnosis is to evoke for observation those forms of behavior which indicate to us how some underlying system is operating" (_ibid_, p363). This means that each child is questioned and prodded about his performance on the tasks Piaget sets for him to test the limits of his understanding about the operation being examined.

The second problem is how to determine that two problems which are similar in logical structure as seen by the logician are similar on the plane of psychological process. In the Genevan experiments if two or more logically similar tasks are mastered around the same age the integration of the underlying processes and operations is inferred. But different tasks have usually been given to <u>different</u> groups of children, or if perchance some of the same children are used in several experiments, within subject results are not examined (e.g. Inhelder & Piaget, 1964).

Thus, Piaget's theory of the underlying organization of children's cognitive operations is empirically based on the similarity in ages of emergence and the improvement with age of each ability as each was assessed in different groups of children. Many British and North American studies have been done testing the interrelatedness of various logical structures by giving different tasks to the same children and analysing the within subject relationships among them. It will be fruitful to review some of these findings here to see how well Piaget's theory fares when put to more rigorous empirical tests than his own.

RESULTS OF NON-GENEVAN STUDIES OF THE INTERRELATEDNESS OF COGNITIVE STRUCTURES.

In summarizing the themes of the articles in the book of readings which he co-edited, Sigel (1968) observes that the studies in the volume do not present strong support for the hierarchical and integrative aspect of Piaget's theory of cognitive growth. Each study investigated different combinations of presumed interlocking operations, but in general there was a failure to find consistency in the performances of individual children across a range of diagnostic tests of concrete operations. First, let us examine several studies which examined performance on different tasks all concerned with the _same_ concept. Kofsky (1966) gave four to nine year olds eleven tasks designed to tap the developmental sequence of classificatory behavior. A scalogram analysis of the results revealed an overall order of difficulty of the tasks which corresponded to the order of development predicted by Piaget. But those children who passed the more difficult items did not necessarily pass all the easier ones. Lunzer (1960) tried to determine the interlocking nature of mathematical elements involved in the ability to measure volume. He found no consistent pattern in his subjects' responses to conservation of volume tasks and their understanding of the notion of infinite subdivision, even though Piaget says the latter is the basis of the former. A study by Dodwell (1960) reveals only moderate degrees of consistency between a variety of tasks in the area of number which Piaget regards as intrinsically interrelated aspects of the number concept. In another study Dodwell (1963) devised seven groups of tests each covering a different aspect of spatial concepts and each made up of a number of items designed to tap various facets of the same problem. Intercorrelations of performances in various areas were low and there was also a general lack of consistency among responses to the items within each of the areas, which merely represented variants of the same task or problems.

The story is similar for studies which examine inter-relationships between structures although Braine (1959) does claim to have found a close association between a child's ability to respond to ordinal position of an element in a series and his success in a measurement task. On the other hand, Lovell and Ogilvie (1961) and Smedslund (1961) fail to substantiate the close relationship between the acquisition of conservation and an understanding of transitivity of weight which Piaget's theory postulates. Piaget claims that conservation is one of the very basic underlying abilities of all concrete operations. Dodwell (1962) found a very low correlation (+.20) between formally equivalent problems involving the concept of cardinal numbers and the logic of classes, two basic operations acquired during the concrete operational period. In a study examining the relationships between the three multiplicative groupings of classes, logical relations, and spatial relations within the same subjects, Shantz (1967) found that the correlations between tasks seldom accounted for more than 25 percent of the variance, thus giving only moderate support for Piaget's hypothesis that a close relationship exists among the multiplicative abilities.

Even studies which test the generality of one operation across various materials fail to find the correlations which Piaget's theory would seem to require. For instance, Lovell and Ogilvie (1960) found that subjects who were non-conservers in one situation involving quantity were not inevitably non-conservers in another: one-third of the subjects who were non-conservers with plasticene balls were conservers with a stretching rubber band. These authors also present evidence from two University of London dissertations (Beard, 1957 and Hyde, 1959) which found that non-conservers with plasticene balls were conservers when liquid was poured from one vessel into another of a different shape or into a number of smaller vessels. Uzgiris (1964) gave 140 subjects from Kindergarten through Grade 6 tests of conservation of quantity, weight, and volume using the same four materials with all three physical properties. The results confirmed the findings of Piaget and others that conservation developed first with quantity, next with weight and last with volume. But this sequence held only with any given material: a child's position on the conservation sequence was not constant across material. That is, just because he demonstrated conservation of weight with one material did not mean he demonstrated it with another. Furthermore, the variation was not systematic: there was no single material on which all subjects were either accelerated or lagging behind. The results of these studies all throw into serious doubt Piaget's claim that once conservation has been attained with one physical property it holds in all situations involving that property.

POSSIBLE REASONS FOR LACK OF CONCORDANCE IN RESULTS CITED

As the above results demonstrated, the patterns which emerge in studies of cognitive development are complex and not completely compatible with Piaget's statements about the topic. From a purely experimental standpoint there are many factors which might tend to disrupt the clear pattern of stages described by Piaget. Such subject variables as special interests and training, the amount of formal instructions, and difficulties of learning to apply a set of rules or operations learned in one context to a new situation may mar an otherwise unified picture (Dodwell, 1963). Furthermore, children's performances are notoriously unreliable. Young children do not feel the need to be consistent and are more likely to perform in a random fashion than adults (Kofsky 1966). Also such task variables as variations in the difficulty and form of the instructions, differences in the familiarity and in the concreteness or abstractness of the stimuli and variations in the manners in which the tasks are structured may mask any real regularities in logical development.

It is equally likely, however, that the reasons lie within Piaget's theory itself as in the empirical tests of it. As Hunt has pointed out, it is probably true "that Piaget is correct in asserting that the concrete operations become Gestalt-like operational structures in which one can find reflected the logical operations he had attributed to them. But it is also probably true that Piaget is wrong in asserting that these emerge quickly and all at once. Rather, they appear to emerge tentatively from coping with a given kind of problem in one situation, then again with that kind of problem in another situation, in another, and another, and so on, then from coping with related kinds of problems in a variety of situations until the rules for the solution become generalized" (Hunt, 1969, p54). Besides, as Dodwell (1963) has noted, the part of Piaget's theory which seems to say that once an operation has been acquired its deployment in novel situations should present few if any difficulties is not a logical requirement of the theory. It is quite reasonable that such response generalization should be incomplete at first and only improves with experience. Furthermore, Wohlwill (1963) reviews evidence showing that often (but not always) when there is a departure from synchrony in the development of concepts, in violation of Piaget's structural model, this departure usually takes the form of one concept being mastered consistently ahead of the other. That is, the deviation is often not random but tends to follow a predictable sequential pattern, in the scalogram-analysis sense.

Pinard and Laurendeau (1969) warn that one should not use the frequency of dysynchronisms found in the literature as a pretext for rejection of the "structures d'ensemble" hypotheses which they interpret as characterizing a relatively final stage of development. When newly emerging structures are in the process of formation, as for instance during the so-called transition period between five and eight years, a child's responses might be expected to oscillate from one occasion to the next and to be maximally susceptible to the effects of task-related variables. That Piaget recognized this is shown in two ways. First, he provided for a stabilization phase in the formation of each stage in which newly formed structures are undergoing consolidation. Second, he proposed the concept of "horizontal décalages" to account for the differentials in performance relating to the particular content of a task or to a variety of situational variables.

Let us examine further this notion of horizontal "décalages" or "differentials" as Flavell translates it (1963). Piaget and many other researchers have found that the concept of conservation emerges with respect to substance at around five or six years of age, for weight two years later, and for volume two years after that. Similarly, Piaget (1953) has found that the notion of transitivity of the equalities of length develop around seven or eight years, of weight around nine or ten, and of volume around eleven or twelve. Piaget's account of horizontal differentials is that "each field of experience (that of shape, size, weight, etc.) is in turn given a structure by the group of concrete operations, and gives rise in its turn to the construction of invariants (or concepts of conservation). But these operations and invariants cannot be generalized in all fields at once; this leads to a progressive structuring of actual things, but with a time-lag of several years between the different subject-matters. Because of this, concrete operations fail to constitute a formal logic; they are incompletely formalized since form has not yet been completely divorced from subject-matter" (Piaget 1953, p17). As is plainly obvious, this account is merely a description of events and not an explanation. Wohlwill (1966b) criticizes this concept for several reasons. First, it is merely an ex post facto explanation and offers no basis for predicting responses. Second, it is not even a complete description of events since Piaget has not indicated, for instance, where the conservations of length, area, or number fit into the substance, weight and volume sequence. Further, it cannot handle the discrepancies in the acquisition of different concepts which, according to Piaget, mutually imply one another--for instance, conservation and transitivity.

Piaget and Inhelder (1941) do give reasons why seriation is not applicable to all dimensions of objects at once but develops in the sequence of quanity, weight, and volume. The explanation deals with a child's "field of simultaneous visual perception" and how this interacts with the level of development of his intellectual operations. This explanation will be examined in more detail below in connection with the results of the present research, in which transitivity and seriation tasks of both size and weight were presented to each child to see the order of development of the two operations in these two physical properties within the same subjects.

DIAGNOSIS FOR PRESENCE OF UNDERLYING STRUCTURES

As has been suggested above one of the reasons for the lack of intertask correlations is the variation in the methods used to test whether a child had a grasp of the various operations implied in the tasks used. The problem of just how one tests for the presence of an underlying structure has generated some discussion but as yet no satisfactory solution upon which all can agree has been found.

A distinction made by Flavell and Wohlwill (1969) between the competence and automaton aspects of cognition is relevant here. These authors claim that one must distinguish between two principle components of a psychological theory which accounts for complex behavior. The first is a <u>competence</u> model which is a formal, logical representation of the structure of some domain. It is the abstract, purely logical representation of what the child knows or could do in a timeless, ideal environment. The second is an <u>automaton</u> model which encompasses the psychological process by which the information embodied in the competence aspect of cognition actually becomes available and is utilized in real situations. These are the mechanisms used for coding, processing, and divulging information rather than the reasoning processes as such. The authors give as illustration of the use of these models the performances of three children in situations requiring transitivity inferences. Child A, aged four years, never makes transitive inferences in any situation and cannot be taught the principle by any method. Child B, aged eight, makes transitive inferences on some tasks but not others, but training attempts on the latter are partially successful. Child C, aged 14, correctly applies the rule in every situation without training. The interpretation of these results in the light of this model would be that for Child A transitivity has not yet become part of his abstract cognitive competence. For children B and C the transitive operation is part of their competence but they differ in their automaton system so that B cannot transfer and generalize the rule to all situations whereas C can. "A competence-automaton approach does not ignore or underplay the undeniable capacity difference between B and C, but it wants to assert that it is not the same kind of difference as that which distinguishes A and B" (Flavell and Wohlwill, 1969, p73).

It is obviously the competence model of cognition with which Piaget's theory is concerned, but it is the automaton model with which we have to deal in order to examine his theory empirically. The various opinions and debates in the literature as to how to diagnose a child's competence with respect to a given operation and across several operations have really been concerned with how to cope with and control the automaton aspect of his intelligence--i.e., those subject and situational variables which may operate to obscure the true nature of a child's reasoning ability. For instance, Smedslund (1964) has some suggestions about controlling the testing situation. He would have us distinguish and control three things in the testing situation: the percept--the stimulus situation as apprehended by the child; the goal object--the pysical property (e.g., length, quantity, weight) with which the task deals; and the inference pattern--the set of premises and the conclusion required in the situation (e.g., transitivity, or conservation). To study any one factor the other two must be held constant. For instance, if one wanted to investigate the relationship between conservation and transitivity, the percept and the goal object of the tasks would have to be exactly the same. It is Smedslund's claim that this approach offers the only chance of finding exact relations between the operations. Presumably then, if exact relations were not found one could conclude that it must be due to the competence aspect of cognition, that is, to a real lack of relationship between the underlying operations themselves, and not to the automaton aspect, that is, methodological and subject variables in the situations.

It could be argued, however, that an experiemntal design fulfilling Smedslund's conditions would be extremely difficult to produce and, moreover, would reveal nothing about the generality or spontaneous applicability of a child's operational understanding. The results derived from such highly artificial situations would be at least as equivocal about the course of natural development as those gleaned from the diverse but more natural situations presented to children in most of the research to date. The present study makes no attempt to control for all aspects of the stimulus situations in the tasks used. Rather, each operation is represented by tasks designed to be as provocative as possible in disclosing the nature of the child's understanding of the operation, and it was hoped that each would be as natural a medium through which to reveal underlying processes as is possible in an already artificial testing situation.

There are subject variable aspects of the automaton model which receive attention in the literature also. For instance, Pascual-Leone (1967), in a study comparing the responses of 10 year olds on certain Piagetian tasks (water-level, conservation of substance, and weight) to their responses on four tests of Field Dependence, presents results which suggest the necessity of including individual differences and cognitive styles among the variables to be controlled experimentally in developmental studies in cognition. He found that a much weaker relationship was found between the Piagetian tasks in field dependent than in field independent subjects. He suggests from this that the frequently reported low intercorrelations among Piagetian tasks hypothesized to belong to the same developmental level may be due to the fact that the samples were not homogeneous with respect to the subject's cognitive style--i.e., their field dependence or independence. Since there are usually many misleading perceptual cues in all the situations, it is possible that the field dependent subjects severely attenuated the intercorrelations patterns of the sample. If only field dependent subjects were tested, they might be able to overcome the confusion present in misleading perceptual cues in the situations and the intercorrelations predicted by Piaget's theory might appear.

One of the most important subject variables is of-course language, not only a child's verbal fluency, but his basic understanding of verbal cues. One important question about language is whether to employ it at all in diagnostic tests and if so to what degree and how. This question is the basis of the Braine-Smedslund debate reviewed below. A second question is to what degree one should depend on the child's verbal responses as a major, if not only, indication of the presence of an underlying structure. Several opinions about this question are also reviewed below. But both questions depend for their answers on what one considers the role of language to be in the <u>development</u> of the cognitive structures in the first place. It is to a discussion of this that we will turn next before dealing with these questions more fully.

ROLE OF LANGUAGE IN DEVELOPMENT OF COGNITIVE STRUCTURES

As far as Piaget is concerned, language is not an intrinsically necessary element of operational thinking. Concrete operations originate in sensory-motor shemata and not in language or perception, except in so far as the latter is integrally bound up with the sensory-motor schemata. Language is only one of the manifestations of the symbolic function which emerges at the end of the sensory-motor period, the others being symbolic play, deferred imitation, and the mental image. All are internalized imitation. The development of the symbolic function is in turn dominated by intelligence in its total functioning. Once language is acquired this does not mean that the child receives the operations ready-made from the outside through the medium of linguistic constructs. The development of all structures demands an active construction on the part of the child. Therefore, "a verbal transmission that gives adequate information relative to operational structures is only assimilated at levels where these structures have already been elaborated on the plane of action or of operations as interiorized actions" (Piaget, 1969, p127).

Language is a symbol system used for communication and is therefore important for socialization. It is a vehicle of symbolization without which thought could never really become socialized. But it is not an indispensible medium for intelligence. For instance, Furth and his colleagues have devised thinking tasks not couched in verbal terms and have found that the basic manifestations of logical thinking in deaf children, who are of course linguistically deprived, were present without any important structural deficiencies (see Furth, 1969).

A contrasting view to Piaget's is that of Bruner (1964), who claims that since language releases the child from dependence on immediate perceptual input, its use as a program for ordering and integrating experience is the mechanism of transition from iconic to symbolic levels of cognitive functioning, i.e., from what Piaget designates as the preoperational to the concrete operational stages.

These two contrasting views on the importance of language in the developing cognitive structure of the child precipitate contrasting experimental approaches as well. Since for Piaget language reflects rather than affects underlying structures he has no hesitation in using it liberally in his investigations of children's thinking. On the other hand, an investigator of Bruner's persuasion would avoid its use in his experimental design lest it obscure or lead a child's thinking in the situation.

ROLE OF LANGUAGE IN DIAGNOSIS OF UNDERLYING STRUCTURES

In his investigations Piaget presupposes adequate linguisitc competence on the part of his subjects and he also believes that children adequately express what they think. Therefore, if a child does not express the logical principle of, say, transitivity it is because that operation has not yet developed. When he has acquired the operation adequate logical definitions will be forthcoming. Piaget, therefore, usually requires a verbal justification of behavioral responses as a basis for diagnosing a child as being "operational" or not with respect to a given operation. Also, he rarely, if ever, bothers to determine beforehand whether or not a child understands the crucial terms in the instructions--for instance, relational terms such as "more", "longer", "same", etc. "Piaget holds that lack of comprehension of these terms indicates that the child has not assimilated this knowledge to the appropriate cognitive structure. Therefore, such lack of comprehension in itself is an indication of cognitive level" (Sigel, 1968, p520). On the other hand, Piaget has shown that just because a child does understand and can use such relational terms in some simple situations does not mean that he can respond correctly in situations which demand the use of conservation or transitivity inferences (see Furth, 1969, Ch6.). These situations demand more of the child than mere verbal competence.

This is where Braine (1962) differs with the Genevan group. As far as Braine is concerned if a child understands these simple verbal cues in the situation, this means he must have developed the concept already. Therefore, one cannot study how a concept develops with methods which employ verbal cues to evoke the concept. "If one seeks to state an age at which a particular type of response develops the only age which is not completely arbitrary is the earliest age at which this type of response can be elicited using the simplest experimental procedures" (Braine, 1968, p187). For Braine this means a completely non-verbal operant conditioning, stimulus control situation where responses are reinforced in the presence of the concept under investigation and then tests are made for stimulus generalization. His rationale for this method is that if the child is presented with a behavioral problem, the solution of which requires the understanding of a logical inference, and he gets it right, then we can say that the child possesses that operation whether or not he can express it verbally. The problem then becomes how to set up a behavioral task which can only be solved by recourse to the underlying inference it is supposed to reveal. This led to the controversy between Braine (1959, 1964) and Smedslund (1963) as to the adequacy of their various procedures to do just this with the transitivity of length operation. Braine's strictly behavioral procedure revealed the "emergence" of the transitivity operation at around five years. Smedslund's procedure, which required the use of language on the part of both E and S, used no reinforcement or corrections of responses, and employed a more stringent criterion of transitivity, revealed the operation to emerge around eight years.

As Gruen (1966) points out, it is no wonder that each finds a different age of emergence of transitivity since each is using different criteria for diagnosing the presence or absence of the operation. But the disagreement is not merely one concerning age norms, nor is it entirely a matter of which experimental procedures are appropriate for assessing cognitive processes. It is basically a disagreement about what are the necessary and sufficient conditions for Piagetian concepts to be formed. The different methods Braine and Smedslund have employed have led them to study qualitatively different phenomena. The crucial question is then what is it that Braine finds in four and five year olds that he calls transitivity and how does it differ qualitatively from what Smedslund finds in seven and eight year olds? This question has not been answered yet. But it is Hunt's (1961) opinion that Braine's interpretation of the transitivity inference as a capacity which can be revealed even as a learning set in a specific situation has little bearing on Piaget's conception of this inference as an operation of thought readily available and serving the end of seeing how things work.

While the present study employs nothing akin to Braine's discrimination method, it does distinguish between behavioral responses and the verbal justifications given for these responses for the operations of transitivity and seriation in the hopes of perhaps discovering some differences in the quality of these kinds of responses.

PURPOSES OF PRESENT RESEARCH

1. RELATIONSHIP BETWEEN SERIATION AND CLASSIFICATION

In "The Early Growth of Logic in the Child", Inhelder and Piaget (1964) studied the development of seriation and classification "to discover the essential formatory mechanisms", i.e., the underlying logical structures, common to both operations. They studied a handful of children of each age from preschool to middle elementary school age and concluded that "the development of seriation (of length) is almost exactly parallel to that of classification and tends to precede it step by step" (p4). This conclusion and the elaborate theories presented to explain the observations place great emphasis on the integrated nature of the development of these two operations. As Donaldson (1960) points out, "it is odd, then, to find instances where the same subjects figure in more than one table and yet not to be told fully and explictly how the two sets of results relate to one another" (p184). One of the aims of the present research was therefore to repeat many of Inhelder and Piaget's procedures and to analyze the within-subject data on the development of the two operations of classification and seriation. Thus we will be able to see if, within the same child, seriation and classification develop temporally as Piaget says they do.

Classification and seriation are two of the most important operations which develop between the ages of five and eight and their presence distinguishes the concrete operational from the preoperational child. The presence of operational classification is evident when a child can define the qualities of a class in terms of a more general class and can give one or more specific differences between subclasses. Piaget calls this class "intension". Further, the child can distinguish between the members of the subclasses and the general class as shown by his mastery of the quantifiers "all", "some", "one", and "none". This is called class "extension". At the concrete operational level this operation can only be performed when actual stimuli are in front of the child. For instance, if a collection of flowers composed of some tulips and roses were put before him, he would be able to reason that "all" of the tulips comprised only "some" of the flowers so that if one took "all" of the tulips away there would be "some" flowers (i.e. the roses) remaining. He can correctly answer the question "Are there more tulips or flowers here?" The preoperational child cannot do this for, says Piaget, he cannot compare the part (tulips) to the whole (flowers) but only to another part (roses). Tasks involving the child's understanding of the quantifiers mentioned above are among Piaget's more famous and intriguing and were repeated with minor modifications in the present research.

Operational seriation is no more than a symbolic representation of a series which is seen by the child to be more than merely temporally (before-after) or spatially (left-right, up-down) bound. A child possessed of operational seriation can deal with asymmetric transitive relations (e.g. $A > B > C$) and recognize their reversibility (i.e. if $A > B$, then $B < A$), according to Piaget's logical analysis of this operation. Each child in this study was asked to line up several sets of objects which varied along length and weight dimensions and his performance on these tasks is compared to that on the classification tasks mentioned above to determine the degree of interrelationship between these two operations.

The relevant groupings in this comparison are Groupings I the primary addition of classes, and V, the addition of asymmetrical relations. To the author's knowledge no other within-subject comparison of the development of these two structures has been done to date.

2. RELATIONSHIP BETWEEN SERIATION AND TRANSITIVITY.

Piaget's Grouping V, the addition of asymmetrical relations, deals with the composition of an ordered series from the coordination of successive transitive relationships among points on a dimension. The behavioral manifestation of this structure would thus be given in the seriation operation or any performance directly based on transitivity or ordering. But is it empirically true that the operations of seriation and transitivity are interchangeable exemplars of this structure? Does the understanding of one imply the understanding of the other? Many authors think so. For instance, Furth notes that around seven or eight years the child "grasps the double principle of seriation which requires that each element must be consistently compared to its neighbour in the one direction and to its neighbour in the reverse direction. When this is mastered and the operation of seriation is reversible and firmly structured, the principle of transitivity follows as a matter of course. If A is bigger than B and B is bigger than C, it is understood through thinking alone that A is bigger than C". (Furth, 1969, p.64). Piaget (1953) implies that seriation develops before transitivity when, in discussing his own data on the transitivity of equality of weight, he says that transitive reasoning will not occur until the weight relationships are structured by a preliminary group of operations, such as seriation.

Piaget and Inhelder (1941) describe investigations done in Geneva on the seriation and transitivity of weight. For the seriation tasks, children were presented with three, four, six, and ten objects to be ordered, sometimes in a free unstructured situation where the child could weigh the objects in his hands or on a balance in any way he chose, and sometimes in a restricted situation where the child was told to weigh the objects two at a time on the balance or in his hands. For the transitivity task, three matchboxes were put in front of the child and he was told, as the experimenter pointed appropriately, that A was heavier than B, and B was heavier than C, and was asked to designate the heaviest and lightest. The information regarding the relationships between the matchboxes was also given in two other forms: $A > B$ and $C < B$, and $B < A$ and $> C$. The results revealed three stages of development towards operational seriation. The child in stage I cannot seriate anything by any method. By stage II he can seriate the 10 elements by trial and error in the free situation only, but he cannot seriate four or six elements when restricted to weighing them two at a time. By stage IIIA he can seriate 10 elements in the free situation "operationally", i.e., by choosing and measuring the heaviest of all which have not yet been placed in line. But he can only seriate three elements and no more using the restricted method, which means he can only coordinate two relations using this method. By stage IIIB he can seriate 10 elements by either method, which means he can coordinate many relations. This stage is reached by nine or ten years of age. It is not until stage IIIB that the child can make transitive inferences about the matchboxes no matter how the information is presented. The hardest method of presentation is when he is told that B is lighter than A but heavier than C. These results seem to demonstrate that operational seriation and transitivity of weight do develop at the same time within the same child. There is no doubt that the tasks presented in this study were very demanding of the child and employed very stringent criteria of operational seriation and transitivity, especially compared to most North American studies of these operations. For instance, Battisti and Simmons (1968) and Murray and Youniss (1968) conducted investigations comparing performances on transitivity and seriation of length tasks within the same subjects using only simple behavioral responses, no verbal justifications, as diagnostic criteria for the presence of operations. The former study found that success on transitivity preceded success on seriation somewhat, although neither ability was examined thoroughly. The latter study found that the seriation task was passed by more subjects than the transitivity task, although it can be argued that the seriation task used here was neither comparable in difficulty, nor in the thoroughness with which it examined the operation, to the transitivity tasks used.

One purpose of the present study is to see the relationship between transitivity and seriation performances in both length and weight dimensions. Each task was designed to examine these operations as thoroughly as possible without becoming tedious. Thus several sets of different kinds of objects are used in each task in order to ascertain something about the generality of the child's understanding and to give the best possible chance of this understanding revealing itself in the relatively artificial experimental situation.

3. RELATIONSHIP BETWEEN TRANSITIVITY AND SERIATION ON SIZE AND WEIGHT DIMENSIONS.

As discussed in an earlier section, previous research using different groups of subjects has found that transitivity and seriation of size develop around two years before transitivity and seriation of weight. Piaget developed his concept of "horizontal décalages" to account for these results. A purpose of the present research is to see how performance on both these operations develops in each of the dimensions of size and weight within the same subjects. It will also be of interest to see, if they do develop at different times, which one is most closely related to the development of the classification operation. Inhelder and Piaget (1964) claim that classification develops around the same time as seriation of length (i.e., size).

4. RELATIONSHIP OF BEHAVIORAL AND VERBAL MEASURES OF TRANSITIVITY AND SERIATION TO EACH OTHER AND TO OTHER PERFORMANCE MEASURES.

In all transitivity and seriation tasks in this study two types of measures are taken--what is referred to as behavioral level responses, and the verbal justifications of these responses, or logic, as it is designated here. A child is considered to be displaying transitive ability at the behavioral level if, given the two pieces of information that A > B and B > C, he can correctly infer that A > C. A child is considered to be demonstrating seriation ability at the behavioral level when he can (1) determine in some adequate manner the relationships among a number of stimuli varying along a dimension (size or weight), (2) correctly line these stimuli up in order according to these relationships, and (3) correctly insert additional stimuli into the already constructed series. These types of responses are used by many investigators of these operations (e.g., Battisti and Simmons, 1968; Inhelder and Piaget, 1941 and 1964; Murray and Youniss, 1968; Smedslund, 1963).

In addition to these measures, the subjects were asked to give verbal explanations about their transitive choices and placings of certain stimuli in the series they had formed. A subject is considered to be displaying transitivity at the "logical" level if he can state verbally that he knows A > C because A > B and B > C. "Logical" seriation is evident when the child can say a certain stimulus goes where he has (correctly) placed it in a series because it is both bigger (or heavier) than the stimulus (or all the stimuli) on one side of it and smaller (or lighter) than the stimulus (or stimuli) on the other side.

As noted earlier, Piaget feels that operations should reveal themselves through language as they develop since the linguistic tools are already present in the child for any operation and it only needs the development within the child of the operation for them to be used to reveal it. Therefore, the present research examines the relationship between the behavioral measures of transitivity and seriation and these verbal explanations about them. Further examination is made to see how each of these types of measures for each operation is related to performance on the classification and other tasks (to be described in the next section) presented to the subjects.

These two types of measures, the behavioral and the logical, were taken primarily because of a reticence on the part of the author to answer at this time the basic questions about the operations being examined, i.e., What are transitivity and seriation? Are they the logical process as expressed in verbal justifications? Or are they the behavioral achievements of giving correct behavioral responses? It was thought that at this point in our knowledge about children's thinking an adequate description of cognitive development demands consideration of both of these sources of data before answers to these questions can be attempted.

5. TWO POSSIBLE LINKS BETWEEN THE CLASSIFICATION AND TRANSITIVITY OPERATIONS.

In analyzing logically the separate structures of transitivity and classification, the author found two general bases of similarity between them. In order to demonstrate concrete operational ability in both a child has to combine two successively presented pieces of information together in order to derive a third piece of information. For instance, with the transitive operation the child must be able to put the two facts, $A > B$ and $B > C$, together in order to infer the third fact that $A > C$. In a classification task where class B is composed of subclasses A and A', the facts that an object in class A is an A and also a B are presented in succession by the experimenter and the child has to manipulate these facts in order to respond correctly to E's questions. This logical ability is designated as "successive combinativity". A more important ability was considered to be what is referred to as "simultaneous combinativity"--the ability to consider that one object possesses simultaneously two qualities, roles, relationships, etc. Thus, in a classification task, the child must understand that an object of class A is both an A and a B and that it possesses the qualities of both classes A and B at the same time. In transitivity, he must consider that object B is both $> A$ and $> C$ at the same time.

Thus, it does not seem too absurd to suggest that, if classification and transitivity do emerge synchronously it is because of the development in the child of an understanding of one or both of these combinativity operations. In order to get some independent test of this hypothesis other problems were sought which required the possession of one or other of these logical abilities for their solution. It was thought that the inference task used by the Kendlers (1962b) required successive combinativity for its completion and this and another inference task of similar structure but different content were used in this study. For simultaneous combinativity several new problems were devised to meet the specifications of the logic involved.

Thus, the final purpose of the present study is to test by independent means two hypotheses concerning the possible logical abilities underlying the attainment of the concrete operations of classification and transitivity. These are:

(a) that one of the necessary underlying logical abilities is successive combinativity, the ability to combine two successively presented pieces of information in order to infer a third piece of information.

(b) that another necessary logical ability is simultaneous combinativity, the ability to consider or accept that one object can simultaneously possess two qualities, relationships, roles, etc.

CHAPTER TWO

Method

SUBJECTS

The subjects used were 93 school children, 31 each in Kindergarten, and grades one and two attending St. Joseph's School in Hamilton. The only criteria for inclusion in the study were that a child had never repeated a grade, that he was fluent in English, and that he fell within certain age ranges set for each grade. Beyond these specifications the children were selected randomly. In the course of the study each child was tested on both Forms A and B of the Full Range Picture Vocabulary test (FRPV). The means and standard deviations of the scores of each grade are presented in Table 1.

PROCEDURE

The research consisted of nine separate tasks. Each S had six sessions of from 10 to 20 minutes long with E and one or two tasks were given per session. Testing took about three months to complete. The tasks which comprised one session were given to all 93 children before the next session

Table 1

A summary description of the sample used in the present study

	Kindergarten	Grade 1	Grade 2
n	31	31	31
Age at beginning of testing:			
Mean	5 yrs, 9 mos.	6 yrs, 9 mos.	7 yrs, 9 mos.
Range	5 yrs, 2½ mos.– 6 yrs, 2 mos.	6 yrs, 2 mos.– 7 yrs, 2 mos.	7 yrs, 1 mo.– 8 yrs, 2 mos.
# Females	15	13	15
# Males	16	18	16
FRPV scores:			
Mean	98.3	95.7	100.0
SD.	11.2	15.1	16.3

with any one child occurred. Ss were tested individually in a small, quiet room in the school. E brought S from his classroom to the testing room for each session and returned him thence when it was finished. For all the tasks S and E sat opposite each other at a low table which was placed beside a higher desk. Before each task was presented E ascertained that S knew the colours, names, shapes, or any other relevant quality of the stimuli to be used. E was also satisfied that S understood the concepts "larger than", "heavier than", and their opposites, or "equal to" where necessary.

TASK 1. TRANSITIVITY OF LENGTH - 3 OBJECTS.

Two sets of stimuli consisting of three objects each were used. The first set was composed of three crayons of different colours all 3 1/8" long and 3/8" in diameter. (A, orange = B, yellow, = C, purple). Crayon A was lying down on the table in front of S; crayon C was standing up on the adjacent desk about four feet away from A. E held crayon B in her hand and said: "Let's pretend that the orange and the purple crayons (A and C) are stuck down--we can't move them at all. I want to find out which one is longer, or if they are both the same length. So I am going to measure them both with the yellow crayon (B) to find out. Watch me." With the crayons separated and in different positions it was hoped that visual comparisons of them would be reduced. E then placed B up against A and said, "What do you see here?" S replied in words to the effect that they were the same length and E reiterated "Yes, they are both the same length. Now I'll measure the purple one (C)." E repeated the same procedure with C. E then said: "So which do you think is bigger, the orange (pointing to A) or the purple (pointing to C), or are they both the same length?" After S gave his reply, E said: "Why do you think that? What made you think that?"

The second set of stimuli in the task were three differently coloured wooden dowels all 1" in diameter but of different lengths (A, yellow 5 3/8" long >B, green, 5 1/8" long >C, red 5" long). Again S was instructed that A and C were "stuck down", A lying on the desk and C standing on the table. E measured each in turn with B making sure that S verbalized each relationship, A > B and C < B, and repeating it clearly after him. E also prompted S to say the converse of the relationship he had stated initially; i.e., if upon observing A and B side by side S declared that "A is longer than B", E would add "And B is ... than A" expecting S to fill in the missing word. Thus both S and E verbalized both relationships in two ways. S was then asked which of A and C he thought was longer or if they were the same length and why. He was then asked to designate the longest, middle, and shortest dowel. He did this either by pointing to the appropriate object or naming it by colour.

TASK 2. SERIATION OF SIZE - >3 OBJECTS

a. Preliminary training.

To insure that S understood what "lining things up in order of their size" meant, two sets of stimuli consisting of only three objects each were presented initially. The first set consisted of three wooden cubes of different colour and size: A, green, 2 5/8" on each side, > B, pink, 2" on each side, >C, white, 1 5/8" on each side. They were placed on the table in front of S in haphazard fashion and E said: "Line these three blocks up so that they make steps. Line them up so that the biggest one is first, then the middle one, and then the smallest, so that they are in order of size." If S seemed to hesitate in his procedure these instructions were repeated. If S was still confused, E gave him all necessary assistance and instruction until the three blocks were lined up correctly.

The second set of stimuli consisted of three wooden posts (1 5/8" in depth and width) of different colours and length: A, pink, 16" long >B, black, 15 3/4" long >C, blue, 15½" long. E said: "Now line these up in order of their length so that the longest one is first, then the next longest, then the shortest. Line them up so that their ends make steps." Again E gave all necessary assistance until the line-up was correct.

Only 10 Ss (eight in Kindergarten, two in Grade 1) needed E's assistance in lining up the objects correctly and all revealed understanding of the task after E had made appropriate explanations and demonstrations. All the other Ss understood what was required of them without aid.

b. Seriation task

Two sets of stimuli were used in this task. The first set consisted of 11 differently coloured wooden stairs, all 5" long but of different heights and widths so that when lined up properly they made a very colourful staircase. The end dimensions ranged from 3 3/8" square to 15/16" square, the difference between adjacent stairs being usually but not always ¼" in length and width. Nine of these were dumped on the table in front of S in haphazard fashion and S was asked to line them up in order of size so that they made stairs. E noted S procedure of doing this and his final assemblage of the objects. If S failed to line the objects up in the correct order, E lined them up for him and proceeded with the next step. E then presented in turn two more stairs, actually the fourth biggest and the seventh biggest, and asked S to insert them correctly into the line up saying: "Put this one where it should go so that they still make stairs". After each insertion S was asked why he had put the inserted stair where he did.

The second set of stimuli was composed of 10 wooden dowels, 1" in diameter, of different colours and ranging in length from 4 7/8" to 3 3/4", with 1/8" difference between adjacent dowels. Eight were dumped on the table initially to be lined up, and then two others, the third and sixth longest, were presented to be inserted into the line-up later. Again S was asked why he had placed the inserted dowels where he had.

TASK 3. TRANSITIVITY OF WEIGHT - 3 OBJECTS

So that S could become familiar with how the balance used in this task worked, S first weighed two objects in his hands to ascertain which was the heavier, then put them on the balance and noted that "the heavier one goes down, and the lighter one goes up". S was then instructed to switch the objects around to the opposite pans and to note that "no matter what side they're on the heavier one always goes down, the lighter one always goes up". This was repeated with another two objects.

Three sets of stimuli of three objects each were used in this task. The first set consisted of three small plastic toy purses of identical size (approximately 3" wide, 2" high, and 3/4" deep), but each of a different colour and design. Their contents (lead shot, plasticene, and paper) determined their weights such that purse A, white >B, blue >C, red. After S had determined what the stimuli were called and their respective colours E said: "Let's find out if some are heavier than others or if they are the same weight. You watch while I weigh them together and we'll see what happens". E weighed B and C on the balance and asked, "What do you see here?" making sure that S stated correctly that B was heavier than C and the reverse, that C was lighter than B. The same was done with A and B. The purses were set back on the table in random fashion and S was asked the following questions:

1. Which two did we weigh together first?
2. Which one was heavier?
3. Which two did we weigh together next?
4. Which one was heavier
5. Which two didn't we weigh together yet?
6. And if we did weigh these two (A and C) together what would happen?
7. How do you know? Why do you think that?
8. Which one is the heaviest of all, next heaviest, and lightest one of all?
9. Line them up on the table in order of their weight?

Questions 1 to 5 were posed to insure that S's retention of the weighings and their outcomes was correct before the questions requiring transitive reasoning (questions 6 and 7) were posed. Any incorrect responses to questions 1 to 5 were corrected by E, and in some cases where total confusion prevailed the objects were weighed again. Question 9 was given merely as a sort of preliminary training for the seriation of weight task to be described next.

The second set of stimuli used consisted of three plastic flowers each a different kind, colour and weight, but of a similar size: A, white carnation > B, red rose > C, purple iris. The procedure was the same as with the purses except that flowers A and B were weighed together first, and B and C second.

The third set of stimuli were three small gold toy trophy cups each standing 3½" high and identical in appearance except that their bases were of different colours. Lead shot and plasticene were stuck inside their bases where S could not see so that A, red base > B, black base > C, white base. The procedure with these was identical to that used with the purses.

TASK 4. SERIATION OF WEIGHT - > 3 OBJECTS.

By the time this task was given all the Ss had already done the transitivity of weight task, the last response of which requires S to line up the three objects in order of their weight, and the seriation of size task which requires S to line up many objects in order of their size. It was thought that these tasks sufficed as "preliminary training" for the present one and that no special pre-task was necessary to ensure that S understood what "to line up in order of weight" meant.

Two sets of stimuli were used in this task. The first consisted of five small plastic purses of identical size and shape (3" wide, 2 5/8" tall, 1" deep) but of different colours and patterns and weighted so that A, red > B, green > B', blue > C, purple > D, white. Four of these purses (A, B, C, and D) and the balance were set in front of S and E said: "I want you to find out which of these purses is the heaviest one of all, which the next heaviest, which the next, and which is the very lightest one of all. You can weigh them two at a time on the balance in any order and as many times as you wish until you are finished and can tell me for sure which is the first heaviest, second heaviest, third heaviest, and lightest one of all". If S did not seem to catch on immediately E repeated these instructions in various ways until S understood what he was to find out. E went on: "When you have found this out line the purses up in order of weight so that the heaviest one is first, then the next heaviest...(etc)... and the lightest is last in the line". E recorded S's procedure and his final designation of the stimuli. E pointed to the object S had designated as the third heaviest (C if he had it correct) and asked why it went where S had placed it in the line-up. E then presented the fifth purse, B', the third heaviest of the five, and said: "Here is another purse which I forgot to give you. Find out where it should go in the line-up. Do whatever you have to do so that you know for sure where this purse should go according to its weight". When S had completed his procedure and made his decision, E asked him for an explanation of why he had placed it where he did.

The second set of stimuli were eight 2½" wooden cubes weighted and coloured so that A, buff colour >B, red >C, blue >C', orange >D, green >E, white >F, pink >F', yellow.

The procedure was similar to that with the purses. Six blocks (A, B, C, D, E and F) were introduced first for S to line up according to their weights. This done, E pointed to the objects S had designated as the second and fourth heaviest (blocks B and D if the line-up was correct) and asked why he had placed them where he had. E then presented C' and F' in turn for insertion, noted S's procedure of discovering where each belonged, and then asked why he had inserted C' where he had.

TASK 5. CLASS INCLUSION: ALL-SOME

This task is one used by Piaget in his classification experiments (Inhelder & Piaget, 1964). The paper cardboard stimuli were five circles (diameter 2") and four squares (2" sq.). All the circles were blue, two of the squares were blue and two were red. These were laid in a horizontal row in front of S in random order and he was asked the following questions about them. Half the S's received the questions in the order 1 to 4 and the other half in the reverse order.

1. Are all the circles blue?
2. Are all the red ones square?
3. Are all the blue ones circles?
4. Are all the squares red?

If S did not spontaneously offer an explanation for his "Yes" or "No" responses he was asked to do so.

TASK 6. CLASS INCLUSION: PART-WHOLE

This task is very similar to that used by Inhelder and Piaget (1964) and is designed to see if a child can compare a subset with the total set of which it is a part. E arranged 16 plastic flowers in the following order (from E's point of view) in a horizontal row: 4 pink tulips, 4 yellow tulips, 2 roses, 2 carnations, 2 dahlias, 1 straw flower, and 1 bluebell. The underlying hierarchical structure of these objects can be schematized as:

C - flowers

B - tulips   B' - other flowers

A - pink tulips   A'- yellow tulips

On the table in front of S members of sub-classes are grouped together, although the subsets are not set apart from each other.

S and E together labelled all the flowers and named their colours. E then asked the following questions in the order given. S was asked to explain or give reasons for all his responses.

1. If you make a bouquet out of the tulips will you use these...? (pointing in turn to the pink tulips, yellow tulips, carnations, and roses.)
2. Are there more pink tulips or tulips here? (also phrased for all Ss as: If I made a bouquet out of all the pink tulips, and you made one out

of all the tulips, whose bouquet would be bigger or would they be the same?

3. Are there more tulips or flowers here? (also phrased for all <u>S</u>s: If I made a bouquet out of all the tulips, and you made one out of all the flowers, whose bouquet would be bigger or would they be the same size?)
4. If you pick up all the tulips will there be any flowers left on the table?
5. If you pick up all the flowers will there be any tulips left?
6. If you pick up all the yellow tulips will there be any tulips left?

TASK 7. ONE OBJECT, TWO QUALITIES

This task was designed as a test of simultaneous combinativity, i.e., to see if <u>S</u> could consider that one object possesses simultaneously two attributes or roles. To answer correctly on each trial <u>S</u> must consider some of the stimuli presented in two different ways.

The stimuli were six cards placed in front of <u>S</u> one at a time. Each card had a different set of objects pasted on it. As each card was presented <u>E</u> ascertained that <u>S</u> knew the necessary names and colours or any other relevant qualities of the objects, as well as their quantities, before questioning him about them. A description of each card and the questions asked about them follows. S was required to explain each of his responses.

Card 1. There were seven tulips in a horizontal row, three red and four yellow. Four of the tulips (two red and two yellow) had short stems, three (one red and two yellow) had long stems.

E asked: "Which are there more of

a. red flowers (3) or flower with short stems (4)?

b. yellow flowers (4) or flowers with long stems (3)?"

The number in brackets is the number in each category specified. On this card the two relevant qualities, colour and length of stems, were visually separated. That is, S could respond by first looking along the heads of the flowers for colour, and then along their stems for length. On the next three cards the two relevant attributes were not visually separated thus, but inhered in the entire gestalt of each object. It was thought that perhaps this would make these questions more difficult.

Card 2. Eight flowers were scattered on the card, five daisies and three tulips. Two of each kind were yellow, and three daisies and one tulip were white.

E asked: "Which are there more of...

a. circles (4) or big things (3)?

b. squares (4) or little things (5)?"

Card 4. Five apples and two triangles were arranged on the card. Two apples and both triangles were red, three apples were green.

E asked: "Which are there more of...red things (4) or apples (5)?"

Card 4 differs from the previous three in that one set of objects, the triangles, was homogeneous. There were only red triangles, no green ones.

Card 5. Six squares were lined up horizontally on the white card. One square was all red, three squares were all green, two squares were half red, half green.

E asked: "Which are there more of squares with red on them (3) or squares with green on them (5)?"

This card is similar to Card 1 in that the relevant qualities are visually separated within each object.

Card 6. Six female figures were represented on the card. They were all pictures cut out from a Grade 1 reader. The figures were arranged and described to S as follows:

```
                  old lady
           mother of two daughters

    daughter 1              daughter 2 (both young women)
    mother of               mother of

little girl 1          little girl 2   little girl 3
```

E asked: "Which are there more of mothers (3) or daughters (5)?"

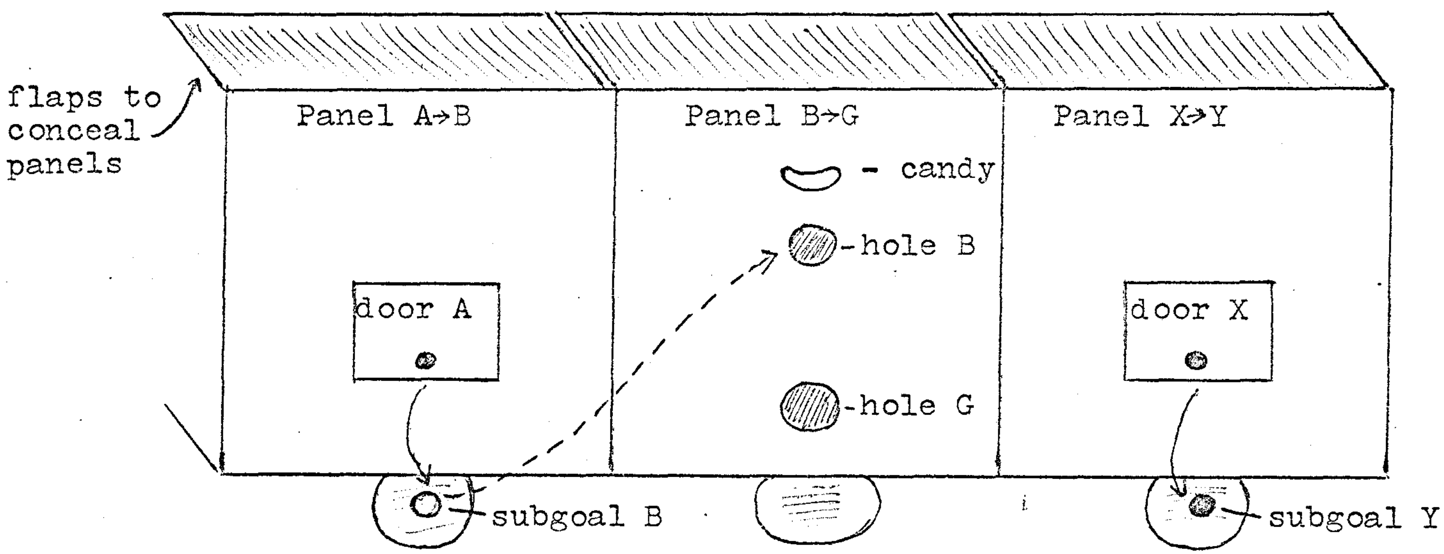
TASK 8. INFERENCE I

Tasks 8 and 9 were designed to be tests of successive combinativity, i.e., to to see whether S could put two successively presented pieces of information together in order to derive a third and crucial piece of information. It is thought that this general ability underlies the operations of transitivity and classification.

Task 8 was identical in design to the one used by the Kendlers in their research on inferential behavior (1962b). The apparatus used is illustrated in Figure 1.

S sat in front of a cardboard box the front of which was divided into three differently coloured vertical panels. E was sitting behind the box facing S. Each panel had a black cardboard flap hinged at its top which could be lowered to hide the panel or raised to reveal it. The middle panel had two holes, one about 4" above the other, in its centre. There was a candy stuck above the higher hole. The two outer panels each had a small door which S could open by pulling the door knob. Each panel had a small plastic tray at its base to catch the subgoals emanating from the doors and holes. The doors are designated as A and X, the subgoals emanating from them as B and Y respectively. Which panel was the A→B panel and which the X→Y panel was alternated: for half of the Ss it was as in Figure 1; for the other half the sides were reversed. Similarly, which of the two stimuli, a green glass marble and a silver steel ball bearing, was designated as B or Y was alternated also. The middle panel is designated the B→G panel.

Figure 1. Illustration of apparatus used in Task 8: Inference I.



The procedure was as follows:

1. Introduction of outer panels.

When S first sat in front of the apparatus all the flaps were down so that none of the panels could be seen. E opened the flap covering panel A → B and said, "Pull this door open and see what happens". S did so and subgoal B (the green marble for half the Ss, the ball bearing for the other half) dropped out. E said, "Pick up that thing and look at it... Now give it back to me and you can have another turn. Pull the door again--see here's the same thing again. Give it back to me". Panel A → B was closed and the procedure was repeated with panel X → Y. E never labelled either of the subgoals.

2. Training on panels A → B and X → Y.

With both the outer panels open to S's view, E held up subgoals B and Y in random order each time saying, "Open the door that will get you one like this". This continued until S got six successive trials correct. The two outer panels were then closed, and the middle panel was revealed for S.

3. Training on panel B → G.

E pointed to the candy stuck on the front of the panel and said, "See this candy? You may get one for yourself. If you drop the right thing through this hole (B) a candy will come out of this hole (G) and fall into the dish". E held out the subgoals B and Y in her hand: "One of these two little things is the right one. Take one of them and drop it in the hole to see if it makes candy come out". S did this until he put the correct subgoal (B) into the upper hole, whereupon E made a candy, the major goal, come out of the lower hole (G). S was encouraged to repeat this performance of choosing either B or Y and dropping it through the upper hole until he correctly chose and dropped B through the hole and received the candy on four successive trials. S was allowed to keep and eat the candy only after the fourth correct trial. The left-right position of subgoals B and Y in E's hand as she offered S to choose one was randomized.

4. Test of Inferential Behavior, or Successive Combinativity.

Now S was tested to see if he could string together the two pieces of previously learned behavior, i.e., that opening the door in panel A produced subgoal B, and that inserting B into the upper hole of the middle panel would produce the major goal, a candy. This is called an "integration response". S must ignore panel X → Y in this operation. All panels were opened and E said: "Would you like another candy? This time I won't put out any little things in my hand, but I will open all the shutters. If you do what you are supposed to do, you can make a candy come out. Go ahead."

S was given 60 seconds in which to open one of the doors. If he did not E asked, "Which door should you open to help get a candy?" S was then given another 60 seconds to respond. If he did not the task was terminated.

E recorded S's first and second choice of doors, whether or not he made an integration response, and the time taken for each of his responses.

TASK 9. INFERENCE II.

This task has a similar structure to task 8 with the following differences: there is nothing comparable to panel X → Y in task 8, and the correct performance of the task depends on a bit more verbalization on S's part. This task was designed to be a more direct test of successive combinativity than task 8 in which it was thought that the presence of the X → Y panel may have been distracting.

On the table in front of S were put a toy garage, a green truck which would not fit through the garage door, and a blue racing car which would. The procedure was as follows:

1. Training with A → B response.

E said: "Let's pretend that this is your green truck. Any time you give me your green truck, I'll trade you this blue racing car for it. Go ahead, give me the green truck... Here's the racer for it." This trade was made four times and each time E asked a question like, "If you give me the green truck what will you get?" or "If you wanted me to give you the blue racer, what would you have to give me?" etc. At the end of the fourth trial S kept the blue racer, and E kept the green truck.

2. Training with B → G response.

E said: "Here's a garage with a little doorway. Whenever you can drive a car or anything through the doorway into the garage, you'll get a candy. Try the racer and see if it will fit through the doorway. If it does you'll get a candy". S did this for four trials with on each trial E asking a question like, "If you put the racer through the doorway into the garage what will you get?" or "If you wanted to get a candy right now what should you do?" After S gave his verbal answer to these questions he was allowed to drive the racer into the garage and received a candy. He was only allowed to keep and eat the candy after the fourth trial. After the fourth trial E retrieved the racer.

3. Test for inferential behaviour, or successive combinativity.

E said: "Would you like another candy?... Here's your green truck back again. Remember, whenever you can drive anything through the doorway into the garage you'll get a candy. See if you can get another candy." The correct response was for S to trade his green truck for E's blue racer and then insert the racer into the garage. All of S's behavior at this point was recorded. He was allowed 60 seconds to make an initial response. If he had done nothing in this 60 seconds E said, "Is there anyway I can help you get a candy?" S was then given another 60 seconds in which to respond. If he did nothing again in this time the task was terminated.

ORDER OF PRESENTATION OF THE TASKS.

The tasks were given in six sessions in the order outlined below. The tasks are numbered here as they are above.

Session 1. 5. Class Inclusion: All-Some
8. Inference I

Session 2. 1. Transitivity of Length - 3 Objects

Session 3. 3. Transitivity of Weight - 3 Objects
6. Class Inclusion: Part-Whole

Session 4. 2. Seriation of Size - > 3 Objects
7. One Object, Two Qualities

Session 5. 4. Seriation of Weight - > 3 Objects

Session 6. 6. Inference II.

The FRPV intelligence test was given at the end of the sixth session.

CHAPTER THREE

Results

The results of each task will be presented separately first before the relationships between the tasks are considered.

TASK 1. TRANSITIVITY OF LENGTH - 3 OBJECTS.

RESPONSE MEASURES.

For each of the two sets of stimuli we have the following measures:

1. Behavioral Transitivity (BT). S's prediction as to which object, A or C, was longer or if they were the same length.
2. Logic. S's reasons as to why or how he came to this conclusion. In both transitivity tasks in this study, for an explanation to be considered logical S had to make verbal reference to both relationships relevant to the transitive conclusion, i.e., he had to indicate somehow both that A>B, and that B>C, and that that's how he knew A>C. Scoring an explanation as logical is quite straightforward--S either made reference to these two relationships or he did not. If S made reference to only one of the relationships E would

question him about this further to see if he was implying the other relationship as well or not. For instance, an S might say, "A was longer than B so it's longer than C too". E would then say, "Just because A is longer than B does that mean it will be longer than C too? Why?" S might respond affirmatively and leave it at that, in which case his explanation would be scored as a non-logical; or he might say, "No, but B was longer than C too, so A must be longer than C." This explanation would be scored as logical. This procedure was always followed whenever only one of the two relevant relationships was alluded to by S.

3. Verbal Seriation (VS). For the second set of stimuli of unequal lengths we also have S's designation by pointing or naming of the longest, middle and shortest objects. S does not actually move any of the objects. All three objects must be designated correctly for VS to be counted as correct.

DESCRIPTIVE RESULTS

Table 2 is a summary of the percentage of correct responses given in this task. Since there were no differences between the two sets of stimuli in the percentage of correct BT responses or in the percentage of logical explanations given, these percentages have been combined to yield one BT percentage and one Logic percentage for each grade. The percentage of correct BT responses is high in every grade. Only the difference between Kindergarten and Grade 2 is significant

Table 2

Descriptive results for Task 1: Transitivity of Length - 3 Objects

	Percentage of Correct Responses			
	Kindergarten	Grade 1	Grade 2	Mean
1. Behavioral Responses Behavioral Trans. (BT) (mean of 2 sets of stimuli)	80.6	91.9	100	90.9
Verbal Seriation (VS) (unequal set)	74.2	77.4	100	83.9
BT & VS both correct (unequal set)	67.7	77.4	100	81.7
2. Logic (mean of 2 sets of stimuli)	24.2	64.5	93.5	60.8

($p < .01$, using a two-tailed test of significance between proportions for independent samples). The "BT and VS both correct" measure is the percentage of Ss who had both the Behavioral Transitivity and Verbal Seriation responses correct with the second set of stimuli which were of unequal lengths. For both the Verbal Seriation alone percentages and the BT and VS combined percentages, the differences between Grade 2 and each of the other two grades are significant (all p's $< .01$, two-tailed). The Kindergarten-Grade 1 differences are not significant.

With the unequal set of objects in several instances an S predicted correctly that A>C but then failed to seriate the three objects A, B, and C in the correct order. This is designated a Type A error--correct BT but incorrect VS. Three Ss in Kindergarten made this type of error, and three Grade 1 Ss made it. With five out of the six of these errors made the ordering of objects A and C is correct, that is, S acknowledges that A is still longer than C, but object B is designated incorrectly as either the longest or the shortest object. Only one S (in Grade 1) indicated that C was now longer than A, having said the opposite for his BT response. Therefore, this type of error seems mainly due to confusion with regard to the "middle" object rather than blatant contradictions in S's thinking. A type B error occurs when S incorrectly predicts that C>A but then verbally seriates the three objects in the correct order--incorrect BT but correct

VS. Two type B errors were committed, both by Kindergarten Ss. With this type of error S's thinking is contradictory since he says first that C>A and almost immediately after that A>C.

Since it occurs on only three occasions that the VS response is blatantly contradictory to the BT response with the unequal set of stimuli, this error seems insignificant in this task. Therefore, only the BT responses are considered hereafter when dealing with behavioral level performance on this task.

For the Logic responses all the differences in percentage of the logical explanations given between grades are significant at the .01 level or better. With the Grade 2 Ss only one S failed to give at least one logical explanation for his predictions. With the Kindergarten Ss only three Ss could give two logical reasons for their two BT responses. Logical performance is not as good as Behavioral performance in any of the grades and the differences between the BT and Logic percentages are significant (both p's <.01) for the Kindergarten and Grade 1 Ss. The small difference in Grade 2 is not significant.

CATEGORIES OF PERFORMANCE.

On each of the transitivity and seriation tasks in this study the Ss are divided into Categories of Performance on the basis of their behavioral and logical performances respectively on that task. It is then possible to compare a S's performance on the different tasks by comparing his Categories of Performance on these tasks.

1. LOGIC

In the transitivity of length task S was asked to give two explanations for his Behavioral Transitivity responses, one for each set of stimuli, equal and unequal. Thus an S could give two, one, or no logical explanations. The Categories of Performance decided upon were Good, Some, and None; they are defined and their distributions are given in Table 3. Over half the sample has Good logic, the majority of these being in Grade 2. There is a sort of all-or-none effect, i.e., most Ss either show perfect logic or no logic at all, with only a few Ss showing Some logic. There is a significant Grade by Category relationship in that the higher the grade the more logic displayed ($\chi^2 = 41.03$, $p < .001$).

2. BEHAVIOR

The Behavioral Transitivity measure is the only one considered in dividing the Ss into the Categories of Performance for behavior. There is a certain probability that S could get a BT response correct by chance alone, and this is why there are more than one set of stimuli and thus more than one BT response requested on the two transitivity tasks in this study. In designating the Categories of Performance, Good transitivity will be considered to exist only in those Ss who get all the BT responses correct.

Table 3

Logic Categories of Performance for Task 1: Transitivity of Length - 3 Objects

Description	Score # logical expl.	Distribution K	1	2	Total
Good	2	3	17	28	48
Some	1	9	4	2	15
None	0	19	10	1	30
Total		31	31	31	93

Table 4 presents the definitions and distributions of Categories of Performance for behavior for this task. Behavioral performance in this task is very good, with a large majority of Ss making perfect predictions. Only two Ss had no BT responses correct and they are both in Kindergarten (and they both displayed no logic also). The Grade by Category relationship is significant ($\chi^2 = 11.92$, $p < .01$). (For this task only there is no Fair category for behavior. This is because the frequencies in the Fair and Poor categories, 13 and 2 respectively, would have been too low to permit statistical comparison with other tasks.)

3. RELATIONSHIP BETWEEN LOGIC AND BEHAVIOR

Table 5 is a contingency table showing the relationship between the Categories of Performance for logic and behavior for Ss in all three grades combined. The relationship is highly significant ($\chi^2 = 22.03$, $p < .001$). Nobody with Good logic has Poor behavior, and the majority of Ss with Poor behavior reveal no logic at all. Thus in this task an S's logical ability is highly correlated with his behavioral responses. If his logic is perfect, his behavior is also perfect. If his behavior is poor, chances are he lacks logic in explaining his responses. On the other hand, even if he gives no logical explanations there is a good chance that his behavioral responses were all correct nevertheless. Thus there is a sizeable group of Ss (18) whose behavior is perfect but who reveal no logic. Their performance on other

Table 4

Behavior Categories of Performance for Task 1: Transitivity of Length - 3 Objects

Description	Score #BT correct	Distribution K	1	2	Total
Good	2	21	26	31	78
Poor	0.1	10	5	0	15
		31	31	31	93

Table 5

Relationship between Logic and Behavior Categories of Performance for Task 1: Transitivity of Length - 3 Objects.

		Logic Good	Some	None	Total
Behavior	Good	48	12	18	78
	Poor	0	3	12	15
	Total	48	15	30	93

tasks will be examined later to see if this is a consistent pattern of performance for these Ss or if it occurs with them only on this task.

TASK 2. SERIATION OF SIZE - >3 OBJECTS

RESPONSE MEASURES

For each of the two sets of stimuli the following response measures were recorded:

1. Behavioral Seriation: Whether or not S correctly lined up the objects, no matter how he went about doing it.[1]

2. Behavioral Insertions: Whether or not S inserted the two additional objects into each set correctly, no matter how he went about it.[1]

3. Logic: Whether or not S gave logical or non-logical explanations in response to E's query as to why he had inserted each of the two objects in each set where he did. For

---

[1] Neither the S's manner of going about lining up the objects initially nor his method of finding out where to insert the extra objects into the line-up is considered in this task. It is very difficult to discriminate between true operational behavior and merely good perceptual discrimination when the dimensional differences between the stimuli are visible as here. Also, the types of trial and error behavior are many, ranging from one or two errors of placement which are quickly corrected by S, to total hit and miss behavior which, after much shuffling, finally results in a correct line-up or insertion. Furthermore, there was only one observer present to record all of S's behavior and make the necessary judgements about which type of behavior S was displaying. For these reasons the reliability of the mode of attack measures would be open to doubt and thus it was thought best to abandon them.

an explanation to be scored as logical the following criteria must be met: S must mention in some fashion that the object he inserted is both smaller than the adjacent larger object or objects preceding it in the line-up and larger than the adjacent smaller object or objects following it in the line-up. That is, he must mention two relationships: the inserted object is both larger than some objects and smaller than others. As in the transitivity tasks, if S mentioned only one of these relations and stopped, E kept up an interaction by asking him questions (e.g., "Just because it is bigger than this one, is that the only reason it goes in here?") to see if S was aware of the other relation but had failed to mention it at first. This measure is very easy to score, as are the two behavioral responses.

DESCRIPTIVE RESULTS

Table 6 presents the percentage of correct responses given in this task. Since there were no significant differences between the two sets of stimuli in the percentage of correct responses with any of the response measures the results for the two sets have been combined to yield one percentage for each response measure in each grade. The percentage of correct behavioral seriations increases with grade, being almost perfect in Grade 2. The differences between Kindergarten and the other two grades are significant (both p's $< .01$ in a two-tailed test of the difference between the two proportions for

Table 6

Descriptive results for Task 2: Seriation of Size - > 3 Objects

	Percentage of Correct Responses			
	Kindergarten	Grade 1	Grade 2	Mean
1. Behavioral Responses Seriations (mean of 2 sets of stimuli)	54.8	85.5	96.8	79.0
Insertions (mean of 2 sets of stimuli)	74.2	94.4	99.2	89.2
2. Logic (mean of 2 sets of stimuli)	16.9	59.7	75.8	56.5

for independent samples). The Grade 1-Grade 2 difference is not significant.

Insertion ability also improves with increasing grade. Again the differences between Kindergarten and the other two grades are significant ($p < .05$ for the Kindergarten-Grade 1 difference; $p < .01$ for the Kindergarten-Grade 2 difference), whereas the Grade 1-Grade 2 difference is not significant. In all three grades the percentage of correct insertions is higher than the percentage of correct seriations. Although none of these differences attains significance this result is very interesting. In the procedure if S did not seriate the objects correctly initially E did it for him and then gave him the objects to be inserted. Therefore it was possible, and indeed occurred, that S could not seriate the objects correctly himself but could correctly insert objects into an already constructed series. Other experiments in the literature on seriation (e.g., Battisti and Simmons, 1968; Murray and Youniss, 1968) usually terminate a trial if S fails to seriate the objects correctly by himself. Thus they have not examined the relationship between the abilities to insert objects and to seriate them initially. It seems to be easier for the Ss, especially for the Kindergarteners, to deal with a situation already structured for them than to structure it initially for themselves.

Logical performance also improves with grade, and again only the difference between Grades 1 and 2 is not signifificant (both other p's $<.001$).

CATEGORIES OF PERFORMANCE

1. LOGIC

S had four opportunities to give logical explanations: with each of the two sets of stimuli he was asked to explain why he inserted the two extra objects into the series where he had. Table 7 presents the Categories of Performance for these logical explanations with definitions and distributions. The higher the grade, significantly more logic is displayed ($\chi^2 = 30.57$, $p < .001$). Again, as with the transitivity task, there is a sort of all-or-none effect with regard to logic: most Ss either display perfect or near-perfect logic or none at all. The six Ss with three out of four logical explanations were grouped in the Good category because it was judged that their one non-logical explanation was a lapse, i.e., it was nearly logical but not quite, being not completely off base.

2. BEHAVIOR

According to the definition of correct seriation performance at the behavioral level employed in this study an S must not only be able to line objects up correctly according to the relationships he has discovered among them, but he must also be able to correctly insert additional objects into this line-up. Therefore, both behavioral seriation and insertion responses were considered when devising the Categories

Table 7

Logic Categories of Performance for Task 2: Seriation of Size - >3 Objects

Description	Score # logical expl.	Distribution K	1	2	Total
Good	3, 4	3	18	23	44
Some	1, 2	5	3	4	12
None	0	23	10	4	37
		31	31	31	93

of Performance for behavior for this task. Table 8 presents these categories for the behavioral responses combined--the two initial seriations of the stimuli, and the four insertions of the extra objects. Most of the Grades 1 and 2 Ss had perfect performances on this task but less than 25 percent of the Kindergarteners did. (Note that an S could be categorized as Fair if he had four behavioral responses correct and all four of these could have been insertions only; i.e., he could have failed to line up both sets of stimuli in the first place and yet be categorized as Fair. This only occurred with two Ss, both in Grade 1). The Grade by Category relationship is significant ($\chi^2 = 32.75$, $p<.001$). In general, behavioral performance is quite good on this task, with nearly two-thirds of the Ss (60/93) giving perfectly correct responses.

3. RELATIONSHIP BETWEEN LOGIC AND BEHAVIOR.

Table 9 is a contingency table showing the relationship between the Categories of Performance for Logic and Behavior for Ss in all three grades combined. The relationship between logic and behavior is highly significant ($\chi^2 = 30.43$, $p<.001$). But note that behavior is much better than logic - 33 Ss lie above the diagonal from the Good to Poor-None cells, and only five lie below it. Many Ss displayed only some or no logic but still had perfectly correct behavioral responses, whereas nobody who had Good or Some logic turned in a Poor behavioral performance. As in the previous transitivity task, logical ability seems to develop later than behavioral ability in this task.

Table 8

Behavior Categories of Performance for Task 2: Seriation of Size - >3 Objects

Description	Score total # sertns & insertns correct	Distribution K	1	2	Total
Good	6	8	24	28	60
Fair	4, 5	14	5	3	22
Poor	0- 3	9	2	0	11
		31	31	31	93

Table 9

Relationship between Logic and Behavior Categories of Performance for Task 2: Seriation of Size - >3 Objects.

		Logic Good	Some	None	Total
	Good	39	7	14	60
Behavior	Fair	5	5	12	22
	Poor	0	0	11	11
	Total	44	12	37	93

TASK 3. TRANSITIVITY OF WEIGHT - 3 OBJECTS

RESPONSE MEASURES.

The following measures were recorded in this task:

1. Behavioral Transitivity (BT): S's response to the question, "If we weighed A and C together what would happen?" In other words, S's prediction about the outcome of weighing A and C together after he had seen A and B and then B and C weighed together.

2. Logic: S's explanation as to why he made the BT response he did. In order for this response to be judged "Logical" S must make explicit reference to the outcome of the two weighings which E made and indicate that these are the reasons for his saying that A>C. For example, he would say something to the effect that "A was heavier than B, and B was heavier than C, so A will be heavier now." Instead of "heavier than" he could say "went down with" (i.e., on the balance) etc.

3. Verbal seriation (VS): S was asked to point to or name the heaviest, next heaviest, and lightest objects without moving them. All three objects must be correctly designated for the VS response to be correct.

DESCRIPTIVE RESULTS

Table 10 presents the percentages of the responses which were correct. The results for the three sets of stimuli have been combined since there were no differences between them. The percentage of correct BT responses increases with

Table 10

Descriptive results for Task 3: Transitivity of Weight - 3 Objects

	Percentage of Correct Responses			
	Kindergarten	Grade 1	Grade 2	Mean
1. Behavioral Responses				
Behavioral Trans (BT)	67.7	80.6	84.9	77.8
Verbal Seriation (VS)	50.5	63.4	63.4	59.1
BT and VS both correct	41.9	59.1	63.4	54.8
2. Logic				
% logical explanations	1.1	24.7	41.9	22.6

the grade. The differences between Kindergarten and both Grades 1 and 2 are significant ($p < .05$ and $p < .01$ respectively), but that between Grade 1 and Grade 2 is not. The VS responses also improve between Kindergarten and Grade 1 ($p < .05$) but remain the same in Grades 1 and 2. In all the grades the percentage of correct BT responses is greater than the percentage of correct VS responses and all these differences are significant (all p's $< .01$). Similarly in all the grades the percentage of trials on which the BT and VS responses were both correct is significantly lower than the percentage of trials on which the BT response alone was correct (all p's $< .01$). The discrepancies among the percentages of correct responses of these three behavioral measures reveals that on one trial a BT response could be correct while the VS response was wrong (obvious from the discrepancy between the BT and VS scores alone), and also that a BT response could be incorrect while the VS was correct (evident from the discrepancy between the VS alone score and the BT and VS combined score).

These behavioral errors deserve closer examination. A type A error occurs where the BT response was correct but the VS was wrong. A Type B error occurs where the BT response was incorrect but the VS correct. Table 11 presents the frequency with which each type of error was made in each grade. The number of Type A errors made does not decrease appreciably with age. Type A errors occur on 22.9 percent of the total

Table 11

Behavioral errors made in Task 3: Transitivity of Weight - 3 Objects

Frequency of Errors (n = 93 Trials)

	Kindergarten	Grade 1	Grade 2	Total (279 Trials)
A. BT correct, VS wrong Total	24	20	20	64
1. A-C order correct	19	18	18	55
2. A-C order reversed	5	2	2	9
B. BT wrong, VS correct Total	8	4	3	15

trials. But the ordering of objects A and C is correct in 85.9 percent of these errors; that is, S indicates that A is heavier than C but says that B is either the heaviest or the lightest of all three. The A - C order is reversed so that S's ordering is either C>A>B, B>C>A, or C>B>A, in 14.1 percent of these errors. Type B errors are made on 5.4 percent of the total trials.

Thus on 71.7 percent of all the trials (100-22.9-5.4) BT and VS are both correct or both incorrect. (We know from Table 10 that they are both correct on 54.8 percent of the trials; therefore, they are both incorrect on 16.9 percent of the trials). On the 28.3 percent of the trials that BT and VS outcomes do not concur, chances are greater than 4 to 1 (64:15) that the BT was correct and the VS wrong. Of the total number of trials where one of these responses was correct and the other incorrect, the ordering of objects A and C was contradictory in the two responses on only 8.6 percent of the trials (all the Type B errors plus the Type A errors #2). Therefore, here as in the transitivity of length task, BT and VS responses are contradictory on an insignificant number of trials and we are justified in ignoring the VS responses hereafter when considering behavioral level performance on this task.

In Table 10 the percentage of logical explanations increases with increasing grade, the differences between all grades being significant (all p's <.01). In each grade the percentage of each of the behavioral responses correct is much greater than the percentage of logical explanations given. All these differences between behavior and logic are significant (all p's <.01).

CATEGORIES OF PERFORMANCE.

1. LOGIC.

Table 12 presents the definitions and distributions of the logical Categories of Performance for this task. The higher the grade, significantly more logic is displayed ($\chi^2 = 21.31$, $p < .001$). In fact only one Kindergarten S gave one logical explanation; all the rest gave non-logical reasons for their BT responses. Logical performance is not very good on this task: only 29 percent of the sample show any logic at all.

2. BEHAVIOR

Only the BT responses are considered in devising the Categories of Performance for behavior. Since there is a certain probability that the BT response can be correct by chance alone, it is necessary to have several trials and to categorize so that most or all of the BT responses must be correct for an S's performance to be considered Fair or Good. Table 13 presents the definitions and distributions of the behavioral Categories of Performance for this task. Behavioral performance is quite good with over half of the sample (around 53 percent) getting all the BT responses correct and

Table 12

Logic Categories of Performance for Task 3: Transitivity of Weight - 3 Objects

Description	Score # logical expl.	Distribution K	1	2	Total
Good	3	0	7	10	17
Some	1, 2	1	2	7	10
None	0	30	22	14	66
		31	31	31	93

Table 13

Behavior Categories of Performance for Task 3: Transitivity of Weight - 3 Objects

Description	Score # BT correct	Distribution K	1	2	Total
Good	3	13	16	20	49
Fair	2	8	12	8	28
Poor	0, 1	10	3	3	16
		31	31	31	93

only 17 percent having a Poor performance. Behavior does improve with grade but the relationship is not significant ($\chi^2 = 8.78$, $p < .10$).

3. RELATIONSHIP BETWEEN LOGIC AND BEHAVIOR.

Table 14 presents the distribution of the Ss among the logic and behavior categories in contingency fashion with the grades combined. The relationship between logic and behavior is highly significant ($\chi^2 = 22.28$, $p < .001$). Obviously though, behavior is far superior to logic on this task (58 Ss lie above the Good - Good to Poor - None diagonal and only one S lies below it). There are no Ss with Good logic but Poor behavior, but the reverse is found--Ss with no logic and Good behavior. There is only one S with Some logic but Poor behavior. So it would seem that the display of logical reasoning would indicate an S whose behavior was well above chance performance. The opposite is not true, however. Just because an S displays no logic does not mean that his behavior is poor--he could have perfect, fair, or poor behavior. In fact the chances are that his behavior will be Fair or Good. The performances of these Ss on other tasks will be examined in a subsequent section.

TASK 4. SERIATION OF WEIGHT - >3 OBJECTS.

RESPONSE MEASURES.

For both sets of stimuli the following response measures were recorded:

Table 14

Relationship between Logic and Behavior Categories of Peformance for Task 3: Transitivity of Weight - 3 Objects.

		Logic			
		Good	Some	None	Total
Behavior	Good	17	7	25	49
	Fair	0	2	26	28
	Poor	0	1	15	16
	Total	17	10	66	93

1. Seriation Mode of Attack: how S went about discovering the relationships among the objects. This could range from weighing them all in pairs on the balance in a systematic and thorough manner and then lining them up in order, to lifting some of them up in one's hands and then arranging them as one saw fit, with several intermediate modes of attack in between. The only quality assessed about the procedure was whether or not they were "Adequate" to discover the relationships among the objects and hence their correct order. In order for a weighing procedure to be "Adequate", S must have weighed on the balance at least each adjacent pair of stimuli, whatever else he weighed as well. For instance, with the first set of stimuli where four objects are presented to be lined up initially--A>B>C>D--S must weigh the adjacent pairs A-B, B-C, and C-D in order to discover all the relevant relationships among the objects. He probably will also have weighed other pairs together in his procedure--e.g., A-C, or B-D etc., but if his procedure included the weighing of the three adjacent pairs it was labelled "Adequate"; if it did not include these pairs it was "Inadequate". With the second set of stimuli where six objects are to be lined up initially--A>B>C>D>E F--S must again at least weigh adjacent pairs of stimuli together for his procedure to be "Adequate".

2. Behavioral Seriation: Whether or not S lines the objects up in correct order after he has completed his procedure of discovery, whatever it may have been.

3. Insertion Mode of Attack: S's procedure of finding out where in the line-up the insertion object should go. This ranged again from correctly weighing the insertion object on the balance with the appropriate objects already in the line-up until he discovered between which two it belonged, to simply shoving the object somewhere into the line-up without weighing it or comparing it at all. Again the weighing procedure was categorized as "Adequate" or "Inadequate". In order for it to be "Adequate" S must have compared the insertion object with enough of the others to be able to discover its relationship to them and therefore where it belonged in the line-up. For instance, with the first set of objects $A>B>C>D$, B' is the insertion object and it belongs between B and C. In order for S to know this he must have weighed B' together with objects B and C at least. For his weighing to be "Adequate" then, these two pairs, B'-B and B'-C, must have been weighed. If they were not his weighing procedure was "Inadequate". With the second set of stimuli, insertion object C', which belongs between C and D, must be weighed at least with these two objects for an "Adequate" procedure. In fact most S's starting comparing C' from one end of the line-up or the other, and so many more pairs than the two crucial ones were actually weighed. With insertion object F' which belonged at the lightest end of the series, S had to weigh it only with F to discover its proper place and have an "Adequate" procedure.

4. Behavioral Insertion: Whether or not S inserted these objects in the appropriate place, whatever mode of attack he had used to discover this.

5. Logic with Seriation: When S had lined up the objects as he saw fit E asked him how he knew that certain of the objects should go where he had put them. With the purses, the first set of stimuli presented, the object inquired about was C, or whichever one S had placed third. With the blocks, enquiries were made about objects B and D, or whichever ones S had placed second and fourth.

6. Logic with Insertions: After S had inserted the insertion object E asked him how he knew it belonged wherever he had put it. With Set 1 the insertion object was B'. With Set 2 there were two insertion objects, C' and F', but only C' was inquired about.

In Task 2 the definition of a logical explanation stated that S must make reference to the two relevant relationships which determined an object's placement in the line-up. The same definition is employed here. Therefore, a logical explanation is one which includes mention of the two adjacent relations of the object inquired about--e.g., for insertion B' in Set 1 S must say words to the effect that $B' < B$ and $B' > C$ as reason for its being placed between B and C; or for object B in the original line-up S must say that $B < A$ and $> C$ and so it belongs in second place, or words to this effect.

DESCRIPTIVE RESULTS.

Table 15 presents the percentage of correct responses for the various response measures in this task. Since there were no differences between the two sets of stimuli on any of the measures their results have been combined in this table. For behavioral seriations, both the percentage of adequate weighings and the percentage of correct seriations increase significantly between Kindergarten and Grade 1 (both p's $<.01$) and then stay stable through Grade 2. Similarly, with the insertions, the biggest improvement in the percentage of adequate weighings and of correct insertions occurs between Kindergarten and Grade 1 (both p's $<.001$), with smaller improvements between Grades 1 and 2 (for Adeq. Wghings $p<.05$; Insertions is nonsignificant). With seriation behavior in all grades the percentage of correct seriations is either the same as or slightly greater than the percentage of adequate weighings. With insertion behavior in all grades the percentage of adequate weighings is slightly greater than the percentage of correct insertions.

Table 16 examines the errors made which cause these discrepant percentages within each grade. Type A errors are those where the weighing procedure was Adequate but the seriation or insertion was incorrect. These occur least frequently in Kindergarten because of the very few Adequate weighings in the first place compared with the other two grades. Type B errors are those where the weighing procedure was Inadequate

Table 15

Descriptive results for Task 4: Seriation of Weight - >3 Objects

	Percentage of Correct Responses			
	Kindergarten	Grade 1	Grade 2	Mean
1. Behavioral Responses				
a) Seriations				
Trials with Adeq Wghings	6.5	53.2	58.1	39.3
Trials with correct sertns.	12.9	59.7	58.1	43.7
b) Insertions				
Trials with Adeq Wghings	17.2	76.3	88.2	60.6
Trials with correct Instns.	14.0	67.7	73.1	51.6
c) Sertns correct given Adeq Wghings	1.6	46.8	53.2	33.9
d) Insertns correct given Adeq Wghings	10.8	63.4	72.0	48.7
e) Mean Sertns and Insertns correct given Adeq Wghings	7.1	56.8	64.5	42.8
2. Logic				
Seriations	4.3	53.8	76.3	44.8
Insertions	3.2	50.0	72.6	41.9
Mean	3.9	52.3	74.8	43.7

<u>Table 16</u>

Behavioral errors made in Task 4: Seriation of Weight - >3 Objects

Type of Error	Frequency of Errors			
	Kindergarten	Grade 1	Grade 2	Total
Type A - Wghing Adeq but incorrect.-				
Seriation	4	4	3	11
Insertion	3	8	9	20
Total	7	12	12	31
Type B - Wghing Inadeq but correct.-				
Seriation	7	8	3	18
Insertion	3	3	0	6
Total	10	11	3	24

but S got the seriation or insertion correct anyway. This is really chance behavior--S had no way of knowing the correct order of the objects if his procedure was Inadequate and so his final outcome was merely fortuitously correct. There are many fewer of these errors made by Grade 2 Ss than by others. On 83.3 percent of the total number of trials there is agreement between the weighing procedure and the final outcome--i.e., an adequate procedure leads to a correct seriation of the objects, or an inadequate procedure leads to an incorrect one. For the insertion performance this percentage of agreement between the procedure and outcome is 90.2 percent.

In general, performance is better with the insertions than with the original seriations. A similar finding occurred in the length dimension. It was possible there because if S could not line the objects up correctly himself, E did it for him and then requested him to insert the extra objects, which he could often do correctly even though he was unable to seriate them initially. With this weight dimension task not one S thought he was unable to line the objects up correctly. All Ss made some sort of line-up and because there was no visual feedback as to is accuracy, all assumed that their line-up was correct. Many, however, were not, as can be seen in Table 15. Nonetheless it was possible to insert the insertion objects correctly despite the fact that the line-up was not wholly correct. For instance, with the first set of stimuli S could have lined the objects up thusly: B A C D, which of course is wrong. However, he could insert object B' between A and C and be perfectly correct. That is, if he started comparing B' with B he would see that B'<B. He would then go on to A and see that B'<A also. When he weighed B' with C and saw B'>C he would conclude that B'>C but <A and therefore belongs between them. His logic would be impeccable and his positioning of B' perfectly correct. However, if he had lined up the original four objects thus: A C B D, C A B D, or in several other ways he could not possibly insert B' appropriately unless he changed the order of the other four objects. This often occurred in which case S could again get the insertion correct even though the original seriation based on his original weighing procedure was wrong. Thus, when we look at the most crucial behavioral statistics of all, the percentage of correct seriations and insertions given that the procedures were Adequate (Table 15 rows c and d) we see that the percentages for insertion behavior are higher than for seriation behavior in all three grades (p's for Kindergarten and Grade 1 both < .05; p for Grade 2<.01). This finding reveals something about the differences in ability when S is required to organize the objects in his environment himself as compared to when they are already organized for him and he has merely to operate upon them in some fairly well prescribed manner.

The percentage of logical explanations given for both seriations and insertions increases significantly with grade, all between grade differences being significant (all p's $<.01$ or better).

If we compare the mean percentages for behavior and logic--i.e., the mean percentage of correct seriations and insertions given adequate weighings (row e) and the mean percentage of logical explanations given, we see that the overall percentages are close (42.8 and 43.7 respectively) and that in Grade 2 more logic is displayed than there are correct behavioral responses to explain. (None of the within-grade differences is significant). This is the only task in which many S's logical performance surpass their behavioral performance. This is possible because an S can give a logical explanation for an incorrect behavioral performance. For example, suppose with the first set of stimuli S has lined the objects up incorrectly--A C B D. E would then ask him to explain why he had placed B third in line and he might say because $B < C$ (which it was not) and $B > D$ (which it was) and so it should go between them. He may not have actually weighed B and C, or he may have weighed them but have forgotten the outcome. In any case, he has placed B and C incorrectly. But he has made a line-up--he has imposed order on the objects--and when asked anything further about the relationships among the objects he reasons from the order he has derived and not from the original weighings which gave him the information to discover this order.

This could only occur in the weight dimension where there is no visual feedback: S does not know his line-up is incorrect. In the length dimension he can see when he has the objects correctly seriated and inserted so he would only give logical explanations about correct behavioral responses. It does not occur in the transitivity of weight task because S does not actually impose physical order on the three objects before he is asked to explain his BT response. The three objects are not lined up in order of their weight, that is, in the order he thinks they should go in, but are scattered on the table in front of him. Therefore, he has only his memory of the outcome of the original weighings to go by when explaining his behavioral response.

Thus there exist children who fail to organize the objects adequately in such a way initially so as to order them correctly, but who once having imposed some order, even though it is incorrect, on the objects can then reason in a perfectly logical manner about the already ordered situation. It seems they can operate perfectly correctly in a structured situation but are at a loss in an unstructured one.

CATEGORIES OF PERFORMANCE

LOGIC.

Table 17 presents the definitions and distributions of the logical Categories of Performance. In all, justifications were requested for five objects, three in the original line-ups, and two insertion objects. As grade increases, the amount of logic displayed increases significantly ($\chi^2 = 43.94$, $p < .001$).

Table 17

Logic Categories of Performance for Task 4: Seriation of Weight - >3 Objects

Description	Score # logical expl.	Distribution K	1	2	Total
Good	4, 5	0	15	22	37
Some	1-3	3	4	6	13
None	0	28	12	3	43
		31	31	31	93

There is quite a strong all-or-none effect with few Ss displaying Some logic. (Only six Ss had four out of five logical explanations, and in all cases the one explanation that missed being logical was a near miss--i.e., one relationship was mentioned but the second one was not made explicit. Therefore, these Ss were grouped in the "Good" category along with those whose logic was perfect.)

BEHAVIOR.

According to the definition of correct seriation performance at the behavioral level employed in this study an S must demonstrate: (1) the ability to organize objects in such a way so as to adequately discover the relationships among them (i.e., in this task, to weigh the pairs Adequately) and, (2) the ability to line the objects up correctly according to these perceived relationships. These two abilities are involved in both initial seriation and insertion behavior. The ability to correctly insert objects into a series is included in Piaget's definition of the seriation operation. Therefore, the measure used as the basis for the behavioral Categories of Performance are those where seriations and insertions are correct given that the weighing procedure were Adequate.

Table 18 presents the definitions and distributions of the Categories of Performance for behavior for this task. Again there is a significant grade effect ($\chi^2 = 40.99$, $p < .001$).

Table 18

Behavior Categories of Performance for Task 4: Seriation of Weight - > 3 Objects

Description	Score # correct Sertns & Insertns given Adeq. Wghings	Distribution K	1	2	Total
Good	5	0	6	14	20
Fair	3, 4	0	12	3	15
Poor	0-2	31	13	14	58
		31	31	31	93

Notice that all the Kindergarten Ss have Poor behavior. In fact, 22 out of 31 of them had no behavioral responses correct, seriations or insertions, and only two Ss had a score of 2.

RELATIONSHIP BETWEEN LOGIC AND BEHAVIOR

Table 19 presents the distribution of the categories for logic and behavior combined. There are some things occurring in this task which occur in no other. First of all, five Ss display Good logic but Poor behavior. That is, one could say they have the concept but they cannot apply it. We will note later how these five Ss fared on the transitivity tasks. With seriation it looks as though they have a grasp of the operation of seriation once the series are established for them, but they are overwhelmed by the number of objects presented and simply cannot organize them all themselves. Also there are 12 Ss who display Some logic but Poor behavior, another unusual finding. Unlike the previous tasks there are no Ss here with No logic but Good behavior. Much logical understanding is obviously necessary in this task before it can be dealt with adequately on the behavioral level. The only Ss with Good behavior are those who also have Good logic. But all of those with Good logic do not have Good behavior. Logical ability is a necessary but not a sufficient condition for good behavioral performance. What else is needed? One could speculate at length about this, but by observing the

Table 19

Relationship between Logic and Behavior Categories of Performance for Task 4: Seriation of Weight - >3 Objects.

		Logic			
		Good	Some	None	Total
Behavior	Good	20	0	0	20
	Fair	12	1	2	15
	Poor	5	12	41	58
	Total	37	13	43	93

Ss one could see that one quality the Ss with Good behavior demonstrated which was lacking in varying degrees in the others was an organizational ability. The Ss who went through all the trials perfectly were very organized and directed in their behavior. They weighed the pairs together in a systematic fashion keeping track of which ones had been weighed together and what the outcomes were in some very precise way--usually by placing the objects in appropriate order according to the weighings completed so far. In other words they were spatially organized: they arranged the objects in appropriate spatial relations as they went along. Other Ss may have done this too up to a point, but not as consistently or clearly enough so as to eliminate all ambiguities. This task would seem to divide those who are truly operational with regard to seriation from those who are not. Trial and error behavior will not lead to ultimately correct solutions as it will in all the other tasks discussed so far. There is no visual feedback to help one as there is in the length dimension tasks, and there are too many objects for a solution to be correct by chance, unlike in the transitivity of weight task where this is a possibility. Therefore, according to this reasoning those 20 Ss with Good behavior and Good logic in this task should be the cream of the sample and should have perfect performance in the other tasks if they are really related. This hypothesis will be examined later.

The relationship between logic and behavior is highly significant ($\chi^2 = 63.98$, $p < .001$). Logical performance outstrips behavioral performance, the number of Ss in Table 19 falling below the Good-Good to Poor-None diagonal (29) far exceeding the number above (2).

TASK 5. CLASS INCLUSION - ALL-SOME.

1. ANALYSIS OF THE QUESTIONS.

The first question is:

(1) CB - Are all the circles blue? - to which the answer is Yes, because there are five circles and they are all blue; there are no red circles, etc.

The second question is:

(2) RS - Are all the red ones squares - to which the answer is also Yes, since there are only two red things and they are both square, there being no red circles, etc.

The form of both of these questions is the same and it is designated as form AB: i.e., what is being asked is, "Are all the As Bs?", where A<B. The structure of the relationship is: B = A + A', which can be considered in two ways:

1. Blue objects = blue circles + blue squares,
   (relevant to the first question, CB), or
2. Square objects = red squares + blue squares,
   (relevant to the second question, RS).

A is included in B, hence we are examining class inclusion. What is being asked is "Are all the As some of the Bs?" In form AB the subject of the question (A) is smaller than and included in the class of the predicate (B).

The third question is:

(3) BC - Are all the blue ones circles? - to which the answer is No, because there are blue ones which are square.

The fourth question is:

(4) SR - Are all the squares red? - to which the answer is also No, because there are two squares which are blue.

The form of these two questions is BA, "Are all the Bs As?", where A<B as outlined above. In this case the subject (B) is larger than the predicate which it includes (A).

In the first and fourth questions, CB and SR, the subject of the question is a shape and the predicate is a colour In the second and third questions, RS and BC, the subjects are colours and the predicates are shapes. This distinction is an important one in the results which follow.

2. DESCRIPTIVE RESULTS

Since this task is the same as one given by Inhelder and Piaget (1964) to 52 children, it might be worthwhile to describe briefly their results as a basis of comparison with our own. Their scoring method is not clearly described, but there seems to have been some leniency for self-corrected responses. For each of the four questions, between the ages of five and eight years there is a consistent increase in the percentage of correct responses with increasing age. Question RS is the hardest question, with 54 percent of the five-year olds responding correctly, this percentage increasing every year up to 80 percent at eight years. Question CB is the easiest with the percentage of correct responses ranging from 67 percent at five years to 100 percent at eight years. Questions BC and SR are of intermediate difficulty. Besides this consistent increase in correct responses with age, two other trends are evident in Inhelder and Piaget's results. The first is that, on the basis of the percentage of Ss responding to both questions correctly, questions of the form BA (BC and SR) are easier than questions of the form AB (CB and RS) at all ages, except age seven years where they are equal. The second trend is that the mean percentage of correct responses for questions CB and SR combined, questions in which the subject is shape, the predicate colour, is higher at all ages (except at five years where they are equal) than the mean of questions BC and RS where the subject is colour, the predicate shape.

Table 20 presents the percentage of correct responses given to the four questions by the 93 children in this study. The first four columns present these percentages for each question separately. The column headed "AB" presents the percentage of Ss who answer both questions CB and RS (of the form AB) correctly. Similarly, the column headed "BA" presents the percentage of Ss who answer both questions BC and SR (of the form BA) correctly. The responses recorded were the S's initial responses to the questions, but if a S changed his

Table 20

Descriptive results for Task 5: Class Inclusion: All-Some

Grade	CB	RS	BC	SR	AB	BA	all 4 correct	CB & SR both corr.	RS & BC both corr.
	Percentage of Correct Responses								
Kindergarten	94	55	32	97	52	32	13	90	16
Grade 1	90	55	81	97	48	77	39	87	48
Grade 2	97	52	55	90	52	45	29	87	35
Mean %	94	58	56	95	51	52	27	88	33

mind immediately without any prompting this second response was the one scored.

With all four questions there is no consistent improvement with grade. There is only one improvement between grades, the increase in the percentage of correct BC responses between Kindergarten and Grade 1 ($p<.001$), but then performance worsens between Grades 1 and 2 ($p<.05$). Any percentages in the last five columns which include the BC response also reflect this result. In column BA the Kindergarten-Grade 1 difference is significant ($p<.001$) as is the Grade 1-Grade 2 difference ($p<.01$). In the last column, RS and BC both correct, only the Kindergarten-Grade 1 difference is significant ($p<.01$). Similarly, only the difference between Kindergarten and Grade 1 is significant ($p<.05$) in the percentage of Ss getting all four questions correct.

Questions CB and SR are answered correctly by almost all Ss in all grades. Questions RS and BC are more difficult. The percentage of Ss responding correctly to both questions CB and SR (subject shape, predicate color) is significantly greater than the percentage responding correctly to both questions RS and BC (subject colour, predicate shape) in all grades (all p's<.01 or better). Inhelder and Piaget found the same thing, but the trend in the present results is much more marked. This will be discussed shortly.

There is no overall difference between the percentage of correct responses to questions of Forms AB and BA. There are differences within the grades (the only significant one is in Grade 1 where $p<.05$), but these are not all in the same direction and seem highly unpredictable.

Inhelder and Piaget found questions of the form BA easier than of the form AB at all ages five through nine years except seven years where they were equal. These authors expound at great length as to why, if $B = A+A'$, Ss find it easier to answer correctly that all Bs are As is wrong than to recognize that all As are Bs is right. They reason that in the question "Are all Bs As" which is really asking "Are all the Bs some of the As?", most Ss will interpret it to mean "Are all the Bs all the As?" They make the same mistake with form AB also. They quantify the predicate incorrectly. Because S sees the collection $B = A+A'$ in front of him he can see that $A \neq B$ since there are some A' present and so he can easily say No to questions of the form BA - "Are all the Bs As?" But since he has not mastered class inclusion, he cannot see that $A = B-A'$ and so when asked "Are all the As Bs?" he still says No, because there are some A's. He interprets the question to be asking "Is $A = B$?" instead of "Is $A<B$?" With neither form is S quantifying the predicate correctly, but he gets away with it with form BA and not with form AB. This interpretation does shed light on what might be occurring in the non-operational child's mind, but it does not explain the data in this task since no such differences between the two forms were found.

Our most dramatic result, and one that does agree with Inhelder and Piaget, is that a question tends to be easier when the collection as a whole indicated by the word "all" is defined by shape than when defined by colour. Obviously it must be easier to form a non-graphic collection (i.e., one formed in S's mind and not actually assembled physically) on the basis of shape than of colour. That is, S can more easily image collections based on shape than based on colour in this task. Thus he can keep the graphic collection of the "circles" or the "squares" in his head very well and make colour the subordinate quality of the objects in front of him. The reverse is much harder to do.

This finding of the increased saliency of form over colour for children in this age range is a common one in the literature. In a study by Kagan and Lemkin (1961), Ss aged 3 years, 9 months to 8 years, 6 months were asked to select among comparison stimuli varying in form, colour and size, the stimulus which was the "same as" a standard stimulus. For all ages and with all the stimuli used form was disctinctively preferred to colour as a basis for similarity. Suchman and Trobasso (1966) tested 3 to 6 year olds in a similar but more complicated situation where hue saturation and figure contour were varied. They found that the younger Ss chose colour over form and that the older Ss chose form over colour as the basis of similarity among the stimuli. The median age at which their Ss began to choose form over colour was 4 years and 2 months. Within the age range tested the trend for form increased and the difference between form and colour increased significantly beyond this transition age.

CATEGORIES OF PERFORMANCE

Since there were no significant between-grade differences in this task it would appear not to discriminate among the Ss very well. However, there may be some differences between Ss who respond perfectly on this task and those who do not in their performances on the other tasks. Therefore, the Ss have been divided into Categories of Performance whose definitions and distributions are presented in Table 21. The relationship between Grade and Category is not significant ($\chi^2 = 6.94$, $p > .10$), as would be expected.

TASK 6. CLASS INCLUSION: PART-WHOLE.

DESCRIPTIVE RESULTS.

Table 22 presents the percentages of correct responses given to each of the six questions by each grade. Question 1, which was a sort of preliminary class-inclusion exercise to see if S understood the difference between tulips and non-tulips, is responded to correctly by 70 percent of all Ss and no regular trend with grade can be noted.

Questions 2 and 3 are the same form: S is asked to compare one part with the whole. In question 2 he must compare the pink tulips with all the tulips; in question 3 he

Table 21

Categories of Performance for Task 5: Class Inclusion: All-Some

Description	Score # correct responses	Distribution K	1	2	Total
Good	4	4	11	9	24
Fair	3	17	15	11	43
Poor	0-2	10	5	11	26
		31	31	31	93

Table 22

Descriptive results for Task 6: Class Inclusion: Part-Whole

Question / Grade	Percentage of Correct Responses						
	1	2	3	4	5	6	Mean
Kindergarten	58.1	38.7	3.2	93.5	38.7	93.5	53.3
Grade 1	83.9	67.7	16.1	100	67.7	100	72.6
Grade 2	67.7	87.1	45.2	100	74.2	96.8	78.5
Mean	69.9	64.5	21.5	97.8	60.2	96.8	68.4

must compare the tulips with all the flowers. Despite this similarity in form, question 2 is significantly easier than question 3 for all grades (all p's$<.01$ or better). Each question also shows considerable improvement with increasing grade (for all between grade differences for both questions $p<.05$ or better, two-tailed except the Grade 1-Grade 2 difference for question 2, and the Kindergarten-Grade 1 difference for question 3, where both p's $<.01$). In order to determine why question 2 is so much easier than question 3 let us look at the types of errors made in response to each question.

For all grades combined 64.5 percent of the responses are correct for question 2. The most usual kind of error made was to say that there were the same number each because there were four pink tulips and four yellow tulips. This response was given 19.4 percent of the time. The remaining 16.1 percent of the responses were either perceptually based variations of this type (i.e., one bunch looked bigger or took up more space etc.) or responses which can only be termed idiosyncratic or complete nonsense (e.g., "More pink tulips --they're nicer. I have some just like them at home" etc.). Ninety percent of these are made by Kindergarten Ss. Therefore, although the question calls for a comparison of a part with a whole, 19.4 percent of the Ss compared one part with the other part. This accounts for around 56 percent of the errors made in question 2. For question 3, comparing the tulips with all the flowers, only 21.5 percent of the responses were correct. Over two thirds of the Ss (66.7 percent) compared one part, the tulips, with the other part, the non-tulips, instead of with the whole. This accounts for 85 percent of all the errors made, the others being nearly all nonsense responses made mostly Kindergarteners. Why this greater tendency to compare the two parts instead of one part with the whole in question 3 than in question 2 (66.7 percent of the responses compared with 19.4 percent of the responses)? The confusion on the S's part may be verbal rather than conceptual. Question 2 refers to two sets of "tulips", one set being modified by "pink". The two sets are not really verbally distinct, and the fact that one set is modified may help to enhance the notion that it is included in the non-modified and therefore more general set. Question 3 refers to two verbally distinct sets of objects--"tulips" and "flowers"--and, as we have seen, most Ss (67 percent) interpret the "flowers" to be the non-tulips. They are not given the same verbal clues in this question that the one set is included in the other. The tulips are really called by two distinct names here--"tulips" and "flowers"--but most Ss failed to see this. So they naturally thought that E meant them to compare the tulips with the non-tulips, i.e., the "flowers". This verbal factor may account for much of the difference in difficulty between these questions.

Questions 4 and 6 are answered correctly by almost everyone and therefore show no trend by grade. These two questions are of the same form--if you remove one part will there be any of the whole remaining? All one has to do to respond correctly is to compare one part with the other; the concept of the whole need not enter the picture.

Question 5 asks if you remove the whole will one of the parts remain. Thus S has to compare the whole with one of its parts to respond correctly here. The most common error here again was to interpret "flowers" to be the non-tulips and therefore to say that if you took the flowers away the tulips would still remain. The percentages of correct responses to this question are much lower than for questions 4 and 6, and all the within-grade differences between question 5 and each of questions 4 and 6 are significant (for Kindergarten both p's $<.001$; for Grade 1 both p's $<.01$; for Grade 2 both p's $<.02$). The percentage of correct responses to question 5 also increase with increasing grade. The Kindergarten-Grade 1 and Kindergarten-Grade 2 differences are both significant ($p<.05$ and $.01$ respectively). The Grade 1-Grade 2 difference is not significant.

According to Inhelder and Piaget's reasoning, only questions 2 and 3 of this task are decisive tests of whether or not S understands class inclusion. Here S is asked to compare the extension of B with that of A (where B = A+A'), or of C with that of B (where C = B+B'). If S recognizes that there are more tulips (B) than pink tulips (A) in a bunch he must be aware of B as the sum of A+A' and he must simultaneously be aware of A as the difference B-A'. It is this ability to transform the equation B = A+A' into A = B-A' which defines operational ability in this realm. Before this develops S can be intuitively aware that the whole is the union of its parts and that one part is distinct from another (and therefore can correctly answer questions 4 and 6), even though he cannot compare the extension of the part and the whole. When asked to compare the extension of B and A he reduces B to A'.

But Inhelder and Piaget's explanations do not explain why there are large differences between the percentages of correct responses to questions 2 and 3. Indeed, these authors found no such differences. In fact, question 3 is easier for their Ss than question 2 for all ages 5 through 8 years, whereas the opposite is true for our Ss. Why? Perhaps their Ss were more familiar with flowers in general than ours and had no trouble remembering that "primulas" (instead of our "tulips") were also flowers.

Inhelder and Piaget also group question 5 with questions 4 and 6 and declare them all to be the same form--i.e., one which asks S merely to compare the two parts of a whole instead of one part with the whole. Their results suggest that these questions are all of the same order also, since they found no difference in the percentage of correct responses among them. But ours do not, and logical analysis of these questions also discloses differences in the forms of these questions as discussed above. Question 5 does require S to compare the extension of the whole with one of its parts and is therefore harder to answer than questions 4 and 6.

CATEGORIES OF PERFORMANCE

Since questions 2, 3 and 5 are the only ones which discriminate among the grades and which are true tests of class inclusion reasoning ability, the Categories of Performance for this task have been based solely upon the responses to these three questions. Responses to questions 1, 4 and 6 have been ignored for this purpose. Table 23 presents the definitions and distributions of the categories for this task. Performance improves significantly with increasing grade ($\chi^2 = 22.91$, $p < .001$).

TASK 7. ONE OBJECT, TWO QUALITIES.

This task was devised as a test of simultaneous combinativity . In order to respond correctly to questions about the stimuli S must count some of the objects twice, regarding them in two different ways.

ANALYSIS OF STIMULI AND QUESTIONS.

The underlying structure of the objects on the six cards presented to Ss will be examined first. Figure 2 presents graphical representations of the structures of each of the

Table 23

Categories of Performance for Task 6: Class Inclusion: Part-Whole

Description	Score # resp.correct to Q's 2,3,5.	Distribution K	1	2	Total
Good	3	1	5	11	17
Fair	2	5	11	13	29
Poor	0, 1	25	15	7	47
		31	31	31	93

Figure 2

Graphical representations of underlying structure of stimuli used in Task 7

Card 1

A. Color; B. Stem Length

A. Color \ B. Stem Length	b-short	b'-long
a-red	ab	ab'
a'-yellow	a'b	a'b'

A,B-dimensions
a,a',b,b'-qualities
ab, a'b, ab', a'b'-quadrants

Card 2

Same as Card 1 where

A. Flower, a - daisy
a' - tulip

B. Color, b - white
b'- yellow

Card 3

Same as Card 1 where,

A. Shape, a - circle
a'- square

B. Size, b - big
b'- little

Card 4

A. Color; B. Thing

A. Color \ B. Thing	b-apple	b'-triangle
a-red	ab	ab'
a'-green	a'b	--

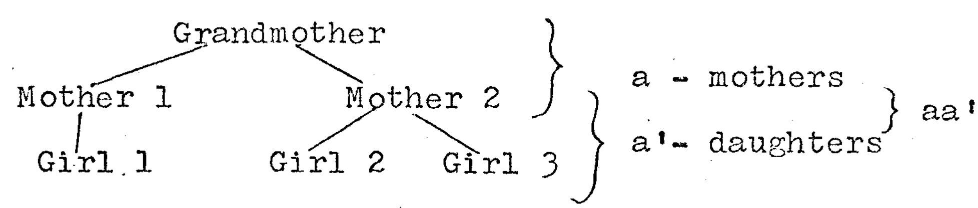
Card 5

A. Color

red } a
red and green } a, a'
green } a'
a, a' } aa'

Card 6

A. Role



six cards. Cards 1, 2 and 3 all have the same structure. Question (a) for each card asks S to compare all a with all b; question (b) asks him to compare all a' with all b'. In other words in both cases he must compare all the objects possessing one quality of one dimension with all the objects possessing one quality of the other dimension. Some objects possess both qualities and must be counted twice.

Card 4 has a different structure from the first three since one of the quadrants is empty--there are no green triangles. S is still asked to compare all a with all b, as with Cards 1-3, but in this case all the objects must be considered in the response whereas with Cards 1-3 some objects, e.g., those in quadrant a'b', must be ignored.

On Card 5 all the objects differ on only one dimension --colour. They are all squares. S must consider all of the objects on the card as he is asked to compare all a with all a'. Card 6 is similar to 5 in that the objects differ along only one dimension, in this case role. But it differs from all the other cards in this task since it deals with the relationships among the stimuli and not with their qualities.

DESCRIPTIVE RESULTS

Table 24 presents the number of correct responses given to each question in each of the three grades. In general, performance is very poor on this task, with only 20.2 percent of the total responses being correct, 5.0 percent in Kindergarten,

Table 24

Descriptive results for Task 7: One Object, Two Qualities

Question \ Grade (n)	Kindergarten (31)	Grade 1 (31)	Grade 2 (31)	Total (93)
	Number of Correct Responses			
1a	3	9	18	30
1b	3	15	21	39
2a	0	2	7	9
2b	0	2	9	11
3a	0	3	9	12
3b	0	2	9	11
4	0	4	5	9
5	8	15	19	42
6	0	3	3	6
Total (279 trials)	14	55	100	169

19.7 percent in Grade 1, and 35.8 percent in Grade 2. However, performance does improve with grade in each question. (The between grade differences which are significant, p .05 or better, two-tailed are: question 1a, both lower grades with Grade 2; question 1b, Kindergarten with both higher grades; question 2a, Grade 1-Grade 2; question 2b, both lower grades with Grade 2; question 3a, Kindergarten-Grade 2; question 3b, both lower grades with Grade 2; question 4, Kindergarten with both higher grades; question 5, Kindergarten-Grade 2; question 6, none.)

Performance is best on questions 1a and b and 5. The differences between the number of correct responses to questions 1a and b and all the other questions (except #5) are significant ($p < .05$ or better, two-tailed) except for the Kindergarteners, and for the Grade 1 Ss between questions 1a and 3a, 3b, 4 and 6. Performance in all grades on question 5 is significantly better than on all the other questions except 1a and b (all p's $< .05$ or better).

Why are questions 1a and b and 5 the easiest? There are two visual factors which could contribute at least partially to this. First of all the pictures on both these cards are arranged in a horizontal row whereas those on the other cards are scattered randomly (except card 6 where the pictures are arranged hierarchically). It is possible that a horizontal row is much easier to scan and also makes it easier for S to count the objects possessing the various qualities specified and keep track of those he has already counted. Secondly, on Card 1 the two dimensions, colour and length of stem, are visually separated making it possible for S to respond by first looking across the heads of the flowers for colour and then along their stems for length. Similarly on Card 5 the two relevant qualities, red and green, were located on separate areas of the stimuli and therefore could be regarded independently. On the other cards the two dimensions are not visually separated. Thus S could not break the objects into two separate dimensions or qualities as easily as with Cards 1 and 5.

Another factor which could account for questions 1a and b and 5 being easier than all the others is the way in which the objects are labelled in the questions. On Card 1 everything is labelled as a "flower" and the distinguishing labels are the modifiers of the word, e.g., "red", "with long stems", etc. Similarly, on Card 5 all the objects are called "squares" and they are distinguished simply by the colours occurring on them. Thus on both cards the groups of objects to be compared are both referred to by their general names modified by the dimensions or qualities to be distinguished. On the other cards (except Card 6 again), in each question one group is called by its specific name (e.g., on Card 2, "daisies" or "tulips") while the other group is referred to by its general name modified by a quality (e.g., "white flowers" or "yellow flowers"). It is highly possible, therefore, that in question 2a, say, S could interpret the "white flowers" to mean the white tulips since the tulips are the obvious opposite to the daisies and he might assume that E was merely labelling the tulips as "flowers". Similarly for questions 3a and b, it would be quite natural on S's part to assume that E was referring to the objects with the shape not specifically labelled as "things" instead of as their specific shape. This would encourage the occurrence of the type of error found to a considerable degree with Cards 2 and 3, but virtually nonexistant with Card 1, namely, the comparison of one quality with one opposite quadrant. For instance, in response to question 2a "Which are there more of daisies (5) or white flowers (4)?" many Ss replied "Daisies, because there are five of them and only one white flower" (meaning the white tulip), thus betraying the fact that they are equating "flowers" and "tulips" despite the fact that in describing the pictures on the card E carefully called them all flowers "some of which are daisies and some tulips". Thus, both visual and verbal factors could account for the different rates of correct responses among the questions. It is impossible to differentiate among these factors in the present study.

CATEGORIES OF PERFORMANCE

In order to compare performance on this task with that on the other tasks the Ss have been divided into three Categories of Performance based on the number of correct responses they gave to the nine questions they were posed about the six cards (two questions each for Cards 1-3, one each for Cards 4-6). Table 25 presents the definitions and distributions of these categories. The relationship between grade and performance is significant ($\chi^2 = 24.42$, $p<.001$).

TASK 8. INFERENCE I.

DESCRIPTIVE RESULTS

1. Preliminary Training.

Ss in all grades quickly learned the subgoal segment, i.e., they learned which door yielded the marble and which yielded the ball bearing. No one made more than seven errors, and the vast majority of Ss made no errors at all. The major goal segment was also easily learned, i.e., whether the marble or the ball bearing yielded the candy. Again the maximum number of errors was seven. Table 26 presents the number of Ss reaching criterion without error for both segments. There are practically no differences among the grades.

2. Test Trial: Initial Choice.

The first component of inferential behavior is the initial choice between doors A and X of which A is the inferential choice. Table 27 presents the number of Ss making the responses described. The top half of the table deals with those Ss who spontaneously opened either door A or X. The bottom half deals with those who did nothing for 60 seconds on the test trial, and then made a choice only after E said: "Which door should

Table 25

Categories of Performance for Task 7: One Object-Two Qualities

Description	Score # responses correct	Distribution K	1	2	Total
Good	6-9	0	2	8	10
Fair	1-5	12	22	18	52
Poor	0	19	7	5	31
		31	31	31	93

Table 26

Number of Ss reaching criterion without error on training trials in Task 8: Inference I.

Grade (n)	Kindergarten (31)	Grade 1 (31)	Grade 2 (31)	Total (93)
Side panels - subgoals	26	29	28	83
Middle panel - major goal	28	27	26	81
Total	54	56	54	164

Table 27

Descriptive results for Task 8: Inference I

Type of Response	Kindergarten	Grade 1	Grade 2	Total
	Number of Ss making each type of response on the Test Trial			
Spontaneous Behavior				
1. Initial Choice:				
choose A first	19	19	21	59
choose X first	9	7	9	25
Total	28	26	30	84
2. Integration Response:				
direct	9	16	7	32
indirect	7	3	13	23
Total	16	19	6	55
After Adult Prompt				
1. Initial Choice:				
choose A first	3	3	1	7
choose X first	0	2	0	2
Total	3	5	1	9
2. Integration Response:				
direct	2	2	0	4
indirect	0	1	0	1
Total	2	3	0	5

should you open to help you get a candy?" This is called the "adult prompt". The number of Ss spontaneously choosing doors A or X first hardly vary at all among the three grades. Ninety percent (84/93) of the Ss do spontaneously open one of the doors within 60 seconds and 70 percent (59/84) of them make the correct inferential choice. Ten percent (9/93) of the Ss require an adult prompt to get started and 78 percent of them subsequently make the correct choice (7/9).

3. Test Trial: Integration Response.

If after his initial choice S inserted a subgoal into hole B he was considered to have made an integration response. This response was correct if the subgoal he used was B and incorrect if he used Y. Of those making an integration response only one S (in Grade 2) inserted subgoal Y into the hole. All other integration responses made were correct and are the only ones discussed from now on. However, a correct integration response could be direct, that is, be made with no unnecessary responses intervening between initial and goal responses, or indirect, that is, occur after one or more unnecessary responses (i.e., pulling door X either before or after S had pulled door A). From Table 27 it can be seen that of all those Ss spontaneously making an initial choice, the total number of Ss who then make an integration response increases slightly with increasing grade (16, 19 and 20 respectively for K, Grades 1 and 2). However, the number of Ss making a direct integration response does not vary in any regular manner with grade. The number of Ss in Grade 1 making a direct response (16/31) is higher than in the other two grades (only the Grade 1-Grade 2 difference is significant at the .05 level, two-tailed). The number in Grade 2 is surprisingly low (7/31). But the number making an indirect response in Grade 2 (13/21) is greater than in Kindergarten or Grade 1 (7 and 3 respectively) (only the Grade 1-Grade 2 difference is significant at the .05 level, two-tailed). Of all those Ss spontaneously making an initial choice, 65 percent (55/84) subsequently make an integration response. The other 35 percent make no further response after opening one of the side panel doors. Of those Ss making the correct initial response of opening door A, 54 percent make a direct integration response. Of those responding initially only after an adult prompt 56 percent (5/9) make an integration response. Of those making a correct initial choice after an adult prompt 57 percent (4/7) make a direct response.

CATEGORIES OF PERFORMANCE

For the purpose of comparing an S's performance on this task with that on the other tasks in this study, the Ss have again been divided into Categories of Performance. Table 28 presents the definitions and the distributions by grade of these categories. Obviously those who make a spontaneous initial choice followed by a direct integration response (Category A) fulfill all that is expected of them in

Table 28

Categories of Performance for Task 8: Inference I.

Description	Distribution K	1	2	Total
A. Spontaneous direct integration response	9	16	7	32
B. Spontaneous indirect integration response	7	3	13	23
C. After adult prompt, direct or indirect integration response	2	3	0	5
D. No integration response	13	9	11	33
	31	31	31	93

this task. They demonstrate the ability, at least in this situation, to link two successively occurring events together (A→B, B→G) in order to reach the main goal. In so doing they can ignore any other stimuli (such as door X) in the situation which will not help them to achieve the goal. Those who spontaneously make an initial choice which is followed by an indirect integration response (Category B) are displaying ability to link two pieces of behavior together to achieve a goal, but they cannot ignore other stimuli and possible responses available in the situation. They can make the integration response only after they have the two subgoals available in front of them as they had during training. They cannot shortcut this process. Those who make initial choices only after an adult prompt (Category C) present a problem for us. They need no prompt to make the integration response and 80 percent of them made a direct integration response (4/5). But initially they were stymied when asked to get a candy for themselves. They did not know where to start. But once E put them in motion the behavior proceeded along alright. Since it is the integration response and whether or not it is direct which is the most crucial aspect of performance in this task, one could argue that these five Ss should be put into Categories A and B. However, it is important that S see exactly how to go about getting to the goal for himself without anyone giving him hints, and this fact prevents us from putting these five Ss into the first two categories. In fact, when these categories are compared with those of other tasks subsequently, Category C will be combined with Category D to insure expected values of sufficient size in the Chi-square tests. For the distribution in Table 28 the Grade by Category relationship is significant ($\chi^2 = 14.32$, $p<.05$) but certainly not regular. Grade 1 performance is superior to that of the other two grades.

TASK 9. INFERENCE II.

Every S went through the two training stages without error. For the test for inferential behavior, every S (with the curious exception of two Ss, one in Kindergarten and one in Grade 2, both of whom achieved the final solution) tried to put the green truck through the garage door. This response was to be expected since it was an obvious first possible way of getting a candy. The fact that the green truck was too large in all its dimensions to fit through nevertheless did not deter many Ss from trying out ingeneous angles and manoeuvers to make the truck fit through the doorway. All these attempts failed. Many Ss then gave up declaring that there was no way they could get the truck in and therefore no chance of getting a candy. However, some then saw that if they could get the blue racer back again they could achieve their goal. They then traded their green truck for the blue racer with E, drove the racer into the garage, and received a candy. These Ss achieved solution spontaneously. If S had tried in vain to get the green truck through the doorway and then given up in despair, E always said: "Is there any way I can help you get a candy?" If S then saw that he could trade the truck for E's racer and then proceeded to achieve the goal, he was considered to have achieved the solution with an adult prompt. Many Ss were not helped by this prompt and never achieved the solution. Most of these suggested somehow that E help them get the green truck into the garage, or simply give them a candy since they were E's to give.

Table 29 presents the Categories of Performance, their definitions and distributions, which are just those responses described above. The number of Ss in the various categories in Grades 1 and 2 are almost identical, and only slightly fewer Ss in Kindergarten reached a solution than in the other two grades. The relationship between Grade and Category is not significant ($\chi^2 = 2.19$). Despite this lack of discrimination between grades this task might discriminate among the Ss within grades in some fashion which is relevant to their performances on other tasks in this study.

Note that performance on this task differs from that on the previous one in the following way. Of those failing to make an integration response in the previous task, all did make an initial choice by opening one or other side panel doors. It was the second response in the chain which they failed to perform. In the present task those Ss who failed to solve the problem failed to make the first necessary response toward

Table 29

Categories of Performance for Task 9: Inference II

Description	Distribution K	1	2	Total
A. Spontaneous Solution	5	8	9	22
B. Solution after adult prompt	8	9	9	26
C. No Solution	18	14	13	45
	31	31	31	93

reaching the goal, i.e., trading the green truck for the blue racer. All of those who did make the trade made the second necessary response, inserting the racer in the garage, to get the candy. Thus there are obviously crucial differences between these two tasks in the saliency of their first or second necessary responses.

RELATIONSHIPS AMONG THE TASKS

INTRATASK RELATIONSHIPS BETWEEN LOGIC AND BEHAVIOR IN TRANSITIVITY AND SERIATION TASKS

The results to be described in this section pertain to the fourth purpose outlined in the Introduction. Table 30 presents a summary of the contingency tables which relate logical and behavioral performances within each of the four transitivity and seriation tasks. The first two columns present the Chi-square scores and p's of each of these tables. The third column gives the number of Ss who fall in the same category on both measures. That is, they either have Good behavior and logic, or Fair behavior and Some logic, or Poor behavior and No logic. (For the transitivity of length task where there were only two categories for behavioral performances, 60 is the number of Ss with either Good behavior and logic, or Poor behavior and No logic). The fourth column, the "Direction of the relationship", gives the number of Ss whose categories of behavior and logic differ and the direction of the differences. For instance, for task 2, seriation

Table 30

Relationship between Logic and Behavior on Transitivity and Seriation tasks: summary of contingency tables

Log. vs Beh. on Task #	χ^2	P	# Ss on Diagonal	Direction of relnship	φ'	λ_{LB}
1. Trans Lgth	22.03	<.001	60	Beh > Log 30:3	.49	.20
2. Sertn Lgth	30.43	<.001	55	Beh > Log 33:5	.40	.22
3. Trans Wgt	22.28	<.001	34	Beh > Log 58:1	.35	.10
4. Sertn Wgt	63.98	<.001	62	Log > Beh 29:2	.59	.53

of length, behavioral performance is better than logical performance (stated as Beh<Log in the table) because 33 Ss had either Good behavioral performance but showed Some or No logic, or had Fair behavior and No logic. That is, they had a higher category of performance for behavior than for logic. Only 5 Ss did the reverse. The fifth column presents the phi-coefficient (φ'), which is a measure of the strength of association between two variables in a contingency table. It is based on the χ^2 for the table and can range from zero, reflecting complete independence, to one, showing complete dependence, of the variables. Finally, the last column gives λ_{LB}, which is an index of predictive association between two attributes, L and B (for Logic and Behavior). That is, knowing the S's category of performance on either measure L or B reduces the probability of making an error in predicting his category on the other measure by λ_{LB} X 100 percent. It is a measure of how much one attribute tells us about the other, a measure of our power to predict where neither attribute is specifically designated as the thing predicted from or known first. It is calculated independently of the χ^2 for each contingency table (Hays, 1963, p606ff).

The pertinent question here is what is the relationship within each task between the behavior and logic measures? Are they interchangeable measures in identifying Ss who do or do not possess the operation? The results in Table 30 indicate that these measures are not independent. All the χ^2s are significant beyond the .001 level, and the φ's range from .35 to .59. But they are not interchangeable either since one measure develops earlier than the other in at least one-third of the Ss in each task (in task 3 this proportion is nearly two-thirds). In tasks 1 to 3 behavior develops before logic; in task 4 logic develops before behavior. The directions of all these relationships are quite stable, reflected in the large ratios among those Ss who change categories. In every task, of those Ss who change, a very large majority change in one direction and only a few change in the other.

All the Ss in the two transitivity tasks, and nearly all in the seriation tasks, who have Good logic also have Good behavior. But Ss with No logic are almost evenly divided among the behavior categories: in fact there is slightly more chance of their having Good behavior than any other category. Logical reasoning ability is not a prerequisite for Good behavioral performance on these tasks. The percentages of Ss with Good behavior who also have Good logic in tasks 1, 2 and 3 are 61, 65 and 35 respectively. In task 4, seriation of weight, this percentage is 100. Also in this task 41 of the 43 Ss with No logic have Poor behavior, and 12 of the 13 Ss with Some logic have Poor behavior. These results no doubt make for the large λ_{LB} (.53) for this task. The error of predicting an S's category of performance on one measure is greatly reduced by knowing his category on the other measure. In the seriation of weight task logical reasoning is obviously a prerequisite to good behavioral performance.

In a study of the development of transitivity of length Smedslund (1963) found that any S who gave four out of four correct behavioral transitivity responses also gave a logical explanation for at least one of these responses. This is not the case in the present data. The percentage of Ss with perfect behavior giving at least one logical explanation is 76 for transitivity of length and 49 for transitivity of weight. In Smedslund's study 89 percent of all correct behavior transitivity responses are given logical explanations. In the present study the percentages are 69 and 28 for transitivity of length and weight respectively. In the seriation of length task 57 percent of all correct insertions are logically justified (see Table 31).

How can S give a correct behavioral response without a logical justification? As Smedslund points out, a correct prediction that A<C need not be based upon transitive reasoning. A non-transitive S may give a correct response for several reasons, one of which is guessing. The probability of guessing correctly is lessened if there is no reinforcement and no information about errors given in the situation, as here, but it is still greater than zero. Although guessing was controlled for as much as possible in the experimental situation, it was not controlled for statistically in the

Table 31

Type of explanation given for correct behavioral responses in Tasks 1, 2 and 3

Type of explan.	Percentage of Responses Given		
	Trans length	Sertn length	Trans wght
Logical	68.6	56.9	28.3
Semi-logical	2.4	18.1	33.8
Other	29.0	25.0	37.9

present research.[1] Another possible basis of getting a correct behavioral response is perceptual discrimination of the relationship between objects A and C. This was very unlikely in the present transitivity of length task where the differences between stimuli A and C were very small or nil and where A and C were placed several feet apart and on different horizontal levels. Perceptual discrimination was impossible in the transitivity of weight task where the weights of the stimuli were not related to their sizes or to any other of their features. It is possible in the seriation of length task, though, where S can line up the objects and insert others by trial and error solely on the basis of perception. A third non-logical way of being behaviorally correct is by the use of what Smedslund calls a non-transitive

---

[1] Guessing is a possible factor in the two behavioral transitivity measures only: it is certainly not a possible basis for any verbally logical responses, nor for any behavioral seriation responses where the coordination of so many relationships precludes a correct outcome by chance. Therefore, the chance of miscategorizing an S because he guessed the correct response instead of figuring it out some other way only occurs with the behavioral transitivity measures. This source of unreliability could account for some of the instability of these behavior measures in this study, but it can be noted in Table 33 that behavior across seriation tasks where guessing is nil has no more predictive association than across the transitivity tasks. Also, the relationships in Table 30 demonstrate that the degree of association between behavior and logic is no less on the transitivity tasks than on the seriation tasks. Therefore, any miscategorizing error, if it exists, is probably small and affects such a small area of the results, namely two response measures out of a total of 13, that it can be discounted as affecting the results to any significant degree.

hypothesis, or what is referred to in Table 31 as a semi-logical explanation. That is, S reasons only on the basis of one relationship--e.g., he says A>B, therefore A>C - instead of the two relationships necessary for a truly logical explanation. Table 31 shows that this type of explanation was given very little in the transitivity of length task (2.4 percent) but one-third of the time in the transitivity of weight task. It is thus more likely in the latter than in the former task that correct behavioral transitivity responses were based on non-transitive hypotheses. However, there were more correct behavioral responses in each task arrived at by neither logical nor semi-logical hypotheses, indeed by no regular sort of hypotheses but rather by perceptual, tautological, nonsense, or no reasons at all (see "other" row in Table 31). This might suggest that there is either a lot of guessing going on in these tasks or that Ss can use other types of hypotheses, or no hypotheses which can be verbalized at all, to derive the correct response.

It is important to remember that S is asked to justify his behavioral response after he makes it; therefore, what he says after making the response may not be the reasons he used before to help him derive his response. In fact, S may not have used any sort of reasoning which he can readily report beforehand--merely intuitive reactions may lead to the response. But then E asks for a verbal justification and S has to come up with something since he sees that E expects it. This something may be pure nonsense, or S may quickly seek a perceptual basis for his response, or he may remember that A went down before on the balance when put on with B so it will go down again with C, or he may even come up with a logical explanation. But none of these justifications necessarily reflect the reasoning process that S used to arrive at his behavioral response initially.

In tasks 1, 2 and 3, many Ss had perfect behavior but no logic (n = 18, 14 and 25 respectively; see Table 32). Their behavior is well above chance performance so it cannot be all attributed to guessing. And yet all their reasoning is not semi-logical either. They do give proportionally more semi-logical explanations than the sample as a whole, but still not enough to account for all their correct behavioral responses. Semi-logical explanations are most prominent in the transitivity of weight task where many responses are justified by "A went down with B (or was heavier than B) so it will go down with C (or is heavier than C)". Actually this explanation need not really be semi-logical, i.e., S may not really be considering that A is actually heavier than B and therefore is heavier than C. He may merely have noted that A went down before when put on the balance with B and so he figures it always goes down (or he may reason similarly about C always going up). In this case he is merely assessing the probabilities of the situation and responding accordingly.

Table 32

Type of explanation given for correct behavioral responses in Tasks 1, 2 and 3 by Ss with perfect behavior but no logic

Type of explan.	Percentage of Responses Given: Trans length	Sertn length	Trans wght
Semi-logical	5.5	44.6	58.7
Other	94.5	55.4	41.3
n	18	14	25
# Ss giving all semi-log explans	0	2	8

These *S*s whose behavioral performance was perfect but who displayed no logic on tasks 1, 2 and 3 are not the same *S*s on each task: the probability of one of these *S*s who has No logic and Good behavior on one task falling into the same category on one of the other tasks is no greater than that for the sample as a whole. Thus, one cannot single out a group of *S*s who consistently perform behaviorally very well while displaying no logic. This type of performance just happens to occur on one task. The chances are very good that either their logical category or their behavior category or both will be different on the next task.

If with these same *S*s in each of the three tasks mentioned we compare their distributions among the categories of performances for the other tasks with the distribution of the entire sample using a Chi-square goodness-of-fit test we can see if they differ significantly from the whole sample in the categories in which they occur. Thus, each of the No logic, Good behavior groups of *S*s in the two transitivity tasks and the seriation of length task were compared with the entire sample on their categories of performance for logic and behavior on the other transitivity and seriation tasks combined and on all the other tasks 5 through 9 combined. The only significant difference in distribution found was that those *S*s who had No logic and Good behavior on the transitivity of length task displayed significantly less logic in all the other tasks combined ($\chi^2 = 9.94$, p .01).

This is a reasonable finding in view of the fact that in the sample as a whole more logic was displayed in this task than in any other. None of these groups of Ss show any difference in their performance on any of the other tasks.

The results are quite different in the seriation of weight task. Here less than one percent of all correct behavioral responses (correct seriations and insertions given Adequate weighings) are not justified with a logical explanation, and nearly all of these are semi-logical reasons. In this task there are many wrong behavioral responses which are given perfectly logical justifications. How this can happen has been discussed previously. Again one can single out for special comparison two groups of Ss to see if their performance on this task is part of a consistent pattern they may display on other tasks. There were five Ss who had Good logic but Poor behavior in the seriation of weight task. These five did not differ significantly from the sample as a whole in their performances on any of the other tasks. Also there were 20 Ss who had Good logic and Good behavior, and it was hypothesized that these Ss were the most advanced of the sample and would have superior performances on all the other tasks. When their distributions among the various categories on the other tasks are compared with those of the entire sample the only significant difference found is that they display significantly more logic on the other three transitivity and seriation tasks combined than does the sample as a whole ($\chi^2 = 13.3$, $p < .01$). Their performances are not significantly superior on the behavioral measures of these tasks nor on any of the other tasks 5 through 9.

Thus any search for specific groups of Ss who show consistent behavior of a given type across several tasks is quite fruitless. No aspect of performance on the tasks presented in this study is that clear-cut.

RELATIONSHIP BETWEEN PERFORMANCE IN LENGTH AND WEIGHT DIMENSIONS ON TRANSITIVITY AND SERIATION TASKS.

Table 33 presents the results pertaining to the third purpose of this study: to see the relationship in performances within the same Ss on transitivity and seriation tasks between the size and weight dimensions. The aim was to see in what order performance on these dimensions develops. Performance in the length dimension develops before that in the weight dimension on both measures of the transitivity and seriation tasks. The direction of all the relationships is quite stable--i.e., the ratios of Ss performing better on the length task to Ss performing better on the weight task are large in all four pairs of tasks. One could predict with a high degree of accuracy that whatever an S's category of performance on a weight task it will be the same or higher on an analogous length task. But that is all one can predict with respect to behavioral performance where predictive association between dimensions is nil (both $\lambda_{LW} = 0$). An S's behavior

Table 33

Relationship between performance in Length and Weight dimensions on Transitivity and Seriation tasks: summary of contingency tables.

Lgth vs Wgt on Measures	χ^2	P	# Ss on diagonal	Direction of relnshp	φ'	λ_{LW}
Trans - Log	37.92	<.001	45	Lgth>Wgt 46:2	.45	.19
Trans - Beh	0.89	N.S.	36	Lgth>Wgt 46:11	.10	0
Sertn - Log	58.43	<.001	69	Lgth>Wgt 20:6	.56	.58
Sertn - Beh	15.56	<.01	29	Lgth>Wgt 61:3	.29	0

category in one dimension is no more predictable when given his behavior category on the other dimension than when the latter is not known. This is true even though behavior performances in the two seriation tasks are not independent (χ^2 = 15.56, p<.01) and show some degree of association (φ' = .29). But a significant χ^2 result, especially with a relatively large sample like this one, does not mean that a relationship "observable to the naked-eye, which will be applicable in some real-world situation" has been discovered. Lambda does suggest, however, "just how much the relationship found implies about real predictions, and how much one attribute actually does tell us about the other" (Hays, 1963, p610).

On the other hand, logical ability is more closely associated across dimensions (φ' = .45 for transitivity, .56 for seriation). In the transitivity tasks logic in length develops much earlier than in weight (46:2 ratio of Ss who change categories between dimensions), but there is some predictability (λ_{LW} = .19) between dimensions in some categories: e.g., 28 out of 30 Ss with No logic in length have No logic in weight, and 16 out of the 17 Ss with Good logic in weight have Good logic in length. The degree of association between the logical performances on the seriation tasks is the highest one in the entire study by whatever means it is measured, - χ^2, number of Ss on the diagonal, φ' or λ_{LW}. Predictability between dimensions is high (λ_{LW} = .58) because over two-thirds of the sample (65 Ss) falls in either the Good-Good (32) or None-None (33) categories on both tasks: 32 of the 36 Ss with Good logic in weight also have Good logic in length; 33 of the 43 Ss with No logic in weight have No logic in length; 33 of the 37 Ss with No logic in length have No logic in weight. All these results indicate that knowing an S's logical category of performance in one dimension greatly reduces the probability of error in predicting his category in the other.

Logical abilities in the seriation tasks are the most nearly equal in development of all four relationships in Table 33. In both cases S is reasoning about objects already lined up in a series in front of him, one according to length the other according to weight, and as far as he knows each series is correct. Perhaps it is true that once the objects are lined up they could be along any dimension. Only the words used to express the relationships among the objects (i.e., "longer", or "heavier") are different; the reasoning for an object's placement in the series is the same. The very close relationship between the logical performances on these two tasks suggests that this is a reasonable notion.

Why do the operations of transitivity and seriation develop earlier in the size dimension than in weight? When an operation is revealed with respect to one sort of quality why isn't it immediately applicable to a second and a third? Piaget and Inhelder (1941) offer an explanation for their results, which are similar to ours, on the seriation of size and weight. According to these authors, when objects differ along a perceptual dimension the child can usually perceive this all at once--that is, it is apparent in his "field of simultaneous visual perception" (p240). But when a series of objects differ in weight this is not visually apparent and the field of simultaneous perception established by the hands or the balance which weight the objects is much more confined or limited than the visual field. Therefore, the seriation of weight will imply a much larger number of intellectual operations in order to coordinate the perceived relations. (One could add here that it would also seem to imply a more highly developed memory system in order to remember all the perceived relations). The obstacle to operational attainment is the child's egocentrism which makes him deal with absolutes rather than with relations. The qualities of weight remain egocentric longer than the visible dimensions. Vision puts the child in the universe of simultaneous givens where an appreciation of relations of size and space develops readily. He comes to understand rapidly that the big and small, the wide and narrow, etc., do not exist in themselves but only relative to other perceptually apparent things. But because the heavy and the light are more difficult to structure perceptually there occurs not only a delay but also a fixation of egocentric habits which render these qualities more resistant to relativity.

RELATIONSHIP BETWEEN TRANSITIVITY AND SERIATION.

Table 34 presents a summary of the contingency tables relating the Ss' performances on the transitivity and seriation tasks within the length and weight dimensions. In the length dimension behavioral performances on the two tasks are independent ($\chi^2 = 4.28$, N.S.) and there is no predictability between categories ($\lambda_{TS} = 0$). Transitivity develops a little before seriation but the ratio of Ss who change categories between tasks is not large (25:11). In the weight dimension behavioral performances are not independent ($\chi^2 = 10.59$, $p<.05$, $\varphi' = .24$) but again there is no predictability from one set of categories to another ($\lambda_{TS} = 0$). Those Ss with Good behavior in the transitivity task (49) are not divided among the seriation categories any differently than how the marginals in the contingency table would suggest. Similarly those Ss who have Poor behavior on the seriation task are divided among the transitivity categories in proportion to the marginal totals. Therefore, knowing a S's category on one task does not reduce the probability of error in predicting his category on the other. Although transitivity performance is much better than seriation performance, the ratio of Ss who change categories being large (54:5), not all those Ss with Good behavior on seriation have Good behavior on transitivity: out of 20 Ss with Good seriation behavior, 15 have Good, three have Fair, and two have Poor transitivity behavior. Good seriation behavioral performance usually

Table 34

Relationship between Transitivity and Seriation in Length and Weight dimensions: summary of contingency tables.

Trans vs Sertn of Tasks	χ^2	P	# Ss on diagonal	Direction of relnshp	φ'	λ_{TS}
Length - Log	31.45	<.001	59	Trans>Sertn 21:13	.41	.38
Length - Beh	4.28	N.S.	57	Trans>Sertn 25:11	.21	0
Wgt - Log	32.54	<.001	59	Sertn>Trans 31:3	.42	.27
Wgt - Beh	10.59	<.05	34	Trans>Sertn 54:5	.24	0

implies Good transitivity performance, but not always.

Logical abilities on the two tasks within each dimension are related (both χ^2's significant <.001 level; $\varphi' = .41$ and .42 for length and weight respectively). In the length dimension the fact that 57 out of the 58 Ss who fall on the diagonal occur in either the Good-Good (35) or None-None (22) cells helps account for the high degree of predictability between the transitivity and seriation categories ($\lambda_{TS} = .38$). A few more Ss display logical reasoning on the transitivity task, but the ratio of Ss changing categories between the tasks is very small (21:12). The relationship between the logical performances on the two weight dimension tasks is the only one in which seriation performance surpasses transitivity performance, and the direction of this relationship is very stable (the ratio of change is a large 31:3). The fair degree of predictability between the two ($\lambda_{TS} = .27$) is no doubt accounted for by the fact that 41 out of 43 Ss with No logic on the seriation task also have No logic on the transitivity task, and 16 out of 17 Ss with Good logic on transitivity also have Good logic on seriation. These results suggest a Guttman scale relationship between the logical abilities on these tasks: if S is logical on transitivity, he's logical on seriation, but not vice verse. But if he shows no logic on seriation, he shows none on transitivity either. It is thought that more logic is displayed with seriation than with transitivity because in the seriation task S is asked to give logical explanations about objects which are already lined up in order along the dimension (whether the order be correct or not does not matter since S thinks it is correct) and therefore presents S with a structured situation to contemplate, whereas in the transitivity task the three objects are scattered randomly on the table and S must rely on his memory of the original weighings in order to respond logically.

Piaget and Inhelder (1941) found that at the behavioral level operational seriation and transitivity of weight were achieved around the same time. That is, it was not until S could correctly seriate up to 10 objects by weighing them two at a time in his hands or on a balance that he could also answer transitivity questions about three objects of different weights. No justifications were requested for any of the Ss' responses. This operational behavior was achieved around nine or ten years of age. In the present study transitivity behavior is much easier than seriation. The present seriation task is very similar to Piaget and Inhelder's, but the transitivity task here is probably less demanding than theirs. In their task the three objects are not actually weighed in front of S: he is merely told by E that A<B, and B<C and he must designate the heaviest and lightest. The information is also presented in two other ways--A>B, C<B, and B<A but >C--and S must respond correctly in all three situations to be considered operational. In the transitivity if weight task employed in this research the three sets of stimuli are always weighed on the balance for S and, as noted earlier, he can derive his response by means other than transitive reasoning. This seems much less likely in Piaget and Inhelder's task and probably accounts for the later development of transitivity in their study.

Murray and Youniss (1968) in comparing the development of transitivity and seriation of length using behavioral measures only, found that the seriation task was passed by more of their Ss, aged 5½ to 8 years, than the transitivity task. Their easy seriation task required S merely to line up five objects initially and then insert one more. Their transitivity task had four trials altogther in which the middle stick B was either of intermediate length between A and C, equal to A, equal to C or absent altogether. Three out of four trials correct was considered a "pass". It would appear just at face value that the transitivity task was quite a bit harder than the seriation task and this could account for the direction of their results. Battisti and Simmons (1968) also compared transitivity and seriation of length performance at the behavioral level and found their transitivity task to be easier than the seriation task. In the seriation task S had to line up eight sticks initially and then insert eight more. In the transitivity task E compared B with A and C, whose lengths were appropriately disguised, and S had to say which was longer, A or C, and designate the correct order of all three stimuli. It would appear that perhaps these tasks are more nearly comparable in difficulty than those used by Murray and Youniss.

All these studies, and the present one, point up the experimental problems one inevitably runs into in trying to compare abilities across different tasks. How do you know if the tasks are of equal difficulty to begin with? That is, does each survey the operation it is meant to be testing with the same thoroughness? Does each probe the Ss' depth and breadth of understanding of the operations to the same degree? These are extremely difficult questions to answer and the diversity of experimental situations employed to date and the discrepancies among their results indicate that no one has found the solution yet.

RELATIONSHIP BETWEEN THE CLASSIFICATION AND SERIATION OPERATIONS.

In this study we presented two of Inhelder and Piaget's measures of classification ability both of which are claimed to be tests of the understanding of class-inclusion, namely, the All-Some task and the Part Whole task. The first row in Table 35 presents a summary of the measures of association made on the contingency tables relating the Ss' categories of performance on these two tasks. The Chi-square test reveals them to be independent ($\chi^2 = 8.79$, N.S.), the degree of association between them is low ($\varphi' = .22$), and the pre-

Table 35

Relationship between Classification and Seriation-Transitivity operations. 1. Task 5: All-Some compared with Seriation and Transitivity tasks

All-Some vs	χ^2	P	#Ss on diagonal	Direction of relnship	φ'	λ_{LB}
Part-Whole	8.79	N.S.	37	A-S>P-W 39:17	.22	.02
Len-Sertn-Log	3.86	N.S.	35	Ser>A-S 32:26	.14	.03
Len-Sertn-Beh	5.97	N.S.	33	Ser>A-S 47:13	.18	0
Len-Tran-Log	3.90	N.S.	32	Tran>A-S 36:25	.14	0
Len-Tran-Beh	0.42	N.S.	26	Tran>A-S 57:10	.07	0
Wgt-Sertn-Log	5.59	N.S.	28	A-S>Ser 33:31	.17	.05
Wgt-Sertn-Beh	2.43	N.S.	32	A-S>Ser 44:17	.11	0
Wgt-Tran-Log	5.38	N.S.	31	A-S>Tran 47:15	.17	0
Wgt-Tran-Beh	6.92	N.S.	27	Tran>A-S 47:19	.19	.02

dictive association between them practically nil ($\lambda_{CS} = .02$). The All-Some task is easier than the Part-Whole: 39 Ss perform better on All-Some while only 17 perform better on Part-whole. This lack of psychological association between these two tasks might suggest that class inclusion is not a unitary operation and that these two tasks are measuring different aspects of it. At any rate it means that they cannot be regarded as one measure and therefore they will have to be compared with the other tasks in the study separately. Similarly since we have just seen how seriation and transitivity abilities can vary within the same Ss, these tasks will have to be treated separately also from now on in comparing them with the others in the study.

Table 35 presents a summary of the relationship between performances on the All-Some task and the seriation and transitivity tasks. The Chi-square scores reveal that performance on the All-Some task is independent of ability on all the other tasks. There is no predictive association between it and any of the other tasks. In general, performance on length dimension tasks is superior to All-Some performance, and performance in the weight dimension is inferior (except for behavioral transitivity). But these relationships are very unstable: there are few Ss on the diagonals and the ratios of Ss whose performances change between tasks are not large.

It can be concluded on the basis of these results that the classification operation, when measured by the All-Some class inclusion task, is psychologically unrelated to both the seriation and transitivity operations, measured in size and weight dimensions, as they are understood both behaviorally and logically by the child. Furthermore, since All-Some performance does not improve with increasing grade, it is doubtful that it is a measure of class-inclusion understanding at all, if it is to be assumed that this understanding increases across the age range tested here.

Table 36 presents a summary of the contingency tables relating the Part-Whole classification task to the transitivity and seriation tasks. It is evident from the Chi-square scores that Part-Whole performance is not independent of performance on these tasks as All-Some performance was. The degree of association between Part-Whole and the other tasks is fair (as measured by φ'), but the predictive association between them is poor (as measured by λ_{CS}). Thus, while performance on the tasks are not independent, one cannot predict much about an S's ability on one of the tasks by knowing his level of ability on the other.

The Part-Whole task would appear to be a real test of class-inclusion since S is asked to compare the parts and wholes in a hierarchical set-up with objects which have meaning for him. Inhelder and Piaget (1964) claim that classification ability, an aspect of which is the understanding of class-inclusion, develops around the same time as the seriation of length. Table 36 reveals that in the present study

Table 36

Relationship between Classification and Seriation-Transitivity operations. 2. Task 6: Part-Whole compared with Seriation and Transitivity tasks

Part-Whole vs	χ^2	P	# Ss on diagonal	Direction of relnship	φ'	λ_{CS}
Len-Ser-Log	11.15	<.05	42	Ser>P-W 38:13	.24	.14
Len-Ser-Beh	17.47	<.01	29	Ser>P-W 61:3	.31	.03
Len-Tran-Log	21.97	<.001	43	Tran>P-W 43:7	.34	.20
Len-Tran-Beh	3.91	N.S.	26	Tran>P-W 63:4	.20	0
Wgt-Ser-Log	13.14	<.02	45	Ser>P-W 33:15	.27	.19
Wgt-Ser-Beh	10.73	<.05	49	P-W>Ser 26:18	.24	.05
Wgt-Tran-Log	23.26	<.001	57	P-W>Tran 26:10	.35	.15
Wgt-Tran-Beh	16.15	<.01	39	Tran>P-W 50:4	.29	0

the length dimension tasks are easier than the Part-Whole task and all the relationships are quite stable since the ratios of Ss changing categories are quite large. There is no clear-cut direction of the relationships with the weight dimension tasks: Part-Whole ability develops earlier or later than transitivity and seriation of weight depending on whether one looks at behavioral or logical measures of these two tasks. This lack of direction might suggest that Part-Whole performance develops closer in time with weight dimension tasks than with length dimension tasks which develop earlier.

Part-Whole performance seems to be related to both seriation and transitivity tasks to the same degree. Within these tasks, however, there is more predictive association with the logical measures than with the behavioral measures, although no λ_{CS} is very large. Like the logical measures the Part-Whole responses require verbalization of the operation involved and this basis of similarity between the tasks might account for some of the predictability between them.

SIMULTANEOUS COMBINATIVITY AS A COMPONENT ABILITY IN CLASSIFICATION AND TRANSITIVITY OPERATIONS.

To test the hypothesis that one of the logical abilities underlying both the classification and transitivity operations is the ability to understand that one object can possess two qualities, roles, attributes, relations, or whatever at the same time, the One Object-Two Qualities task was devised. Table

37 presents a summary of the relationships between this task and those measuring the seriation, transitivity, and classification operations. The significance of the Chi-square scores indicates that this task is not independent of these other operations. However, the degree of association between it and the other tasks is not great (φ' ranges from .21 to .37) and the predictive association between them is low. This weak relationship is also reflected in the small number of Ss falling on the diagonals of the contingency tables relating their categories of performance on the One Object-Two Qualities task with those on the other tasks. It would appear that most of the tasks are easier than the One Object-Two Qualities task, but the only places where the direction of the relationship is stable are in the two behavioral transitivity measures and in the behavioral seriation of length measure (the ratios of Ss changing categories between tasks are very large here).

By the measures we have used one could say that understanding of simultaneous combinativity is certainly not independent of the operations of classification, seriation, and transitivity, but that the strength of the associations shown between abilities on these measures would certainly not lead one to conclude that it is a crucial underlying logical ability of these operations.

Most of the stimuli used in the One Object-Two Qualities task have logical structures similar to All-Some task stimuli.

Table 37

Test for simultaneous combinativity: relationship between One Object-Two Qualities task and Classification-Transitivity operations.

1 Obj-2 Quals vs	χ^2	P	# Ss on diagonal	Direction of relnship	φ'	λ Si-CT
Len-Ser-Log	9.93	<.05	30	Ser>1 Ob-2Q 43:20	.23	.10
Len-Ser-Beh	25.70	<.001	29	Ser>1 Ob-2Q 62:2	.37	.12
Len-Tran-Log	19.35	<.001	37	Tran>1 Ob-2Q 44:12	.32	.20
Len-Tran-Beh	9.52	<.01	20	Tran>1 Ob-2Q 68:5	.31	.09
Wgt-Ser-Log	20.39	<.001	42	Ser>1 Ob-2Q 31:20	.33	.18
Wgt-Ser-Beh	19.27	<.001	43	1 Ob-2Q>Ser 34:16	.32	.05
Wgt-Tran-Log	14.90	<.01	41	1 Ob-2Q>Tran 38:14	.28	0
Wgt-Tran-Beh	8.42	N.S.	34	Tran>1 Ob-2Q 50:9	.21	.05
All-Some	11.75	<.02	30	A-S>1 Ob-2Q 41:22	.25	.05
Part-Whole	22.72	<.001	48	1 Ob-2Q>P-W 28:17	.35	.07

The questions asked about the stimuli are phrased differently though: in the former task S is asked "Which are there more of, a's or b's?" whereas in the latter task he is asked "Are all the A's B's?". The results in Table 37 indicate that they are psychologically related to some degree (χ^2 = 11.75, p<.02; φ' = .25), and although the ratio of Ss changing category indicates that the All-Some task is easier than the One Object-Two Qualities task, this ratio is not large and they could be developing around the same time. But performance on the One Object-Two Qualities task is not independent of performance on the other tasks whereas All-Some performance is. Also, One Object-Two Qualities ability improves with grade; All-Some does not. Therefore, the One Object-Two Qualities task is tapping some ability which improves with age and is related to other concrete operations, but the All-Some task is not.

SUCCESSIVE COMBINATIVITY AS A COMPONENT ABILITY IN CLASSIFICATION AND TRANSITIVITY OPERATIONS.

To test the hypothesis that another of the possible logical abilities underlying the classification and transitivity operations is the ability to link together two successively presented pieces of information or behaviors in order to derive a third and crucial piece of information, the Kendlers' Inference task was presented to the Ss as well as another task devised for this study having the same logical structure as the Kendlers' task. Table 38 presents the relationship between the Inference I task (Kendlers') and all the other tasks in this study. The results presented in the first row indicates that the two Inference tasks are tapping two quite independent abilities (χ^2 is N.S., φ' is low, and λ_{Su-CT} is practically nil), and the ratio of Ss changing categories between the two tasks indicates that they are of practically equal difficulty. Thus, even though the two inference tasks possess very similar logical structures, they are psychologically unrelated. That is, a child's ability to correctly solve one of the tasks has absolutely nothing to do with how he performs on the other.

The second row of Table 38 reveals that the same thing is true with regard to the relationship between successive and simultaneous combinativity abilities, as measured by the Inference I and One Object-Two Qualities tasks respectively. They are also of nearly equal difficulty, but are quite independent of each other.

The rest of Table 38 likewise indicates that performance on Inference I bears little or no relationship with performance on any of the other tasks in the study. The results presented in Table 39 indicate exactly the same thing about performance on the Inference II task. Thus one might well conclude that ability with successive combinativity, as measured by the Inference tasks presented in this study, plays no part in the development of the operations of seriation, transitivity, or

Table 38

Test for successive combinativity. 1. Relationship between Inference I task and Classification-Transitivity operations.

Inference I vs	χ^2	P	# Ss on diagonal	Direction of relnship	φ'	λ_{Su-CT}
Inference II	5.30	N.S.	19	Inf I>Inf II 35:29	.17	.03
1 Obj-2 Qual	8.77	N.S.	27	Inf I>1 Ob-2Qu 37:29	.22	.04
Len-Ser-Log	6.59	N.S.	41	Ser>Inf I 31:21	.19	.08
Len-Ser-Beh	4.07	N.S.	34	Tran>Inf I 46:13	.15	.01
Len-Tran-Log	5.68	N.S.	35	Tran>Inf I 38:20	.17	.03
Len-Tran-Beh	23.63	<.001	41	Tran>Inf I 48:4	.50	.04
Wgt-Ser-Log	3.65	N.S.	38	Ser>Inf I 29:26	.14	.07
Wgt-Ser-Beh	2.69	N.S.	38	Inf I>Ser 37:18	.12	.02
Wgt-Tran-Log	13.18	<.02	44	Inf 1>Tran 36:13	.27	.10
Wgt-Tran-Beh	2.95	N.S.	27	Tran>Inf I 45:21	.13	0
All-Some	5.30	N.S.	37	Inf I>A-S 26:20	.17	.04
Part-Whole	4.25	N.S.	32	Inf I>P-W 46:25	.15	.02

Table 39

Test for successive combinativity. 2. Relationship between Inference II task and Classification-Transitivity operations.

Inference II vs	χ^2	P	# Ss on diagonal	Direction of relnship	φ'	λ_{Su-CT}
1 Obj-2 Qual	3.96	N.S.	34	1 Ob-2Q>Inf II 32:27	.14	.02
Len-Ser-Log	5.62	N.S.	41	Ser>Inf II 35:17	.17	.06
Len-Ser-Beh	5.99	N.S.	27	Ser>Inf II 56:10	.18	0
Len-Tran-Log	6.60	N.S.	39	Tran>Inf II 41:13	.19	0
Len-Tran-Beh	2.70	N.S.	29	Tran>Inf II 59:5	.17	0
Wgt-Ser-Log	10.12	<.05	46	Ser>Inf II 28:19	.23	.10
Wgt-Ser-Beh	4.59	N.S.	42	Inf II>Ser 32:19	.16	.02
Wgt-Tran-Log	6.00	N.S.	42	Inf II>Tran 35:16	.18	.01
Wgt-Tran-Beh	1.86	N.S.	30	Tran>Inf II 51:12	.10	0
All-Some	8.85	N.S.	38	A-S>Inf II 33:22	.22	.03
Part-Whole	15.53	<.01	47	Inf II>P-W 26:20	.29	.18

classification, and is also independent of the ability to cope with simultaneous combinativity as operationally defined here.

ORDER OF DIFFICULTY OF THE TASKS.

It is evident from the foregoing analyses that all the tasks are not of the same level of difficulty for the children. In order to get an idea of what the order of difficulty among the tasks might be the following was done. First of all, the three categories of performance for each measure were assigned the arbitrary scores of 1, 0 and -1 for the categories labelled Good, Fair or Some, and Poor or None respectively. (For the behavioral measure of the transitivity of length task which was divided into only two categories initially the Ss in the "Poor" category were divided into two categories for present purposes, the 13 Ss with no correct responses assigned scores of zero, and the two Ss with no correct responses assigned scores of minus one). These scores were multiplied by the number of Ss in each category and divided by 93, the total number of Ss, to give a mean score lying somewhere between -1 and +1 for each measure. The means for the tasks arranged in order of difficulty from the easiest to most difficult are presented in the margins of Table 40. The behavioral aspects of the two length dimension tasks are the easiest and their logical components are also easy compared with most of the other tasks (they rank 4 and 5). The two most

Table 40

Order of difficulty of the tasks

Task / Mean	1. L-T Beh. +.82	2. L-S Beh. +.53	3. W-T Beh. +.35	1. L-T Log. +.19	2. L-S Log. +.08	4. W-S Log. +.06	5. A-S -.02	6. Inf I -.06	7. 10b-2Q -.23	9. Inf II -.25	6. P-W -.32	4. W-S Beh. -.41	3. W-T Log. -.53
1. L-T Beh +.82													
2. L-S Beh +.53	.29												
3. W-T Beh +.35	.47**	.18											
1. L-T Log +.19	.63**	.34	.16										
2. L-S Log +.08	.74**	.45**	.27	.11									
4. W-S Log +.06	.76**	.47**	.29	.13	.02								
5. A-S -.02	.84**	.55**	.37	.21	.10	.08							
8. Inf I -.06	.88**	.59**	.41*	.25	.14	.12	.04						
7. 10b-2Q -.23	1.05**	.76**	.58**	.42*	.31	.29	.21	.17					
9. Inf II -.25	1.07**	.78**	.60**	.44**	.33	.31	.23	.19	.02				
6. P-W -.32	1.14**	.85**	.67**	.51**	.40*	.38*	.30	.26	.09	.07			
4. W-S Beh -.41	1.23**	.94**	.76**	.60**	.49**	.47**	.39*	.35	.18	.16	.08		
3. W-T Log -.53	1.35**	1.06**	.88**	.72**	.61**	.59**	.51**	.47**	.30	.28	.21	.12	

** $p < .01$

* $p < .05$

difficult tasks are the behavioral measure of the seriation of weight task and the logical measure of the transitivity of weight task.

In order to test the significance of the differences between these means Tukey's method of employing the studentized range to find the necessary critical differences between means was used (see Winer, 1962, p87). The critical differences arrived at for these means were .38 for p = .05 and .44 for p = .01. The body of Table 40 presents the differences and indicates the level of significance they attain.

Behavioral performance on both of the transitivity tasks and the seriation of length task is significantly better than logical performance on each task. The opposite is true for the seriation of weight task.

Within each operation of transitivity and seriation performance in the length dimension is significantly better than that in the weight dimension with all measures except logical performance on the seriation tasks which are very similar. In general, within each dimension performance on the transitivity task is superior to performance on the seriation task, although most of these differences do not attain significance. The one glaring exception to this trend is of course logical performance in the weight dimension where significantly more logic is displayed in the seriation than in the transitivity task.

As for the two measures of class inclusion, the All-Some task is easier than the Part-Whole task but not significantly so. The tests of combinativity, both successive and simultaneous, are of similar degrees of difficulty.

CHAPTER FOUR

Disccussion and Conclusions

The problem of whether to use behavioral or verbal measures when testing for the presence of a concrete operation in children has been examined in this study. The question of what to do with each type of response naturally arises. One could decide that behavioral responses alone are sufficient to indicate operational understanding and leave it at that. However, it is quite probable in many experimental situations that correct behavioral responses can be achieved by guessing, by perceptual discrimination, and many tactics other than truly operational reasoning. Thus the error of diagnosing many false positives--i.e., children who appear to be operational but really are not--could arise in this situation. On the other hand, one could depend solely on verbal logical reasoning as the criterion of operational understanding. But then loud cries are heard on behalf of those children who "lack verbal sophistication" or do not understand exactly what is expected of them in the testing situation. These children may really be operational, it is argued, but are unable to demonstrate this fact. A third approach is to consider the behavioral and verbal responses as integral and to discount those behavioral responses which are not given logical explanations and vice verse. This approach is admirably cautious but does neglect much data. For instance, what about those children who are behaviorally competent but logically unsound? Or, how does one account for children who are verbally logical but behaviorally confused (as in the seriation of weight task in the present study)? To what level of development are these children assigned? Or should they be considered unreliable and ignored?

Solutions to this problem do not come easily and it was one purpose of the present research to examine the issue more closely. We chose to look at each measure separately and to examine how each relates to measures on other tasks and to each other within tasks before making any judgements as to which is the best measure or combination of measures. This approach has uncovered some heretofore unexamined aspects of concrete operational thinking. It also has revealed several reasons for suspecting behavioral responses alone as revealing operational understanding.

Operational transitivity implies the ability to use the two relationships A>B and B>C to derive the third A>C. Operational seriation implies the understanding that each object in a series possesses more of the relevant quality than all those preceding it and less than all those following it in the series. Verbal logical justifications of behavioral responses in tasks designed to test this understanding can only reveal truly operational understanding. But correct behavioral responses can be derived from bases other than operational reasoning, e.g., perceptual discrimination, semi-logical reasoning, probability assessment of the experimental situation, or luck. All these things confound the assessment of operational understanding at the behavioral level. This is not to deny the possible existence of a non-verbal, intuitive level of operational ability. But the behavioral measures used here and in other studies are not necessarily measuring it: they are ambiguous measures.

First of all, behavioral responses are not always given logical, or even semi-logical, explanations and it seems they can be based on non-verbal intuitive reasoning which is impossible for an examiner and even for the child himself to decipher. Further, if one assumes that performance on the transitivity and seriation tasks should be related at least to some degree if only because understanding of each is increasing over the age range tested, the behavioral measures are disappointing. Children with perfect behavior but no logic in one task do not necessarily show the same pattern of performance in another. Therefore, even perfect behavioral performance in one task may not necessarily reflect a general and stable understanding of an operation. Behavioral ability is not predictable across dimensions: transitivity performances in the length and weight dimensions are completely independent and unpredictable; seriation ability is related but also unpredictable from one dimension to the other. On the other hand, logical performances across dimensions are highly associated ($\varphi' = .45$ for transitivity, .56 for seriation) and predicting a child's category of performance in one dimension is aided by knowing his category in the other to some extent for transitivity ($\lambda_{LW} = .19$) and to a large extent for seriation ($\lambda_{LW} = .58$). Behavioral ability is also not predictable across tasks within dimensions: within both the length and weight dimensions predictability regarding behavioral performances between transitivity and seriation tasks is nil. But logical ability is predictable ($\lambda_{TS} = .38$ for length, .27 for weight) and the degree of association across tasks is high ($\varphi' = .41$ for length, .42 for weight). Thus the logical measure shows more stability and reliability both across tasks and across dimensions than the behavioral measure does which would tend to suggest that it is tapping a more general and well-rooted ability than that revealed through behavioral performances. Furthermore, there is more predictive association between classification understanding (as measured by the Part-Whole task) and the logical measures of transitivity and seriation than the behavioral measures. According to the order of difficulty of the tasks behavioral performances on the transitivity of length and weight and seriation of length tasks are significantly better than logical performances on each of these tasks. The opposite is the case for seriation of weight. Thus a child may or may not appear to understand the operation depending upon whether it is measured verbally or behaviorally. One cannot say that one type of measure will always be a consistently earlier (or later) indicator of the presence of the operation than the other measure since the direction of the relationships between the measures is not the same in all tasks. If the behavioral measures were as highly related across operation and dimensions as the logical measures then the importance of the confounding factors mentioned above could easily be discounted. But this is not the case. All this evidence demonstrates that either behavioral measures are not "pure" measures of operational understanding or that there is a non-verbal operational ability which does not generalize across tasks in the same way that verbal logical ability does. Until the confounding factors can be controlled for, the behavioral measures should be supplemented with verbal measures in these tasks.

Among the logical measures for transitivity of length and weight and seriation of weight, one is not significiantly easier than another and all show considerable degrees of association amongst each other. These facts suggest that these abilities develop around the same time. But all are significantly easier than logical performance in the transitivity of weight task which is the latest developer of all measures in this study. Why is this so? The reason probably lies primarily in the experimental situation. In the transitivity of weight task the three objects are weighed two at a time by

E and are then placed randomly on the table in front of S. Their arrangement represents no order among the stimuli and is different for each of the three sets of objects. Thus when S is asked to justify his behavioral transitivity response he must reason from his memory of the original weighings of the objects or from whichever way he concluded A>C in the first place. There is nothing in the situation to help him remember, or indeed perceive for the first time, the relationships among the objects. This is not true in the other three transitivity and seriation tasks. In the latter the objects are lined up in order in front of S and when asked to justify his behavioral responses he can reason from the relationships revealed in these orders and does not have to go back to his (probably faulty) memory of how he discovered this order in the first place. That this is in fact the case is evident from the seriation of weight task where, because S does not know if his series is correct or not, many logical reasons are given for the incorrect placement of objects in the series. In the Transitivity of length task the objects are spread apart but B is always physically midway between A and C when S is asked how he knows A>C, i.e., A and C do form the extremes of the three object continuum. Thus the physical facts of the situations no doubt aid logical thinking in these latter three cases but hinder it in the former.

This finding is just one of several in this study which indicate that children perform differently in what we can label as structured and unstructured situations. For instance, in the seriation of length task insertion ability was (nonsignificantly) better than seriation ability, especially among the Kindergartners. (This contrasts with some findings of Piaget (1952) who observed that five year olds could make a correct series of many sticks after considerable trial and error but were unable to insert a second set of sticks within the completed series). In the seriation of weight task the percentage of correct insertions given Adequate weightings is significantly higher than the percentage of correct seriations given Adequate weighings in every grade. These results suggest that there is a difference in a child's ability to behave appropriately in a situation where he must organize the objects in his environment himself, such as lining up objects correctly along a dimension, compared to when they are already organized for him and he has merely to relate to them in some well defined manner, such as inserting a few objects appropriately into already constructed line-ups. This difference is relfected in logical ability too, as implied above, since many children can reason logically about objects which are lined up in front of them but are at a loss when the objects are randomly scattered as in the transitivity of weight task. Also in the seriation of weight task the fact that there are several children who cannot organize the objects adequately initially to seriate or insert them correctly but who can nevertheless reason logically once the objects are ordered is further evidence of this dichotomy in performance between structured and unstructured situations.

Transitivity and seriation, as measured here, are not the same operation. At the behavioral level they are barely related at all, and transitivity seems a little easier to grasp than seriation but not significantly so. At the verbal level of logical understanding they are related but the order of their development is not clear cut. Thus although they may both have a logical basis in Piaget's Grouping V, psychologically they lack sufficient association to be considered synonymous operations. Understanding of each operation is achieved earlier in the length dimension than in the weight dimension and only logical ability is related across the two dimensions. An understanding of class-inclusion appears to develop closer in time with transitivity and seriation of weight than of length, but is not closely associated with the development of either. The ability to combine two successively presented pieces of information to derive a third does not appear to be a component ability of the transitivity, seriation, or classification operations. However, the ability to regard one object as simultaneously possessing two qualities or relationships does appear to be at least related in its development to these operations. Whether or not it is a crucial component of them cannot be answered definitely on the basis of the present results.

Finding psychological relationships among operations which according to Piaget are logically and structurally related is not easy. One reason may be because in the age range usually tested, five to eight years, what operational understanding a child does have is very unstable and influenced greatly by the experimental situation. Different experimental procedures lead to different results as to what is related to what, and which operations develop before others. The task of differentiating between competence and automaton models of cognition is a crucial one. We cannot know a child's true competence level until we solve the problems posed by the automaton aspect of his intelligence. In order to do so, however, must we resort to the type of experimental set-up suggested by Smedslund (1964) where everything is identical in the tasks except the inference patterns being compared? Even equating the tasks physically and procedurally does not guarantee that the operations are being examined to extents equal in depth, generality and difficulty. Furthermore, the present research has shown that within each task different measures yield different results as to when an operation develops and to what others it is related. But all the difficulties in confirming Piaget's theory may not be the fault of methodology. Even when all these experimental problems are overcome, we may find in the end that human intelligence does not develop according to structures devised by logicians, whose minds are presumably already fully developed and therefore capable of structuring reality in a myriad of manners.

References

Battisti, C. and Simmons, R. Transitivity and seriation performance for five, six and seven year old children. Child Study Center Bull., State Univ. Coll. at Buffalo, 1968, 4 (3), 64-69.

Beard, R. An investigation of concept formation among infant school children. Unpublished Ph.D. dissertation, University of London, 1957.

Braine, M. D. S. The ontogeny of certain logical operations: Piaget's formulation examined by nonverbal methods. Psychol. Monogr., 1959, 73, No. 5. (Whole No. 475)

Braine, M. D. S. Piaget on reasoning: a methodological critique and alternative proposals. In W. Kessen and C. Kuhlman (Eds.), Thought in the young child. Monogr. Soc. Res. Child Develpm., 1962, 27, (2, Serial No. 83)

Braine, M. D. S. Development of a grasp of transitivity: a reply to Smedslund. Child Develpm., 1964, 35, 799-810.

Braine, M. D. S. The ontogeny of certain logical operations: Piaget's formulation examined by nonverbal methods. In I. E. Sigel and F. H. Hooper (Eds.), Logical thinking in children. New York: Holt, Rinehart and Winston, 1968, Pp. 164-206.

Bruner, J. S. Inhelder and Piaget's "The Growth of Logical Thinking". II. A psychologist's viewpoint. Brit. J. Psychol., 1959, 56, 363-370.

Bruner, J. S. The course of cognitive growth. *Amer. Psychologist*, 1964, *19*, 1-15.

Dodwell, P. C. Children's understanding of number and related concepts. *Canad. J. Psychol.*, 1960, *14*, 191-205.

Dodwell, P. C. Relation between the understanding of the logic of classes and of cardinal number in children. *Canad. J. Psychol.*, 1962, *16*, 152-160.

Dodwell, P. C. Children's understanding of spatial concepts. *Canad. J. Psychol.*, 1963, *17*, 141-161.

Donaldson, M. Review of J. Piaget and B. Inhelder "La Genèse des Structure Logique Élémentaires: Classification et Seriations". *Brit. J. Psychol.*, 1960, *51*, 181-184.

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

Flavell, J. H. and Wohlwill, J. F. Formal and functional aspects of cognitive development. In D. Elkind and J. H. Flavell (Eds.), *Studies in cognitive development*. New York: Oxford University Press, 1969, Pp. 67-120.

Furth, H. G. *Piaget and knowledge*. Englewood Cliffs, N. J.: Prentice-Hall, 1969.

Gruen, G. E. Note on conservation: methodological and definitional considerations. *Child Develpm.*, 1966, *37*, 977-983.

Hays, W. L. *Statistics for psychologists*. New York: Holt, Rinehart and Winston, 1963.

Hunt, J. McV. Intelligence and experience. New York: Ronald, 1961.

Hunt, J. McV. The impact and limitations of the giant of developmental psychology. In D. Elkind and J. H. Flavell (Eds.), Studies in cognitive development. New York: Oxford Univer. Press, 1969, Pp. 3-66.

Hyde, D. M. An investigation of Piaget's theories of development of the concept of number. Unpublished Ph.D. dissertation, University of London, 1959.

Inhelder, B. Comments in discussion on the criteria of the stages of mental development. In J. M. Tanner and B. Inhelder (Eds.), Discussions on child development. Volume I. The first meeting of the World Health Organization Study Group on the psychological development of the child, Geneva, 1953. London: Tavistock Publications, 1956, Pp. 75-107.

Inhelder, B. and Piaget, J. The early growth of logic in the child: classification and seriation. Translated by E. A. Lunzer and D. Papert. New York: Harper and Row, 1964.

Inhelder, B. and Piaget, J. La genèse des structures logiques élémentaires: classifications et seriations. Neuchatel: Delachaux et Niestlé, 1959.

Kagan, J. and Lemkin, J. Form, color, and size in children's conceptual behavior. Child Develpm., 1961, 32, 25-28.

Kendler, H. H. and Kendler, T. S. Vertical and horizontal processes in problem solving. Psychol. Rev., 1962, 69, 1-16. (a)

Kendler, T. S. and Kendler, H. H. Inferential behavior in children as a function of age and subgoal constancy. J. exp. Psychol., 1962, 64, 460-466. (b)

Kofsky, E. A scalogram analysis of classificatory development. Child Develpm., 1966, 37, 191-204.

Lovell, K. and Ogilvie, E. A study of the conservation of substance in the junior school child. Brit. J. educ. Psychol., 1960, 30, 109-118.

Lovell, K. and Ogilvie, E. A study of the conservation of weight in the junior school child. Brit. J. educ. Psychol., 1961, 31, 138-144.

Lunzer, E. A. Some points of Piagetian theory in the light of experimental criticism. J. child Psychol. Psychiat., 1960, 1, 191-202.

Murray, J. P. and Youniss, J. Achievement of inferential transitivity and its relation to serial ordering. Child Develpm., 1968, 39, 1259-1268.

Pascual-Leone, J. Piaget's period of concrete operations and Witkin's field dependence: a study on college students and children. Mimeographed manuscript, 1967.

Piaget, J. The child's conception of number. New York: Humanities, 1952.

Piaget, J. Logic and psychology. Manchester: Manchester Univer. Press, 1953.

Piaget, J. Language and intellectual operations. In H. G.

Furth, Piaget and knowledge. Englewood Cliffs, N.J.: Prentice-Hall, 1969, Pp. 121-130.

Piaget, J. and Inhelder, B. Le développement des quantités chez l'enfant. Neuchatel: Delachaux et Niestlé, 1941.

Piaget, J., Inhelder, B. and Szeminska, A. The child's conception of geometry. New York: Basic Books, 1960.

Pinard, A. and Laurendeau. "Stage" in Piaget's cognitive-developmental theory: exigesis of a concept. In D. Elkind and J. H. Flavell (Eds.), Studies in cognitivie development. New York: Oxford Univer. Press, 1969, Pp. 121-170.

Razran, G. H. S. Conditioned responses in children: a behavioral and quantitative critical review of experimental studies. Arch. Psychol., 1933, No. 148.

Shantz, C. U. A developmental study of Piaget's theory of logical multiplication. Merrill-Palmer Quart., 1967, 13, 121-137.

Sigel, I. E. Reflections. In I. E. Sigel and F. H. Hooper (Eds.), Logical thinking in children. New York: Holt, Rinehart and Winston, 1968, Pp. 504-528.

Smedslund, J. The acquisition of conservation of substance and weight in children: I. Introduction. Scand. J. Psychol., 1961, 2, 11-20.

Smedslund, J. Development of concrete transitivity of length in children. Child Develpm., 1963, 34, 389-405.

Smedslund, J. Concrete reasoning: a study of intellectual development. Monogr. Soc. Res. Child Develpm., 1964, 29, No. 2 (Serial No. 93)

Suchman, R. G. and Trobasso, T. Color and form preference in young children. J. exp. child Psychol., 1966, 3, 177-187.

Uzgiris, I. C. Situational generality of conservation. Child Develpm., 1964, 35, 831-841.

Vygotsky, L. F. Thought and language. Cambridge, Mass.: M.I.T. Press, 1962.

White, S. Evidence for a hierarchical arrangement of learning processes. In L. P. Lipsitt and C. C. Spiker (Eds.), Advances in child development and behavior. Vol. 2. New York: Academic Press, 1965, Pp. 187-220.

Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1962.

Wohlwill, J. Piaget's system as a source of empirical research. Merrill-Palmer Quart., 1963, 9, 253-262.

Wohlwill, J. F. Readiness, transfer of learning and the development of cognitive structures. Paper presented at the meetings of the Canad. Psychol. Assoc., Montreal, 1966. (a)

Wohlwill, J. F. Piaget's theory of the development of intelligence in the concrete-operations period. In M. Garrison (Ed.), Cognitive models and development in mental retardation. _Amer. J. Ment. Def. Monogr. Suppl._, 1966, _70_(4), 57-78. (b).