An Investigation of Methodological Problems in the Cognitive Developmental Diagnosis of Transitivity

Elizabeth S. Thayer
University of Rhode Island

Follow this and additional works at: https://digitalcommons.uri.edu/oa_diss

Recommended Citation
Thayer, Elizabeth S., "An Investigation of Methodological Problems in the Cognitive Developmental Diagnosis of Transitivity" (1978). Open Access Dissertations. Paper 1079.
https://digitalcommons.uri.edu/oa_diss/1079

This Dissertation is brought to you for free and open access by DigitalCommons@URI. It has been accepted for inclusion in Open Access Dissertations by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons@etal.uri.edu.
AN INVESTIGATION OF METHODOLOGICAL PROBLEMS IN
THE COGNITIVE DEVELOPMENTAL DIAGNOSIS
OF TRANSITIVITY

BY

ELIZABETH S. THAYER

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN
PSYCHOLOGY

UNIVERSITY OF RHODE ISLAND
1978
ABSTRACT

Discrepancies in the age of emergence of transitivity have been found by various investigators. Piagetians argue that transitivity emerges at approximately age seven to eight years while other researchers claim that children as young as four to five years have transitivity skills. It can be argued that these differing conclusions about the age of emergence of transitivity are possibly due to differences in methodological variables. The major objectives of the present investigation were to determine whether differences in the age of emergence of transitivity are due to methodological differences in: 1) type of training (active training vs. passive training), b) order of presentation of training pairs during training (random vs. ordered), and/or c) type of response criterion demanded (judgment vs. explanation).

Subjects age 4-5 to 5-0, 6-0 to 6-7, and 7-6 to 8-0 were administered one of four premise pair training tasks: passive ordered pair, passive random block, active ordered pair, or active random block. Subsequent to the training phase subjects were tested on inference pairs using one of two types of response criteria: judgment only or judgment plus explanation.

Significant age and type of training effects were found for both the judgment and explanation criteria on inference pair errors. No significant order of training pairs effect was found for either criterion. These results lend partial support to the major hypotheses. It can thus be concluded
that methodological differences among studies on transitivity may often affect the age at which this skill emerges. Theoretical and research implications are discussed.
ACKNOWLEDGEMENTS

At this, the culmination of my graduate career, I am pleased to have the opportunity to remember all those persons who have shared my goal. I would first like to offer my appreciation to each member of my doctoral committee: Dr. Janet Kulberg, my major professor, Dr. Stewart Cohen, Dr. Charles Collyer, and Dr. William Vosburgh for their support, encouragement and comments and criticisms on all aspects of this dissertation.

In particular, I would like to recognize the contribution of Dr. Kulberg. Her many hours of counsel and thoughtful discussions of this study and manuscript and many other topics in psychology have been invaluable to my professional development.

The selection of an area and a topic for a dissertation that is both interesting and an original contribution to the psychological literature is a difficult process. I would like to specifically thank Dr. Collyer for originally stimulating my interest in transitivity and for sharing his ideas on both the methodology for this project and our review article from which this dissertation was born.

I have always felt that my undergraduate years in the Psychology Department at Goucher College provided me with an indispensable foundation in psychology. Dr. Ruth Wylie and Dr. Barbara Long offered me much needed advice and support in my decisions to major in psychology and to pursue a graduate education.
I would like to express my appreciation to each of the following people for their help: William Zwick, a fellow graduate student who provided hours of assistance and direction in the analyses of the data for this study, Marlene Schillinger, also a fellow graduate student who assisted me in collecting the data for this project, Dr. Christine Riley, who contributed thoughtful suggestions in the initial development of these research ideas, Linda Clark, whose careful attention to detail made this manuscript a typewritten reality, and my colleagues at Wellesley College for their kind words and endless patience during this past year.

I am gratefully indebted to the administrators, principals, teachers, parents, and students who, each in his/her own way participated in this study. I would particularly like to acknowledge the following persons: the teachers at St. Paul's Nursery School in Wickford, Rhode Island, the teachers at the Little Rest Nursery School in Kingston, Rhode Island, and Dr. Brian Walsh and Ms. Rosemarie Kavanaugh of the Lincoln, Rhode Island School Department.

To two very special friends, Max Alovissetti and Valerie Brewer, I extend my deep and sincere appreciation. Their emotional support, hours of empathic listening, and technical assistance helped to make this experience a personally, as well as intellectually, rewarding one.

My parents have always shared my aspirations towards a higher education and a professional career. For their love, guidance, and devotion I am eternally grateful.
Finally, I would like to thank my husband, Jack. Many times during the past year I have found comfort and strength from his love and understanding. We have always believed in each other's individual goals and it is to Jack that this work is dedicated.
# TABLE OF CONTENTS

| Chapter | Page |
|---------|------|
| Abstract | 11-11 |
| Acknowledgements | iv-vi |
| List of Tables | ix |
| List of Figures | x |
| I LITERATURE REVIEW | 1-26 |
| Introduction | 1-3 |
| Qualitative vs. Quantitative Differences | 4-5 |
| Problems in Cognitive Developmental Diagnosis | 6-8 |
| Methodological Variables | 8-23 |
| Choice of Task | 9-15 |
| Response Required | 15-20 |
| Training | 20-23 |
| Feedback | 23 |
| Conclusions | 23-24 |
| The Present Study | 24-25 |
| Hypotheses | 25-26 |
| II METHOD | 27-39 |
| Subjects | 27-28 |
| Materials | 28-29 |
| Procedure | 29-38 |
| Summary and Treatment of the Data | 38-39 |
| III RESULTS | 40-56 |
| Preliminary Analyses | 40-42 |
| Judgment Scores | 42-45 |
| Explanation Scores | 45-46 |
| Errors on Training Pairs | 46-52 |
| Categorization of Explanations | 52-54 |
# TABLE OF CONTENTS (Cont'd.)

| Chapter                  | Page       |
|--------------------------|------------|
| IV DISCUSSION            | 57-69      |
| Overall Implications and Methodological Issues | 59-67 |
| Future Directions        | 68-69      |
| Bibliography             | 70-75      |
| Appendices               | 76-92      |
| A. Raw Data              | 76-86      |
| B. Tables                | 87-89      |
| C. Explanation Categories| 90-92      |
## LIST OF TABLES

| Table | Title                                                                                         | Page |
|-------|------------------------------------------------------------------------------------------------|------|
| 1     | Summary Table of a 3X2X2 Analysis of Variance for Judgments by Age, Order of Training Pairs, and Type of Training | 43   |
| 2     | Summary of Judgment Mean Error Scores and Standard Deviations on Inference Pairs for Age of Subject With Type of Training | 43   |
| 3     | Summary Table of a 2X2X2 Analysis of Variance for Explanations by Age, Order of Training Pairs, and Type of Training | 48   |
| 4     | Summary Table of a 3X2X2 Analysis of Variance for Judgment Errors on Training Pairs During Testing by Age, Order of Training Pairs, and Type of Training | 49   |
| 5     | Summary of Mean Error Scores and Standard Deviations on Training Pair Judgments for Age of Subject With Type of Training | 51   |
| 6     | Summary of Percentage of Subjects Giving Explanations in Each Category                         | 54   |
| 7     | Summary of Mean Error Scores and Standard Deviations on Judgments for Age, Order of Training Pairs, and Type of Training | 87   |
| 8     | Summary of Mean Error Scores and Standard Deviations on Explanations for Age, Order of Training Pairs, and Type of Training | 88   |
| 9     | Summary of Mean Training Pair Error Scores and Standard Deviations During Testing for Age, Order of Training Pairs, and Type of Training | 89   |
| Figure | Description                                                                 | Page |
|--------|------------------------------------------------------------------------------|------|
| 1      | Rack and Screen Apparatus and Adjustable Sticks                              | 30   |
| 2      | Schematic Representation of Treatment Conditions                            | 39   |
| 3      | Interaction of Age and Type of Training for Judgment Errors on Inference Pairs | 44   |
| 4      | Mean Age Error Scores on Inference Pairs for Explanations                   | 47   |
| 5      | Interaction of Age and Type of Training for Judgment Errors on Training Pairs During Testing | 50   |
Chapter I

Introduction

Transitive inference is the type of reasoning in which items are related such that if the first is related to the second and the second is so related to a third then the first is so related to a third (Grove, 1967). More specifically transitive inference is the type of reasoning in which a valid conclusion of the form A>C can be derived from premises of the form A>B and B>C, where A, B, and C possess different values of some variable. One example of the transitive inference task is the following ordered question:

Jack is taller than Jeff.
Jeff is taller than Mark.
Who is taller, Jack or Mark?

Variables with a simple metric (e.g. length, weight) lend themselves easily to transitive inferences; variables whose metric is complex or partly arbitrary (e.g. measures of skill, preference, or creativity; standings in a competition) often do not.

It has long been believed by Jean Piaget and others that children are unable to form transitive inferences until they pass from the stage of preoperational thought to the stage of concrete operations at approximately age seven (Flavell, 1963; Piaget, Inhelder, & Szeminska, 1960). The transitive inference task is one of the most frequently employed barometers of the development of concrete operations. Piaget and his colleagues propose that young children (below age seven) cannot combine separate experiences in order to produce a new inferential solution. In particular, given the
basic premises that A>B and B>C, the young child is dominated by immediate perceptual input and cannot reorganize this input once new perceptual items arrive. Perceptual domination and the inability to understand the reversibility of ordered relations thus prevent the child from making inferences and lead to severe logical thinking constraints (Bryant, 1973). In his study of geometrical conceptions Piaget discusses transitivity within the context of spontaneous measurement (Piaget et al, 1960). The development of transitivity skills comes late in the growth of logical relations. Piaget believes that a child can recognize lengths of individual terms and the relations between two terms in one pair without being able to order or seriate the lengths of more than one pair of terms. The young child can judge the lengths of the individual stimuli but cannot yet work with the relations between and among the pairs of stimuli of varying lengths. Transitivity enters a child's relational structuring when reversibility and passage among three terms can be understood. At this point in development a particular term in a series (B in the given example) is no longer viewed as having a one-way relation (either smaller than or larger than) but is viewed instead as capable of having two relations simultaneously (both smaller than some things and larger than others).

The traditional Piagetian task directly relevant to transitivity of length concerns the development of measuring techniques which presuppose transitivity (Piaget et al, 1960). Children are asked to build a tower of blocks equal
in measured height to a tower already built by the experimenter. The latter tower stands on a table which is higher than the table on which the child is to build his/her tower. Sticks of various lengths are made available but the child is not told how to use them. The criterion of mastery in this task is the child's ability to use sticks of varying lengths as intermediate comparative measures of the towers. The majority of children appear to understand and master the measuring technique only after age seven or eight years. The principle developmental stages are: 1) a crude comparison without taking into account the difference in table heights, 2) attempts to bring the towers together for closer visual comparison and the child's use of his/her own body as a common measure, 3) use of body-independent measures such as a tower or stick which must be exactly the same length as the tower to be measured and 4) use of sticks longer and shorter than the model. Piaget argues that the child's successful use of measurement implies the use of transitivity in the form: if A=B and B=C then A=C.

The focus of the present study is to examine the discrepancies between Piagetian and other researchers over the age of emergence of transitivity. The study specifically investigates the following methodological differences in the research on transitive inference: 1) age of subjects, 2) type of training on premise pairs (e.g. A>B and B>C for A>C), 3) ordering of premise pairs during training, and 4) type of criterion used during testing for transitivity.
Qualitative vs. Quantitative Differences

Piaget's assumptions and findings concerning the transitive inference task have been challenged by several investigators (Braine, 1959; Bryant, 1973, 1974; Bryant & Trabasso, 1971; DeBoysson-Bardies & O'Regan, 1973; Harris & Bassett, 1975; Lutkus & Trabasso, 1974; Riley, 1975, 1976; Riley & Trabasso, 1974; Roodin & Gruen, 1970; Trabasso, 1975; Trabasso, Riley & Wilson, 1975). These authors claim that Piaget's inference task does not, in fact, test whether the young child has inferential ability or not. Their basic hypothesis is that children's differences in making transitive inferences may be more a matter of the memory processes involved than a lack of logical competence. In other words, one must ensure that the child has remembered the comparisons which have to be combined (A)B and B>C) if one is to conclude that the child is or is not actually making transitive inferences. If memory fails to be controlled an erroneous response may have little to do with inferential ability or level of logical thinking.

Piaget typically did not control for memory factors and although he has been criticized for this omission by Bryant and Trabasso, Piagetians (Flavell & Wohlwill, 1969; Smedslund, 1963) would argue that the effect of memory training is to produce only the semblance of a true operation. Piagetians would thus predict that the ability to utilize transitivity acquired through memory training would lack resistance to extinction and would fail to generalize to tasks other than those used in training. The memory training procedures
used for transitivity tasks have also been said to focus on the psychological processes by which knowledge embodied in the child's cognitive structures gets accessed and utilized in real situations, rather than on logical representations, groupings, rules, etc. which actually form these cognitive structures (Flavell & Wohwill, 1969). The basic controversy stems from different theoretical assumptions concerning the nature of cognitive development. Piaget and his colleagues conceive of cognitive development as proceeding in a discontinuous fashion with identifiable qualitatively different stages emerging during various life periods. Each stage represents a unified set of cognitive structures; development is not viewed as the continuous maturation of independent psychological entities. Developmental stage transitions are consequently not defined in terms of purely quantitative changes (Flavell, 1963; 1977). The position taken by Trabasso is precisely the opposite: "Cognitive development may be seen as continuous and qualitatively similar to the target adult-model. The growth of a child's capacity for immediate, short-term and long-term memory is likely to be gradual and quantitative" (Trabasso, 1975, p. 136-137). In his conceptualization of cognitive growth Trabasso makes a distinction between the skills possessed by the young child and the adult only on the basis of accumulation of information and the development of more efficient and adaptive methods for processing information. Adult thinking can then be conceptualized as qualitatively similar to that of the young child but having a larger fund of knowledge, more refined and better
practiced skills, and thus a greater ability to handle cogni-
tively complex problems.

Problems in Cognitive Developmental Diagnosis

Some of the issues addressed in the recent literature on transitive inference offer excellent examples of general problems in cognitive developmental diagnosis (Flavell, 1977; Smedslund, 1953; 1963). Flavell (1977) proposes that there are two sets of questions we may ask about any cognitive acquisition such as transitive inference. Conceptualization questions ask about the cognitive processes which actually comprise acts labeled transitive inference. These kinds of questions are concerned with the kinds of strategies utilized by individuals as they solve any transitive inference type tasks. Practical assessment questions, on the other hand, ask what procedures should be used to test, on any cognitive process, for the presence or absence of transitive inference in a child. These questions focus on the methodology and analysis used in the design of a task of transitive inference.

For the purposes of this dissertation we will focus only on the second question, that of practical assessment. If we initially make the assumption that a child either does or does not possess the ability to make a successful transitive inference then two possible errors could result from an examination of his/her performance on a specific task. If the child actually does have the logical ability to make a transitive inference and yet shows no evidence of this ability on a task we will make a false-negative (Type II)
diagnostic error, whereas if the child does not possess
transitive inference skills and yet somehow makes the
correct "symptom" response we will then make a false-
positive (Type I) diagnostic error (Flavell, 1977). Both
Flavell (1977) and Smedslund (1963) offer a number of poten-
tial reasons for committing Type I and Type II errors.
False negative errors could result from the following sources:
1) the child might fail to understand the words in the initial
instructions, 2) the child may not initially perceive that A
is longer than B and B is longer than C and 3) the child may
not remember the two initial premises (A)B, B)C) by the time
he/she is required to make the inference. The latter is pre-
cisely the area of inquiry pursued by Bryant and Trabasso
(1971). In any of these three cases, the tester will misin-
terpret the absence of the symptom response as the absence of
transitivity. The possible causes of false positive errors
are: 1) the child might just guess, 2) the child might
directly perceive that A was longer than C without using
transitivity, or 3) the child might use another (nontransi-
tive inference) solution strategy. One such strategy (Smed-
slund, 1963) is that the child might code the initial premises
(A)B, B)C) as "A is long but B is not long, and B is long but
C is not long." Since A has been coded as being "long" and
C "not long" the child will choose A when asked to solve the
AC inference relationship. In any of these cases, the tester
will misinterpret the presence of the symptom response as
the presence of "true" transitivity.
Much of the literature on transitive inference has focused on attempting to control for some of these causes of false-negative and false-positive errors. Let us now examine in more depth specific research on the variables affecting the acquisition and/or utilization of transitivity.

**Methodological Variables**

The major controversy in the literature on transitive inference stems from the apparently discrepant findings concerning the age of emergence of transitivity. One group of researchers claims that very young children (age four and five years) are capable of making correct transitive inferences (Braine, 1959; Brainerd, 1973; Bryant, 1973, 1974; Bryant & Kopytynska, 1976; Bryant & Trabasso, 1971; DeBoysson-Bardies & O'Regan, 1973; Harris & Bassett, 1975; Lutkus & Trabasso, 1974; Riley, 1975, 1976; Riley & Trabasso, 1974; Roodin & Gruen, 1970; Trabasso, 1975; Trabasso, Riley, & Wilson, 1975) while another group of researchers supports the Piagetian claim that the age of emergence is around seven to eight years (McManis, 1969; Murray & Youniss, 1968; Smedslund, 1960, 1963; Youniss & Furth, 1973; Youniss & Murray, 1970). The major issues in the studies which examine the transitivity controversy can be most clearly reviewed with respect to four procedural choice-points: 1) choice of task, 2) response required, 3) initial training, and 4) method of feedback.
Choice of Task

The tasks used in research designs for the investigation of transitivity generally differ in five major ways. (1) Some authors (Braine, 1959; Bryant & Trabasso, 1971) have sought to control for false positive diagnostic errors by varying the ways in which the stimuli in stick-length comparisons are presented to the subject. These stimulus variations include adjustments in the length of the sticks used, the number of stimulus sticks presented, the distance between stimulus and/or test sticks, and the use of a Mueller-Lyer illusion effect to help eliminate purely perceptual judgments. (2) Riley (1975, 1976) has extended some of Trabasso's innovations to comparative dimensions other than length, such as "happiness" and "niceness." (3) Another group of researchers (Harris & Bassett, 1975; Youniss & Murray, 1970) have increased the range of stimuli and test comparisons used to include relations such as "equal to" rather than the traditionally used "greater than" or "less than." (4) Most of the literature on transitive inference tasks is based upon nonverbal stimuli presentations but two studies (Harris & Bassett, 1975; Stetson, 1974) have tested for transitivity using an orally presented task. In these studies children were read the premise relations in sentence form and were asked to give a verbal response. (5) Lastly, task presentations have differed along the dimension of active discovery vs. passive presentation of the initial premise conditions (Bryant & Kopytynska, 1976; Bryant & Trabasso, 1971; Piaget et al, 1960). For the purposes of this
dissertation only the last and part of the first of the above differences will be addressed and reviewed.

In one series of experiments (Bryant, 1973, 1974; Bryant & Trabasso, 1971; DeBoysson-Bardies & O'Regan, 1973; Harris & Bassett, 1975; Lutkus & Trabasso, 1974; Riley & Trabasso, 1974; Trabasso, Riley, & Wilson, 1975) a new kind of experimental task presentation was utilized in order to prevent the transfer of absolute responses of verbal labels leading to a false-positive diagnosis on the basis of a nontransitive hypothesis. To control for the mere parroting of the verbal labels A='longer' and C='shorter' Bryant and Trabasso and their colleagues introduced more stimuli and thus more direct comparisons. The initial four direct comparisons involved five sticks of different lengths and colors (A>B, B>C, C>D, D>E). The rods were always presented in pairs and in such a way that they all appeared one inch from the top of a container. This was accomplished by boring holes of varying length in the box. This forced the child to use the color differences and not the absolute lengths in making a choice. Each child was trained on the initial comparisons and tested a number of times on every one of ten possible pairs of colored sticks. The critical comparison was B>D in which both rods had been larger and smaller in the initial comparisons. Bryant and Trabasso (1971) hypothesized that the probability of making a correct inference on the B>D test is the probability of jointly recalling the information for each of the initial training pairs BC and CD. The data were consistent with
the original hypothesis. In summary, the five stick array box attempts to prevent children from making direct visual comparisons. This task thus has promise as a model for controlling the possibility of false-positive diagnostic errors due to perceptual discrimination and/or non-transitive hypotheses.

This type of task presentation has been extended by Riley (1975, 1976) in a novel study which examined children's abilities to map four different comparative dimensions onto linearly ordered referents. Kindergarten and third grade children were shown a row of six faces on individual cards and were instructed that these faces represented a group of children differing in height, weight, happiness, and niceness. Order was indicated by describing the relations between adjacent faces (e.g. Mike is taller than Steve). Each child was tested after training on all possible comparative relationships.

Riley's (1975, 1976) task was designed to test the hypothesis that individuals tend to integrate premise relations such as A>B, B>C, C>D into a single linear ordering, ABCD. This task does examine the generalizability of the linear order representation to new situations other than length (happiness, sadness, etc.) with fully ordered premise information, but it does not examine the processing used when an individual is asked to make a transitive inference from separate premise relations. In a second experiment, Riley (1975, 1976) therefore designed a display where the premise relationships were always directly observable, but
spatially separated. The child thus had to isolate and coordinate the premise relations. Riley hypothesized that this display would force the child to respond deductively to inferential questions. The comparative heavier was always used for the array of faces and the cards were displayed in three conditions. In the linear display the cards were ordered 1, 2, 3, 4, 5, 6. In the vertical display the five premise pairs were presented in a pairwise ascending or descending vertical order (e.g. (1,2), (2,3), (3,4), (4,5), (5,6)). In this way the resulting display formed two vertical columns running in ascending or descending order from card one to card five and card two to card six. In the random display the five premise pairs were presented in a random vertical arrangement (e.g. (5,6), (2,3), (1,2), (3,4), (4,5)). The random condition was designed to hypothetically prevent a child from utilizing an ordered spatial stimulus arrangement to facilitate the construction of a cognitive spatial representation. After presentation of the premise relationships all children were tested under both a Display Test, where the child was told to find the answer while looking at the arrangement of faces, and a Memory Test where no external display cues were provided. The hypothesis that the random condition would interfere with the use of a linear order problem solving strategy was partially supported by the decision time data in Riley’s (1976) work.

The second distinction to make in the task presentation category is that of "active" vs. "passive" inference problems. Bryant and Trabasso's traditional task (Bryant & Trabasso,
1971; Lutkus & Trabasso, 1974; Riley & Trabasso, 1974; Trabasso, 1975; Trabasso, Riley, & Wilson, 1975) is a passive one wherein the child is told A>B, B>C, etc. On the other hand, the Piagetian tower task is an active one which essentially differs from the passive task of the Trabasso group in three major ways. (1) Subjects must discover the initial premise relations for themselves in an active task. No attempt is made by the experimenter to directly ask the child to compare A with B or B with C. The child must figure out for him or herself how to acquire the needed premise information. Children who do not measure will not obtain this information. It is possible that this difference makes the task a more difficult one and thus hinders the expression of transitivity. (2) Subjects are allowed to concretely manipulate the materials in the task. They have the opportunity to interact with the initial premise stimuli. They are even required to actively construct one of the initial premise conditions (tower two or C). The effect of this difference is less clear. If it were not always paired with the above active discovery of premise differences, it may, in fact, make the task an easier one and thus facilitate the expression of transitivity. This would be in keeping with the Piagetian emphasis on active involvement in the learning process especially in the pre-operational and concrete operational stages. (3) The primary indicator of the presence of transitivity in Piaget's active tower measuring task is the child's active ability to use the available measuring sticks as mediators. In Piaget's
terms a mediator, in this problem, is a stimulus which can be conceptualized as both larger than one stimulus and, at the same time, smaller than a third stimulus. The mediator thus allows the subject to determine the relationship between the first and third stimuli. In other words, B must be conceptualized as a mediator in any A>B, B>C, A>C problem.

Bryant and Kopytynska (1976) argue that young children did not spontaneously use the measuring sticks in Piaget's active tower task because they did not realize that a direct comparison of the two towers by eye might be unreliable. These authors thus designed a measuring task where the two quantities to be compared were invisible (holes drilled in wooden boxes). One rod, marked off in different colors, was provided as the only measuring stick. In support of Piaget's findings, all children age five and six failed a version of the tower task and yet, many were successful on the hole measuring task. Bryant and Kopytynska concluded that there is some evidence that young children can spontaneously use intervening measures which implies the use of transitivity skills. It can be argued that the replication of the Piagetian findings and the new Bryant and Kopytynska findings offer further support for the idea that transitivity and/or measurement is not spontaneously used when the task construction allows for other possible solutions ("eyeballing it" or using self as measure). The Piagetian tower measuring task requires the child to actively construct the C (second tower) premise condition and does not restrict the child's measurement stimulus (B) to only one choice. Although the
Bryant and Kopytynska work does begin to examine the question of active spontaneous measurement it does not directly and actively test children's transitivity skills. A comparison of active vs. passive conditions is still necessary to explore the Piagetian conception of learning and development as a process of active discovery rather than of "passive" training. Bryant (1973) reports some unpublished data which appears to be the only direct attempt to address this question. He reports that he recently repeated Piaget's experiment with five year olds and found that "All children had previously done very well in a passive inference task. Yet they were plainly at a loss in our active test" (p. 423). Further examination of this question seems essential to the literature in this area.

Response Required

Investigators diverge again with respect to the type of response from which transitive inference ability, or the lack of it, is to be inferred. The major controversy is between using the child's judgment by itself (Braine, 1959; Bryant, 1973, 1974; Bryant & Trabasso, 1971; Lutkus & Trabasso, 1974; Murray & Youniss, 1968; Riley, 1975, 1976; Riley & Trabasso, 1974; Trabasso, 1975; Trabasso, Riley, & Wilson, 1975; Youniss & Murray, 1970) or using both the judgment and the child's explanation of his/her judgment (McManis, 1969; Smedslund, 1960, 1963). Investigators who require supplementary explanations thus apply a more stringent criterion.
The most lively exchange in this area has been between Braine and Smedslund. Braine (1959) used a nonverbal technique where subjects were instructed to find a piece of candy always hidden under either the longer or shorter of two pieces of wood. Braine's task and conclusions that transitivity and ordinality were acquired between ages five and six were severely criticized by Smedslund (1963) for failure to control for Type I error due to solutions by irrelevant non-transitive hypotheses. Smedslund presented replication data of his own in support of Piaget's norms and Braine (1964) replied that Smedslund's criticisms were far fetched since there was no evidence for the occurrence of non-transitive solutions in his work. Braine (1964) presented new data which failed to reveal non-transitive solutions and replicated his original findings. Smedslund (1965) responded to Braine's reply and added the criticism of "pseudomeasurement" to his previous charges to Braine. Pseudomeasurement refers to tasks where only one of two sticks to be compared is measured. Smedslund asserts that this type of condition does not control for the possibility of non-transitive reasoning of the type A>B therefore A>C or A<B therefore A<C. Smedslund hypothesized instead, that children may apply transitive and non-transitive strategies differentially on measuring trials. He argues that young children who succeed on genuine measurement trials (trials where both sticks are measured) apply non-transitive hypotheses in the pseudomeasurement trials both when the previously measured stick is correct and when it is incorrect, while children who fail in the genuine measurement
trials do not apply non-transitive hypotheses when the previously measured stick is incorrect. Smedslund (1965) reanalyzed Braine's (1964) data and found that subjects who passed the genuine measurement items performed better on pseudomeasurement trials than did subjects who failed on the measurement items.

Brainerd (1973) argues that the discrepancy between Braine's and Smedslund's data might be due, in part, to the different response criteria employed by each for determining the presence of transitivity. Braine's criteria were judgments only while Smedslund's were judgments plus adequate explanations. Brainerd goes on to state "...the basic proof for the presence of a given cognitive structure must, by definition, be a proof that an act of understanding appropriate to that level of structuration has taken place. By implication, if we wish to avoid being unduly conservative, then we require the minimum necessary evidence that an act of understanding has occurred" (p. 177). Explanations provide information about the nature of the underlying structures, but they may confound the diagnosis of the presence of a structure when utilized as part of the criterion measure. For example, a child may be able to make the correct judgment on a transitive inference task but would be unable to verbalize an adequate explanation for the choice. The requirement of a verbal explanation as a necessary condition for diagnosis of correct transitive reasoning would exclude children such as the one above and might thus result in a false-negative error. Much of the discrepancy between age related findings
in the transitivity literature can be accounted for by this major difference in response required. It seems possible that younger children may, in fact, have the transitivity principle and utilize it successfully but may still lack the verbal skills to be able to offer appropriate transitive inference explanations for their response choices. Smedslund (1963), Flavell and Wohlwill (1969), and others would argue conservatively that verbal explanations are necessary components indicative of the true presence of a cognitive structure. Unless this criterion is employed we have difficulty knowing whether or not the subject utilized a non-transitive solution.

Reese and Schack (1974) also oppose Brainerd's (1973) conclusion that children's judgments provide a better criterion than do explanations for inferences about the presence or absence of cognitive structures. They argue that explanations are the best control for Type I error due to non-transitive hypotheses and that this source of error cannot be eliminated by stimulus refinements as Brainerd suggested. Brainerd (1974) strongly answers Reese and Schack (1974) by asserting that their criticism "also may be faulted on the ground that it is predicated on a source of error (irrelevant hypotheses) which is not even known to exist" (p. 71). This heated debate was finally settled by Brainerd (1977) in a recent publication. In this work Brainerd psychometrically analyzed the criterion question with respect to the sequential development of Piagetian concepts in children. Through a series of mathematical
proofs, Brainerd concluded that the appropriate response criterion is the one with the lowest error rate, regardless of whether the actual errors being committed are false positives or false negatives. Brainerd also concludes that the currently available literature shows that error rates for judgments only are lower than corresponding error rates for judgments plus explanations.

Roodin and Gruen (1970) presented the only work which specifically investigated the judgment vs. judgment plus explanation response criteria for transitivity. Groups of kindergarten, first grade, and second grade children were given a task similar to that of Smedslund (1963) and employing Mueller-Lyer illusions to correct for the possibility of a perceptually dominated solution. One-half of the children at each grade level were allowed to use a memory aid for comparisons of $A \rightarrow B$ and $B \rightarrow C$ while the other half were not. Almost all subjects who could verbally explain their judgments also made correct judgments although the converse was not supported. Age differences were found to be significant for both criteria only in the no memory condition. In the memory aid condition age differences were only significant when a judgment plus explanation criterion was employed. Also there was found to be a significant difference in the number of transitive responses depending on the criterion. Subjects provided a significantly greater number of correct judgments than judgments plus explanations. Further examination of the effects of different response criteria needs to be conducted using other types of task presentations such as those used
by Trabasso and his colleagues.

**Training**

The primary hypothesis advanced by Bryant and Trabasso and their colleagues is that children's difficulty in making transitive inferences may be more a matter of memory limitations than of logical competence level (Bryant, 1973, 1974; Bryant & Trabasso, 1971; Lutkus & Trabasso, 1974; Riley & Trabasso, 1974; Roodin & Gruen, 1970; Trabasso, 1975; Trabasso, Riley, & Wilson, 1975). Their standard paradigm for controlling the forgetting of the initial premises is the following: 1) the subject is trained on one initial pair comparison at a time in a choice discrimination task where comparative questions are asked; 2) when one pair is mastered (8 out of 10 successful choices) the next one is learned; 3) when the four pairs are trained to criterion the second training phase begins. All four pairs are now presented in a blocked, randomized order. This continues until the child responds correctly to six successive presentations of each pair; and 4) testing follows immediately after completion of the two training phases.

In all studies where this procedure is used for ensuring retention of the premises, the authors consistently find that children two or more years younger than Piaget's subjects successfully make transitive inferences, even on the critical BD task. This was also found to be true for retarded subjects whose mental age was five to seven years (Lutkus & Trabasso, 1974). Compared to intellectually average subjects of similar MA level, retardates took
longer to learn the initial comparisons in the training phase, but their performance on transitivity tests was above chance, with the majority of the subjects perfect, and their overall performance slightly below their intellectually average control subjects. Bryant and Trabasso consistently found that memory for initial comparisons is highly correlated with performance on transitivity tasks (Bryant, 1973, 1974; Bryant & Trabasso, 1971; Lutkus & Trabasso, 1974; Riley & Trabasso, 1974, Trabasso, 1975). This is especially true when the critical premises for the BD comparison (B>C, C>D) are examined. Although Smedslund (1963) did try to control for this variable in his work by asking the subject to restate the premises immediately before the test question, Bryant and Trabasso continued to argue that the children did not have adequate training to be certain of retention of the initial training comparisons.

Riley and Trabasso (1974) sought to control for two more possible sources of error in the training phase of the Bryant and Trabasso transitivity task. Although the size and position cues were controlled in the original work of Bryant and Trabasso (1971), the sticks were arranged in a display box so that location relative to the ends and distance between members of pairs were correlated with a spatial array. Riley and Trabasso removed such cues by presenting the sticks next to each other for all pairs in both testing and training by asking either "Which is shorter, A or B?, etc." or "Which is longer A or B?, etc." or both, in three different experiments. When both comparatives were used, subjects (age four
to five years) learned adjacent pairs faster and more often reached criterion than when only one comparative term was used in training. In testing, the children were more often successful when double comparative relations were used within pairs during training.

Riley (1976) complemented this finding in a study of three types of comparative dimensions. She varied the comparative questions in both the training and testing phases by asking either the comparative (e.g. "Which girl is taller?") or negative equative question (e.g. "Which girl is not as tall as the other?"). It was found that decision times to negative equative comparisons were slower than to comparatives but no difference in error rates to these two question types was observed. These findings support Piaget's (Piaget et al, 1960) claim that young children reduce the comparative to a label such as A is long, B is not long. As one of the Riley and Trabasso subjects reported: "You have two sizes of sticks, long and short, and you keep changing which ones are which" (1974, p. 197).

Trabasso and his colleagues have continually trained their subjects on training pairs presented first in an ordered pair sequence followed by pairs presented in a randomized block design. It thus remains to be determined whether or not both types of orderings of training pairs are needed to ensure retention of the premise information for subjects of different age levels. Examination of the effects of each type of training pair order would also provide a measure of one of the methodological differences between Piaget's active tower task
and Trabasso's passive task - the child's ability to recognize (B) as a mediator in any two pair comparison (A>B, B>C). A more detailed discussion of this distinction is presented at a later point in this chapter.

Feedback

Part of Bryant and Trabasso's (1971) initial study used visual feedback at the end of each trial during training phases. It was thus possible that a false-positive error could have resulted if the children solved the BD transitive question noninferentially by simply remembering that B was six inches and D four inches. The authors controlled for this possibility in their second experiment (Bryant & Trabasso, 1971) by eliminating visual feedback in the training session and using only verbal feedback instead. The full lengths of the rods were never displayed completely until the whole experiment was finished. Only one inch of each rod was displayed during training. Children learned the stick length comparisons by remembering the verbal feedback given by the experimenter; "Yes, the (red) stick is longer (shorter) than the (blue) stick" for a correct response or "No, the (red) stick is longer (shorter) than the (blue) stick" for an incorrect response. It has generally been found that young children show a relatively poorer performance in those conditions where feedback is verbal (Trabasso, 1975). Yet, their performance is still above chance level and qualitatively comparable to that of their older peers.

Conclusions

The practical assessment of cognitive abilities is not a simple matter. In the absence of complete controls for all
possible sources of Type I and Type II errors, diagnostic decision-making may be tentative and fallible. Smedslund aptly states "The relationship between any set of behavioral indices and a mental process, therefore, is an uncertain one, and a diagnosis will always have the status of a working hypothesis" (1969, p. 247).

The practical assessment of transitivity is also not a simple matter. The controversy over the age of emergence of this cognitive skill is clouded with methodological difficulties and variations. The researchers theoretical frame of reference seems to determine his/her choice of assessment techniques. A controlled examination of the effects of methodological differences in assessment variables such as training and response criteria would help provide some of the much needed clarification in this area of inquiry.

The Present Study

After a thorough review of the literature in this area one could conclude that differing conclusions reached about the age of emergence of transitivity by previous investigators are possibly due to differences in methodological variables; especially type of training and/or type of response criteria. Yet, no one study has specifically manipulated these dimensions simultaneously. It is the purpose of the present study to experimentally examine the effects of training and differing response criteria on the age of emergence.

Once memory factors are controlled through adequate training on premise pairs, expansion of the training variable
can be a useful way in which to examine the active/passive distinctions essential to the differences between the Piagetian and Trabasso tasks. This study examines two of these differences, manipulation of materials and use of (B) as a mediator. Discovery of initial premise relations is not addressed in this work. A unique active training condition addresses the question of whether or not interaction with stimulus materials will affect transitivity.

Separation of Trabasso's ordered pairs and randomized block training phases into two distinct training conditions examines the subject's ability to recognize the middle term (B) as a mediator in any two pair comparison (A>B, B>C). In ordered pairs training, the middle term is used in each successive training comparison (B in A>B, B>C) while in randomized block training each training pair is not necessarily followed by the next training pair in the A>B>C>D>E sequence containing a similar member (B>C followed by D>E). Randomized block training thus requires the child to determine for him/herself the mediating stick in each two pair sequence.

Trabasso and his colleagues have always employed a judgment only criterion in their work while the Piagetians have always employed a judgment plus explanation criterion in theirs. This study evaluates the effects of both a judgment and a judgment plus explanation criterion on a version of the five stick Trabasso transitivity task.

Hypotheses

The major hypothesis of this study was that differing
conclusions reached about the age of emergence of transitiv-
ity by previous investigators are due to differences in
types of training order of presentation of training pairs,
and/or type of response criteria demanded. Specifically the
following predictions were proposed:
1. Transitivity will significantly increase with age for both
the judgment and judgment plus explanation criterion.
2. The judgment plus explanation response criterion will be
more difficult for younger children than for older children.
3. Children who receive the active training conditions will
evidence significantly more transitivity than children
who receive the passive training conditions.
4. Children who receive randomized block training will
evidence significantly less transitivity than children
who receive ordered pairs training.
5. The effect of the more difficult randomized block
training will be more pronounced for younger children
than for older children.
6. The facilitative effect of the active training condition
will be more pronounced for younger children than for
older children.
Chapter II

METHOD

Subjects

Subjects were 120 Caucasian children including 60 boys and 60 girls. Subjects ranged in age from 4-5 to 8-0 with 40 subjects at each of three age levels; 4-5 to 5-0 (median age 4-8), 6-0 to 6-7 (median age 6-2), and 7-6 to 8-0 (median age 7-8). These age levels were selected so that they would span the controversial age of emergence phenomenon.

The children who participated in this study were selected from three rural, middle-class communities in Rhode Island (Towns A, B, and C). The decision to select children from middle-class communities is methodologically consistent with the intent of the investigator to replicate the subject populations used in the literature on transitivity as closely as possible.

The children were selected according to their date of birth; and parental permission was required for participation in this study. Subjects were then randomly selected for participation in the study from the entire group of subjects in each age level who returned parental permission forms. All children age 4-5 to 5-0 lived in Towns A and B and attended one of two privately operated nursery schools located within the communities. The children age 6-0 to 6-7 and 7-6 to 8-0 lived in Town C and attended first or second grade in one of three public elementary schools located in that community.
An attempt was made to select children from communities with similar socioeconomic levels. Evaluation of the socioeconomic level of these communities was based upon the mean level of income of the residents in the community. The State Department of Rhode Island reports that the mean income in 1970 (latest available census data) was $10,216 for Town A, $11,773 for Town B, and $11,957 for Town C. The State Department also reported the following projected median levels of income for 1976; Town A $14,600, Town B $15,400, and Town C $15,600.

All subjects were randomly assigned by sex to one of the four training conditions used in this study. One-fourth of the subjects at each age level received Active Ordered Pairs Training, one-fourth received Active Randomized Block Training, one-fourth received Passive Ordered Pairs Training, and one-fourth received Passive Randomized Block Training. Equal numbers of boys and girls received each training condition. One male age 4-8 and one female age 6-2 were reluctant to participate in the study and evidenced a lack of understanding of the general task instructions. These subjects were thus eliminated from the study and were replaced by randomly selected same-sex children from the appropriate age level subject pool.

Materials

Five metal standard teaching pointers of adjustable length (from 5 inches (12.7 cm.) to 25 inches (63.5 cm.)) were used. The sticks were color coded in red, blue, green, yellow, and brown masking tape. The colored tape was
wrapped around the top two inches (5.2 cm.) of the pointers to a depth of one-eighth inch (.3 cm.). A screened wooden rack was used in some training and in all testing conditions for presentation of sticks. The rack was constructed so that two sticks protruded 2 inches (5.1 cm.) from the top; the full lengths of the sticks were hidden from view by an opaque black felt screen covering the front of the rack. Two one-half inch (1.3 cm.) metal picture hooks were attached flush to the top of the rack. Each hook was 3 inches (7.6 cm.) apart from the other. The sticks were hung by resting them in the hooks so that only the colored taped portion showed above the rack.

The rack stood 14 inches (35.6 cm.) by 8 inches (20.3 cm.) by 1 1/2 inches (3.8 cm.). The black felt screen was also 14 inches (35.6 cm.) by 8 inches (20.3 cm.). The rack was made from two pieces of wood 14 inches (35.6 cm.) by 1 1/2 inches (3.2 cm.) by 1 1/2 inches (3.2 cm.) and one piece of wood 8 inches (20.3 cm.) by 1 1/2 inches (3.2 cm.) by 1 1/2 inches (3.2 cm.). The rack was nailed to a wooden base 12 inches (30.5 cm.) by 3 1/2 inches (8.9 cm.) by 3/4 inches (1.9 cm.). The apparatus used in this study is pictured in Figure 1.

Procedure

Subjects were trained and tested individually in one 20 minute to 45 minute session. All subjects were initially pretested for knowledge of color names and comprehension of comparative length terms "longer" and "shorter." Subjects were asked to name the color of each of five pieces of tape
FIGURE 1. Rack and Screen Apparatus and Adjustable Sticks
pasted to a blank white card. They were then shown four pieces of yellow chalk of different lengths. The subject was asked to select the longer piece in two pairs and the shorter piece from two other pairs. Subjects were then shown the rack and screen apparatus and it was explained how sticks of different lengths can appear the same when placed appropriately on the rack. The first pair of sticks in the subject's designated color order was always used for the demonstration of the apparatus. The instructions given to each subject were:

Today we are going to play some games with sticks of different lengths and different colors. See how these sticks can be made longer and shorter. The experimenter and subject manipulate one of the five sticks. The subject is given the other four sticks and the experimenter continues saying 'Now here are the other four sticks. I want you to make sure that these all work in the same way as the other one.' The subject then manipulates the other four sticks. The experimenter continues 'Now, if I put these sticks onto this rack you can only see the colored top of the sticks. The rest of the stick is hidden behind this piece of black cloth.' The experimenter demonstrates the rack and screen apparatus with the first color order pair. 'You see - these two sticks look about the same length from the front but if I turn the rack around then you can see that one stick is longer than the other.' The experimenter turns rack around and lets the subject see the sticks of different lengths on the back side. 'Now I'm going to show you two sticks at a time and ask you to (Active conditions - 'make one longer than the other one') (Passive conditions - 'tell me which one is longer or which one is shorter'). The important thing is to try to remember which sticks are longer than other sticks and which sticks are shorter than other sticks. I'm going to give you lots of practice with the sticks and try very hard to remember which is longer and which stick is shorter in every two sticks I show you. After you practice for awhile then we'll try to see how much you remember. So be sure you try to remember which sticks are longer than others and which sticks are shorter than others. Now tell me what you are supposed to try to remember while you practice?' After the subject successfully repeated the task directions the training phase began.
All subjects were then trained on the four basic premise pairs (AB, BC, CD, DE).

**Training**

The passive training levels differed from the active training levels only in that the subjects were not requested to manipulate the lengths of the five color coded sticks themselves. The decision here was to define Active vs. Passive training conditions only in terms of the concrete manipulation of the lengths of the initial premise pairs. In this way active has been defined in its simplest form. The ordering of the five color sticks by length was predetermined and held constant by the experimenter in the passive training conditions. The subjects did not actively establish this order for themselves.

It was also decided to manipulate the training levels by comparing training on adjacent pairs of sticks presented in order (AB, BC, CD, DE) with training on pairs of adjacent sticks presented in a randomized design (e.g. BC, DE, AB, CD, or DE, BC, CD, AB, etc.). All pairs were always presented to the subject in blocks of four pairs. This type of manipulation of the training variable varied the pattern or order in which the subjects were required to learn the adjacent stick pairs. Randomized block training also required the subject to determine for him/herself the mediating stick in each two pair sequence while in ordered pair training the mediating stick was presented in both the "less than" and "greater than" positions in each successive two pair training comparison (e.g. B in A>B followed by B>C).
Two different randomly selected color orders were used during both the active and passive task presentations (Green>Brown>Blue>Red>Yellow and Yellow>Brown>Green>Red>Blue). The subjects in each training condition were randomly assigned to one of the different color orders so that five boys and five girls at each age level received each color order presentation.

**Active Ordered Pairs Training** - Adjacent stick pairs (AB, BC, CD, DE) were trained in ordered blocks of four starting with the longest pair (AB) for one-half of the subjects and the shortest pair (DE) for the other half. The order of start was counterbalanced over sex.

The subject was first shown how a color coded stick could be adjusted to different lengths. The subject was then presented with two 5 inch (12.7 cm.) sticks (A and B or D and E) and was asked to "Make the (green) stick look a little longer than the (brown) stick." After the child correctly adjusted the sticks he/she was asked "Which stick is longer (shorter)." Following the child's response both sticks (A and B or D and E) were removed and the same procedure was repeated with the next pairs in the ordered sequence (BC or CD, followed by CD or BC, followed by DE or AB). All subjects were given four adjustments of each pair. It was impossible to have an incorrect trial in the Active conditions due to the nature of the task. The training criterion level for the Active conditions was determined by reviewing the Riley and Trabasso (1974) trial of last error.
data for ordered pair and randomized block training for their subjects age 4-1 to 5-0.

The placement of the sticks in each training pair was random for left and right positions but equal in occurrence over trials. The comparatives used for the questions and verbal feedback ("longer" or "shorter") were random but occurred equally often over trials and left-right rack positions. All feedback during the session was both visual and verbal. While the correctly adjusted stick pair was still visible to the subject the experimenter said "Yes, the (green) stick is always a little longer (shorter) than the (brown) stick."

**Active Randomized Block Training** - The four comparisons of adjacent stick pairs (AB, BC, CD, DE) were trained in a randomized block order. These pairs were presented in blocks of four pairs with pairs randomized within each block. Each stick pair was presented in the same way as those presented in active ordered pairs training. Placement of sticks was randomized in the same way as in active ordered pairs training. The same comparative questions used in active ordered pairs training were used in this condition with questions occurring randomly but equally often over trials and left-right rack positions. All feedback was both visual and verbal and of the same form as that used in active ordered pairs training. All subjects were given four adjustments of each pair.
Passive Ordered Pairs Training - Four comparisons of adjacent stick pairs (AB, BC, CD, DE) were trained in ordered blocks of four starting with the longest pair (AB) for one-half of the subjects and the shortest pair (DE) for the other half. Order of start was counterbalanced over sex. The placement of the sticks in each training pair was random for the left and right rack positions but equal in occurrence over trials.

The five sticks were first adjusted in length by the experimenter to form the correct color A>B>C>D>E sequence. These length adjustments were made behind the rack and were not visible to the subject. Upon presentation of a stick pair in the rack the subject was asked "Which stick is longer?" or "Which stick is shorter?" Questions were random but equal in occurrence over trials and over left-right rack positions. The subject's response was a judgment indicating the color of the correct stick choice. After the subject made a color judgment the rack was turned around to provide visual feedback on each trial. Verbal feedback (with the rack turned around) during the session was of the form "Right, the (green) stick is longer (shorter) than the (brown) stick" for correct responses or "No, the (green) stick is longer (shorter) than the (brown) stick" for incorrect responses. All subjects were trained to a criterion of four consecutive correct judgments on each pair (four consecutive correct blocks) or twelve blocks, whichever came first. The training criterion level for the passive conditions was determined by reviewing the Riley and Trabasso (1974) trial
of last error data for ordered pair and randomized block training for their subjects age 4-1 to 5-0.

**Passive Randomized Block Training** - The four comparisons of adjacent stick pairs (AB, BC, CD, DE) were trained in a randomized block order. These pairs were presented in blocks of four pairs with pairs randomized within each block. The five sticks were first adjusted in length by the experimenter to form the correct color A>B>C>D>E sequence. Placement of sticks was randomized in the same way as in passive ordered pairs training. The same comparative questions used in passive ordered pairs training were used in this condition with questions occurring randomly but equally often over trials and left-right rack positions. The subject's response was a judgment indicating the color of the correct stick choice. All feedback was both visual and verbal and of the same form as that used in passive ordered pairs training. All subjects were trained to a criterion of four consecutive correct judgments on each pair (four consecutive correct blocks) or twelve blocks, whichever came first.

**Testing**

Testing followed immediately after the training condition was completed. Testing was the same as the passive training conditions except that no feedback was given. Each subject was tested two times on each of 10 possible color pairs. These 10 pairs included the four direct comparisons used in training and six new, indirect comparisons (AC, AD, AE, BD, BE, CE). The two test questions on each pair inclu-
ded one "longer" question and one "shorter" question. For each form of the questions the sticks were in either the left or right rack positions equally often. The 20 resulting test questions were randomized in two blocks of 10 pairs with the constraint that the same pair could not be presented in two consecutive questions. The total number of errors on the 12 indirect comparison questions was used as the dependent variable in this study.

**Response Required** - Two different responses were required during the testing phase; judgments only or judgments plus explanations. One-half of the subjects at each age level were asked to make a judgment only and the other half were asked to make a judgment plus an explanation for that judgment on the 12 indirect comparisons (inference pairs). Equal numbers of boys and girls in each of the four training conditions and in each color order were given each of the two testing conditions.

**Judgment Only Response** - Subjects were asked to state the color of the correct stick choice in response to the test question "Which stick is longer?" or "Which stick is shorter?"

**Judgment Plus Explanation Response** - Subjects were first asked to state the color of the correct stick choice in response to the test question "Which stick is longer?" or "Which stick is shorter?" Following each judgment response the subject was asked "How do you know that the (green) stick is longer (shorter) than the (brown) stick?" Further probing such as "Can you tell me a little more?" or
"Do you remember anything from when you practiced before that would help you remember now?" was used when indicated by the subject's response to the explanation question. A complete correct explanation for inference pairs was one which explicitly mentioned at least one of the correct premises for that test pair (e.g., for A>D the subject may say A>B, B>C, or C>D). The criterion for a correct explanation was the same as those used by both Smedslund (1960) and Roodin and Gruen (1970).

Summary and Treatment of the Data

To summarize, sex of subject was controlled and color order was counterbalanced over sex, age, and training condition. Type of comparative ("longer" or "shorter") used was counterbalanced over trials and left and right rack positions during both the training and testing phases. Order of start was counterbalanced over sex and color order in both the active and passive ordered pairs training conditions. Testing condition (judgment only or judgment plus explanation) was counterbalanced over sex, age, color order, and training condition.

The data collected by both the judgment only and judgment plus explanation measures established the basis for two factorial 3 x 2 x 2 analyses of variance. The factors in the design were age (4-5 to 5-0, 6-0 to 6-7, and 7-6 to 8-0), type of training (active and passive), and order of training pairs (randomized block and ordered pairs). The schematic representation of the experimental design is presented in Figure 2. The level of significance was established at $p < .05$. 
| AGE Block | RANDOM | ACTIVE | PASSIVE | JUDGMENT | JUDGMENT PLUS EXPLANATION |
|-----------|--------|--------|---------|----------|--------------------------|
| **4-5 to 5-0** | BLOCK | PASSIVE | ACTIVE |          |                          |
| ORDERED PAIRS | ACTIVE | PASSIVE |         |          |                          |
| **6-0 to 6-7** | BLOCK | PASSIVE | ACTIVE |          |                          |
| ORDERED PAIRS | ACTIVE | PASSIVE |         |          |                          |
| **7-6 to 8-0** | BLOCK | PASSIVE | ACTIVE |          |                          |
| ORDERED PAIRS | ACTIVE | PASSIVE |         |          |                          |

**FIGURE 2.** Schematic Representation of Treatment Conditions
Chapter III

RESULTS

This chapter has been divided into five sections. The first section presents the preliminary analyses of inference pair errors on all variables which were controlled or counterbalanced in the study. The second section presents the analysis of inference pair errors for the judgment criterion. The third section presents the analysis of inference pair errors for the explanation criterion. The fourth section presents the analysis of the training pair error data during the testing period. Finally, the fifth section presents a categorical analysis of the explanation data from the judgment plus explanation testing condition.

Preliminary Analyses

In order to determine whether or not the judgment data from both the judgment only and judgment plus explanation criterion could be combined, an initial one-way analysis of variance was performed comparing the number of judgment errors in the judgment only condition to the number of judgment errors in the judgment plus explanation condition. An F max test (Winer, 1971) indicated no violation of the homogeneity of variance assumption \( F_{\text{max}} (2,59)= 1.014, p>.05 \). There was no significant difference in the number of inference pair errors \( F (1,118)= .044, p>.05 \). Therefore, it could be concluded that the number of judgment errors made by a subject who is not asked to give an explanation for his/her judgment does not differ significantly from the number of judgment errors made by a subject who is asked to give an
explanation. As a result of this analysis the judgment data from both groups were combined and all further analyses on this criterion were performed for the entire sample (N = 120).

In order to control for possible confounding variables sex (male, female), order of start (pair AB or pair DE) for ordered pair conditions only, and color order (green > brown > blue > red > yellow and yellow > brown > green > red > blue) were each analyzed by a one-way analysis of variance for both the judgment and the explanation criterion. As a preliminary procedure, Hartley's F max test (Winer, 1971) was applied to the data and indicated no violation of the homogeneity of variance assumption for any of these variables: sex \( F_{\text{max}} (2, 59) = 1.163, p > .05 \), order of start \( F_{\text{max}} (2, 29) = 1.002, p > .05 \), and color order \( F_{\text{max}} (2, 59) = 1.403, p > .05 \).

No significant sex difference was found for either the judgment \( F (1, 118) = .078, p > .05 \) or the explanation criterion \( F (1, 58) = .597, p > .05 \). The analyses of variance for order of start in ordered pair conditions yielded no significant differences for either the judgment \( F (1, 58) = .001, p > .05 \) or the explanation criterion \( F (1, 28) = .378, p > .05 \).

A significant difference was found for color order for the judgment criterion \( F (1, 118) = 7.068, p < .05 \) but no significant difference was found for the explanation criterion \( F (1, 58) = 1.743, p > .05 \). The mean inference pair error score for judgments for the green > brown > blue > red > yellow color order was 3.317 (SD = 2.318) and the mean error score for the yellow > brown > green > red > blue color order was 4.550 (SD = 2.746).
Since both color orders were randomly selected from all possible color order combinations no viable explanation can be offered for this significant effect. Furthermore, no explanation can be offered for the significance of this effect for only the judgment criterion and not for the explanation criterion. It was thus decided to treat this result as an artifact and proceed with the rest of the analyses without introducing color order as an independent variable.

Judgment Scores

The means and standard deviations for all judgment data are reported in Table 7 in Appendix B. The data collected for the judgment criterion were analyzed by a 3X2X2 analysis of variance. The analysis contained the following factors: age of subjects (4-5 to 5-0, 6-0 to 6-7, 7-6 to 8-0), type of training (passive and active), and order of training pairs (random block and ordered pair). Hartley's F max test (Winer, 1971) was applied to the data and no violation of the homogeneity of variance assumption was found \( F_{\text{max}} (12,9) = 6.302, p>.05 \). The summary of the analysis of variance of the judgment error scores on inference pairs is presented in Table 1.

As can be seen in Table 1, significant differences were found for age \( F(2,108) = 14.885, p<.05 \), type of training \( F(1,108) = 18.168, p<.05 \), and the interaction of age by type of training \( F(2,108) = 3.793, p<.05 \). The summary of the means comprising this interaction are presented in Table 2. This interaction is depicted in Figure 3.
Table 1
Summary Table of a 3X2X2 Analysis of Variance for Judgments by Age, Order of Training Pairs, and Type of Training

| Source                   | Sum of Squares | Degrees of Freedom | Mean Square | F     |
|--------------------------|----------------|--------------------|-------------|-------|
| Age (A)                  | 142.067        | 2                  | 71.033      | 14.885* |
| Order of Training Pairs (B) | 2.700         | 1                  | 2.700       | .566  |
| Type of Training (C)     | 86.700         | 1                  | 86.700      | 18.168* |
| AXB                      | 11.400         | 2                  | 5.700       | 1.194  |
| AXC                      | 36.200         | 2                  | 18.100      | 3.793* |
| BXC                      | 8.533          | 1                  | 8.533       | 1.788  |
| AXBXC                    | 4.467          | 2                  | 2.233       | .468  |
| Error                    | 515.390        | 108                | 4.772       |       |

*P < .05

Table 2
Summary of Judgment Mean Error Scores and Standard Deviations on Inference Pairs for Age of Subjects With Type of Training

| Age of Subjects | Type of Training | Passive | Active |
|-----------------|------------------|---------|--------|
| 4-5 to 5-0      | M=5.35 SD=2.04   | M=5.55  SD=2.41 |
| 6-0 to 6-7      | M=2.35 SD=2.50   | M=4.45  SD=2.02 |
| 7-6 to 8-0      | M=1.55 SD=1.41   | M=4.35  SD=2.38 |
Figure 3. Interaction of Age and Type of Training for Judgment Errors on Inference Pairs.
Due to the significance of the age by type of training interaction it was necessary to analyze both the age effects within the two levels of type of training and the type of training effects within the three age levels. Simple-effects tests (Winer, 1971) showed significant age differences for passive training only \( F(2,108) = 93.4, p<.05 \). Error scores on active training were not found to be significantly different across age levels \( F(2,108) = 1.858, p>.05 \). A Newman-Keuls test (Winer, 1971) indicated a significant difference between the youngest subjects and subjects in each of the other two age groups for passive training on the judgment criterion \( p<.05 \). The middle age subjects and the oldest subjects did not differ significantly from each other. It was thus found that passive training was significantly more difficult for the youngest subjects (age 4-5 to 5-0) than for subjects in either of the other age groups.

Simple effects also indicated that active training was significantly more difficult than passive training for both the middle age group \( F(1,108) = 9.24, p<.05 \) and the oldest age group \( F(1,108) = 16.43, p<.05 \) but neither type of training was found to be significantly more difficult for the youngest age group \( F(1,108) = .40, p>.05 \).

**Explanation Scores**

The means and standard deviations for all explanation data are reported in Table 8 in Appendix B. The data collected for the explanation criterion were analyzed by a 2X2X2 analysis of variance. The analysis contained the following factors: age of subjects (6-0 to 6-7 and 7-6 to 8-0),
type of training (passive and active), and order of training pairs (random block and ordered pair). The 4-5 to 5-0 age group was excluded from this analysis due to the lack of variability in their explanation scores. The mean number of errors on inference pairs for this group was 11.950 (SD = .224). Only one subject in all conditions gave one correct explanation out of a total of 240 possible correct explanations. The means for the subjects age 6-0 to 6-7 and 7-6 to 8-0 were 6.950 (SD = 4.249) and 4.900 (SD = 3.523) respectively. The means for the three age groups are depicted in Figure 4.

Hartley's F max test (Winer, 1971) was performed and no violation of the assumption of homogeneity of variance was found \([F \text{ max } (8,4) = 7.426, p > .05]\). The summary of the analysis of variance of the explanation error scores on inference pairs is presented in Table 3.

As can be seen in Table 3, a significant difference was found for type of training \([F (1,32) = 6.744, p < .05]\). No other significant differences were found. The mean error score on inference pairs for passive training was 4.350 (SD = 3.553) and for the active training mean was 7.500 (SD = 3.898). It can be concluded that active training was significantly more difficult than passive training for the explanation criterion.

**Errors on Training Pairs**

The means and standard deviations for all training pair data during testing are reported in Table 9 in Appendix B. The data collected on training pair errors during testing were
Figure 4. Mean Age Error Scores on Inference Pairs
For Explanations.
Table 3
Summary Table of a 2X2X2 Analysis of Variance
For Explanations by Age, Order of Training Pairs, and Type of Training

| Source                          | Sum of Squares | Degrees of Freedom | Mean Square | F     |
|---------------------------------|----------------|--------------------|-------------|-------|
| Age (A)                         | 42.025         | 1                  | 42.025      | 2.856 |
| Order of Training Pairs (B)     | 1.225          | 1                  | 1.225       | .083  |
| Type of Training (C)            | 99.225         | 1                  | 99.225      | 6.744*|
| AXB                             | 1.225          | 1                  | 1.225       | .083  |
| AXC                             | 1.225          | 1                  | 1.225       | .083  |
| BXC                             | 2.025          | 1                  | 2.025       | .138  |
| AXBXC                           | 3.025          | 1                  | 3.025       | .206  |
| Error                           | 470.797        | 32                 | 14.712      |       |

*p<.05

analyzed by a 3X2X2 analysis of variance for the judgment criterion only. It should be noted that no explanations were requested for training pairs during the testing phase. The analysis contained the same factors as used in the previous analysis on judgments: age (4-5 to 5-0, 6-0 to 6-7, 7-6 to 8-0), type of training (passive and active), and order of training pairs (random block and ordered pairs). Hartley's F max test (Winer, 1971) was used as a preliminary procedure and no violation of the homogeneity of variance assumption was found \( [F \text{ max (12,9)} = 5.022, p>.05] \). The summary of the analysis of variance of the training pair errors during testing is
Table 4
Summary Table of a 3X2X2 Analysis of Variance for Judgment Errors on Training Pairs During Testing by Age, Order of Training Pairs, and Type of Training

| Source                | Sum of Squares | Degrees of Freedom | Mean Square | F     |
|-----------------------|----------------|--------------------|-------------|-------|
| Age (A)               | 47.617         | 2                  | 23.808      | 15.152* |
| Order of Training Pairs (B) | 4.408         | 1                  | 4.408       | 2.806  |
| Type of Training (C)  | 16.875         | 1                  | 16.875      | 10.740*|
| AXB                   | 3.017          | 2                  | 1.508       | .960   |
| AXC                   | 22.050         | 2                  | 11.025      | 7.017* |
| AXBXC                 | .517           | 2                  | .258        | .164   |
| Error                 | 169.699        | 108                | 1.571       |       |

*P<.05

As can be seen in Table 4, significant differences were found for age \( F(2,108) = 15.152, p<.05 \), type of training \( F(1,108) = 10.740, p<.05 \), and the interaction of age by type of training \( F(2,108) = 7.017, p<.05 \). The summary of the means comprising this interaction are presented in Table 5. This interaction is depicted in Figure 5.

Due to the significance of the age by type of training interaction it was necessary to analyze both the age effects within the two levels of type of training and the type of
Figure 5. Interaction of Age and Type of Training For Judgment Errors On Training Pairs During Testing.
Table 5
Summary of Mean Error Scores and Standard Deviations on Training Pair Judgments for Age of Subjects With Type of Training

| Age of Subjects | Passive | Active |
|-----------------|---------|--------|
| 4-5 to 5-0      | M=4.05  SD=1.35 | M=3.60  SD=1.42 |
| 6-0 to 6-7      | M=1.85  SD=1.24 | M=3.35  SD=1.19 |
| 7-6 to 8-0      | M=1.80  SD=1.11 | M=3.00  SD=1.11 |

training effects within the three age levels. Simple effects tests (Winer, 1971) showed significant age differences for passive training only \[ F(2,108) = 21.03, p<.05 \]. Error scores on active training were not significantly different across age levels \[ F(2,108) = 1.16, p>.05 \]. A Newman-Keuls test (Winer, 1971) indicated a significant difference for passive training between the youngest subjects and subjects in each of the other two age groups on training pair errors during testing \( p<.05 \). The middle age subjects and oldest subjects did not differ from each other. It was thus found that passive training was more difficult for the younger children than for the children in either of the other two age groups.

Simple-effects tests also indicated that active training was found to be significantly more difficult than passive training for both the middle age group \( F(1,108) = 13.06, \)
and the oldest age group \( F(1,108) = 9.17, p < .05 \) but neither type of training was found to be significantly more difficult for the youngest age group \( F(1,108) = 1.27, p > .05 \).

In order to determine the relationship between memory of training pairs and transitive inference skills, correlational analyses were computed between errors on training pairs during testing and both judgment and explanation errors on inference pairs during testing. Both the judgment and explanation measures on inference pairs were found to be significantly related to training pair errors. The correlations were \( r = .54 (p < .001) \) and \( r = .49 (p < .001) \) respectively. Partial correlations for the same comparisons were also computed controlling for age and the same relationships were still found to be significant. The partial correlation between errors on training pairs and judgments on inference pairs was \( r = .45 (p < .001) \) and the partial correlation was \( r = .38 (p < .01) \) between errors on training pairs and explanation errors on inference pairs. It can therefore be concluded that there is a significant relationship between a subject's memory of the premise or training pairs and his/her ability to make judgments or explanations on transitivity questions.

**Categorization of Explanations**

The explanations given by all subjects were analyzed qualitatively. Six types of explanations were distinguished by the experimenter after reviewing the data from this study and previous Piagetian literature on conservation (Elkind,
These six types of explanations were:

Category 1 - Don't Know and Miscellaneous Answers
Category 2 - Romantic Explanations
Category 3 - Perceptually Based Explanations
Category 4 - Use of Irrelevant Relationships
Category 5 - Use of Correct Relationships
Category 6 - Use of Correct Relationships and/or Irrelevant Relationships

A more detailed definition of each category is presented in Appendix C.

The twelve explanations for all subjects were randomized and categorized by the experimenter and two other independent raters. Each explanation was categorized in one category only. The percentage of agreement among the three raters was 98.47%. There was agreement between at least two out of three raters on all ratings. The percentage of subjects at each age level giving explanations in each of the six categories was computed. A summary of these percentages is presented in Table 6. It can be seen in Table 6 that the percentage of subjects giving responses in Categories 2 and 3 decrease with age while the percentage of subjects giving responses in Categories 4, 5, and 6 increase with age. The percentage of subjects giving responses in Category 1 is almost equal for ages 4-5 to 5-0 and 6-0 to 6-7 and somewhat less for age 7-6 to 8-0. These findings suggest that younger subjects (age 4-5 to 5-0) tend to give transitive inference explanations which are based upon romantic reasons or reasons which indicate an attempt to utilize direct perceptual cues.
Table 6
Summary of Percentage of Subjects Giving Explanations in Each Category

| Category | 4-5 to 5-0 | 6-0 to 6-7 | 7-6 to 8-0 |
|----------|------------|------------|------------|
| 1        | 27.92%     | 34.58%     | 14.17%     |
| 2        | 28.75%     | 5.00%      | 2.50%      |
| 3        | 37.92%     | 6.67%      | 4.17%      |
| 4        | 5.00%      | 11.67%     | 17.50%     |
| 5        | .40%       | 20.83%     | 41.67%     |
| 6        | 0%         | 21.25%     | 19.58%     |

Children in the middle (age 6-0 to 6-7) and older groups (age 7-6 to 8-0) tend to give explanations which seem to indicate an attempt to utilize a linear ordering strategy where A>B>C>D>E and inferences can be made by simply remembering the order of the five stick sequence.

Examination of the frequency distributions of error scores on each criterion reveals that the normality assumption underlying the analysis of variance is not met for all three age levels. It can also be seen that the departures from normality are not the same for all age groups. The effects of departures from the normality assumption on the probability of a Type I error have been thoroughly discussed by Pearson (1931), Lindquist (1953), Box and Anderson (1955), and Boneau (1960). The conclusions from all of these articles
are that the $F$ distribution is generally insensitive to the form of the distribution criterion measure and therefore the departure from normality will probably have no appreciable effect on the validity of the $F$ test. The normality assumption can be violated and the probability of a Type I error remains almost exactly at the level selected by the experimenter.

However if one were to be conservative then three alternative procedures are available; non-parametric analyses, statistical transformations of the raw data, or use of an approximate test where the obtained $F$ value is tested against a more conservative criterion where $F = \frac{MS \text{ treatment}}{MS \text{ error}}$ is assumed to be distributed as an $F$ distribution with 1 and $n-1$ degrees of freedom (Winer, 1971). The choice of non-parametric analyses is not feasible because there is no non-parametric equivalent available for the multivariate analyses needed in this study. The use of transformations is also not feasible since no one transformation would normalize all of the frequency distributions found for the error data on both criteria. Transformations are also considered to be inappropriate for data where there is homogeneity of variance but non-normality (Lindquist, 1953). The most reasonable alternative procedure is to use an approximate test with a more conservative criterion value for $F$ at $p < .05$. When an approximate test is applied to the data in this study all $F$ values previously found to be significant remain significant except for the following: the age by type of
training interaction for both the judgment criterion and the errors on training pairs and the type of training effect for the explanation criterion. As a result of this conservative test of significance these three findings may be considered tenuous. A more thorough discussion of the tentativeness of the type of training effects is presented in Chapter IV.
Chapter IV

DISCUSSION

The major objectives of the present investigation were to determine whether or not differences in the age of emergence of transitivity are due to the following methodological variables: (a) type of training, (b) order of presentation of training pairs, and/or (c) type of response criteria demanded.

The first hypothesis was concerned with the degree of transitivity evidenced at different age levels when different response criteria are employed. It was specifically predicted that transitivity would increase with age for both the judgment and the explanation criterion. This hypothesis was partially supported. Subjects in the youngest age group (4-5 to 5-0) demonstrated significantly less transitivity than subjects in the other two age groups (6-0 to 6-7 and 7-6 to 8-0). Yet, the subjects in the latter two age groups did not differ significantly from each other on either criterion.

The second hypothesis was concerned with the effect of the more stringent judgment plus explanation criterion on the age of emergence. Specifically it was predicted that the judgment plus explanation criterion would be more difficult for younger children than for older children. The results supported this hypothesis. Only one child in the youngest age group gave a correct explanation. Children in both of the other age groups gave significantly more correct explanations.

The third hypothesis was concerned with the effect of the type of training task utilized on transitivity. It was
predicted that children who received active training would show more transitivity than children who received passive training. The direction of this hypothesis was not supported. Instead, the results suggest that passive training is superior to active training, at least as active and passive have been defined here. On both the judgment and explanation criteria children who received passive training evidenced more transitivity than children who received active training.

The fourth hypothesis was concerned with the effects of the order of the training pairs on transitivity. It was specifically predicted that children who received randomized block training would show less transitivity than children who received ordered pairs training. This hypothesis was not supported by the results. No significant differences were found for order of training pairs for either the judgment or explanation criterion.

The fifth hypothesis was concerned with the interaction of the order of training pairs and the age of emergence of transitivity. It was predicted that randomized block training would be more difficult for younger than for older children. The results did not support this hypothesis. No significant differences were found for age and order of training pairs.

The sixth hypothesis was concerned with the interaction of the type of training and the age of emergence of transitivity. It was predicted that the facilitative effect of the active training condition would be more pronounced for younger children than for older children. The results did not support this hypothesis. Instead, active training was found to be
significantly more difficult than passive training for subjects in the middle (6-0 to 6-7) and older (7-6 to 8-0) age groups on the judgment criterion only. No significant difference in type of training was found for the youngest subjects (age 4-5 to 5-0) on either criterion.

No specific hypotheses were originally made concerning the effect of type of training and/or order of training pairs on judgment errors on training pairs during testing. Yet, these results were also computed and supported the results found for the judgment criterion. No significant difference was found for order of training pairs (random vs. ordered). Active training was found to be significantly more difficult than passive training and more specifically, active training was significantly more difficult than passive training for subjects in the middle (6-0 to 6-7) and older (7-6 to 8-0) age groups. No difference in type of training was found for subjects in the youngest age groups.

**Overall Implications and Methodological Issues**

The evidence in support of the contention that type of response criteria demanded affects the age of emergence of transitivity is consistent with the findings of the few previous investigators who have addressed this question in their work (Brainerd, 1973; 1977; Roodin & Gruen, 1970). In Trabasso's original work (Bryant & Trabasso, 1971) children as young as four years old were able to make correct transitive inference judgments. The findings of this study support Trabasso's conclusions for the judgment criterion only. Yet,
the youngest children were almost always unable to give a correct explanation for the same inference pairs. This finding supports the Piagetian position which places the age of emergence of transitivity at a later point, approximately seven to eight years old. It appears then that the selection of a response criterion will have an important effect upon the age at which subjects will be found to demonstrate transitivity. It should also be noted that the analysis of the judgment data in both the judgment only and judgment plus explanation criterion yielded no significant differences. That is to say that asking a subject to give an explanation on an inference pair did not significantly influence the subject's ability to give judgments on subsequent inference pairs. Thus future investigations of the effects of response criteria on transitivity should not need to employ a judgment only condition for comparison of judgment and explanation data.

A major question to be answered is why the explanation criterion is so difficult for the young child? One possibility is that the young child (age 4 to 5 years) does not yet possess the verbal skills which would enable him/her to give a correct explanation on an inference task. This hypothesis would seem plausible if, in fact, the young children found it difficult to give any verbal explanation at all. The categorization data for the explanation criterion partially support this idea. Yet, although children in the youngest age group did not provide explanations which were considered correct under the criterion used in this study (categories 5 and 6) they did provide a number of other types
of explanations which Piagetians would claim are representative of preoperational level thinking. The young children generally provided explanations which could be labeled romantic (reasons based upon wish fulfillment or other extraneous criteria) or perceptually based (reasons which showed attempts to utilize perceptual cues rather than an overall ordering strategy). In summary, it appears that young children age four to five years do provide verbal explanations for transitive inferences which are appropriate for their structural level. The young child's verbal reasoning is dominated by perceptual cues and he/she cannot yet express an understanding of the reversibility of ordered relations. Thus, although reaction time and error data from the work of Trabasso and his colleagues (Trabasso, 1975; Trabasso, Riley, & Wilson, 1975) seem to indicate that the young child (age 4 to 5 years) uses a linear ordering strategy (where the child orders premise pairs in a sequence e.g. A>B>C>D>E) to make transitive inferences, the reasoning he/she uses to verbally explain the same inference judgments is still at a preoperational level. Trabasso and his colleagues have proposed that children studied do not use operational transitivity (coordination of the end terms via a middle term) to solve the problem. Instead, they believe that subjects integrate the initial information into an ordered, spatial array. This array is constructed during training, stored in memory, and internally scanned when inferential questions are posed. The subject thus makes an inference by "observing" the relation between the members of a pair in the spatial
memory array, and not by operational transitivity. In spite of Trabasso's hypothesis and findings the kinds of preoperational level explanations given by the younger subjects in this study indicate little or no awareness of the possible ordering strategy involved. Although the Piagetian theorists have attempted to determine and validate the age of emergence for transitive reasoning, they have yet to study in detail the actual cognitive processes which comprise this skill. If young children reason in a qualitatively different manner than their older peers then what are the differences, and how can they be demonstrated and measured in transitive inference and other tasks? It would be particularly useful for Piagetians to accept the challenge of formulating a theoretical model of transitive inference that would account for the available data as well as does the linear order hypothesis, and would lead to new predictions.

A question that must also be asked is why the transitive skills of the children age 6-0 to 6-7 did not differ significantly from those of children age 7-6 to 8-0 on either the judgment or explanation criterion. This finding supports the research of Trabasso and his colleagues but Piagetians (who place the age of emergence at seven to eight years) would expect to find a difference between these age groups on the explanation criterion. It is possible that the criterion for a correct explanation (subject states at least one correct premise training pair) may have resulted in too low a ceiling for subjects in these age groups. The categorization data seem to support this conclusion. As can be seen
in Table 6, subjects age 7-6 to 8-0 gave many more explanations in which only correct training pairs were mentioned (category 5) than did the subjects age 6-0 to 6-7. It is possible that a more stringent explanation criterion than the one used in this study and that of previous researchers (Smedslund, 1963; Roodin & Gruen, 1970) would help to differentiate the levels of transitivity for subjects in the middle and older age groups on the explanation criterion. Once again, it can be seen how the choice of a response criterion, including the levels demanded within a type of response criterion, can influence the age of emergence of transitivity.

The levels of judgment errors in all age groups were higher than those previously found by Trabasso and his colleagues (Bryant & Trabasso, 1971; Riley & Trabasso, 1974; Trabasso, Riley, & Wilson, 1975). Three major variations may have contributed to these differences. First, the studies of Trabasso and his colleagues were usually conducted in a laboratory setting, over a period of two or three days (one session a day), and with monetary reinforcement in the form of a payment for participation in the project. The present study was conducted in a school setting, in one session, and without payment for participation. It is a reasonable possibility that this extreme change of training and testing conditions may have effected the demonstrated levels of transitivity. Second, and probably most important, the subjects in the present study usually received many less training trials than did the subjects in the Trabasso work. The subjects in the present study received a minimum of four training blocks
(16 trials; four trials on each training pair) and a maximum of twelve training blocks (48 trials; twelve trials on each training pair). In other words, all subjects in this study were trained to a criterion of four consecutive correct trial blocks or twelve trial blocks, whichever came first. Subjects in Trabasso's work were usually trained to a criterion of six consecutive correct trial blocks. If the subject did not reach criterion in 100 trials (25 trial blocks) training was discontinued for that day and resumed the next day. If the subject did not reach criterion in 100 more training trials he/she was not tested. The time constraints imposed upon the present study by the public and private school settings in which the work was conducted made it impossible to train subjects to the levels used in the Trabasso work. It is more than likely that the lower criterion and training levels used in this study resulted in higher overall error rates on transitive inference pairs. Thirdly, Trabasso and his colleagues have always trained their subjects using ordered pair training first followed immediately by randomized block training. In the present work, subjects received either randomized block training or ordered pair training. It seems plausible that the double training method of Trabasso's group provided greater and more varied exposure to training pairs and thus resulted in lower overall error rates during testing.

These three hypotheses are further supported by the significant correlations and partial correlations between errors on training pairs during testing and errors on inference pairs for both the judgment and the explanation
criterion. These correlations and the hypotheses proposed above for the higher error rates found in this study support the Trabasso position that ability of children to make transitive inferences is dependent upon their memory of the original premise pair information.

The type of training effect is opposite in direction to the original prediction that passive training would be more difficult than active training. In spite of its significance, methodological difficulties would seem to make this finding a tentative one. Due to the nature of the active task no subject could have had an incorrect training trial. In other words, all subjects always correctly followed the active task directions to "Make the (red) stick longer than the (blue) one" and all subjects always correctly identified the longer or shorter stick on every active training trial. Therefore, all subjects in the active condition received only four consecutive trial blocks (16 trials) of training. Subjects in the passive task were much more likely to have an incorrect trial and usually received between eight and twelve trial blocks (32 trials to 48 trials) of training. Given this difference in the number of training trials it is highly probable that the difficulty of the active task is due to the amount of training and exposure to premise pairs and is not necessarily due to type of training. At this point it seems that the effect of type of training (active vs. passive) still needs to be demonstrated. It can be argued that the designs of the active and passive training tasks themselves are reasonable measures of this dimension provided that the
number of training trials is held constant. The tasks themselves do seem to adequately vary the active/passive dimension but the differences in number of training trials did not allow a direct and controlled comparison to be made.

The lack of significance for order of training pairs (random block, ordered pair) would seem to indicate that the manner in which training pairs are arranged has little effect upon a subject's ability to make transitive inferences. It was originally hypothesized that since random block training requires subjects to discover the order of the training pairs for themselves (as did the Piagetian tower task) it would thus be more difficult than ordered pairs training. Riley (1977) reports unpublished work in which she found significant reaction time differences between random block and ordered pairs training for third grade subjects. Subjects who received random block training took three times longer to make a judgment during testing on inference pairs. Yet, no significant order of training pairs differences were found on the number of judgment errors on inference pairs during training. Riley hypothesizes that random block training may force subjects to organize the premise pairs in an ordered linear sequence while ordered pairs training may not. Since reaction time data was not collected in this study, the results of this work with younger subjects can only partially support Riley's (1977) findings.

Given the setting in which this study was conducted and the time constraints imposed by the schools themselves, both visual and verbal feedback were used to facilitate training.
It is thus possible that subjects' judgments could have been based upon "nontransitive" hypotheses (Smedslund, 1963). Subjects may have remembered the lengths of the sticks (especially in passive training where the lengths remained constant) and thus based their responses upon their memory of these absolute measurements. The categorization data on Table 6 for subjects in the judgment plus explanation condition provides a source of analysis for this hypothesis.

If this hypothesis were to be supported, one would expect a high percentage of subjects' explanations to fall into category 3 (perceptually based explanations). Subjects who gave explanations in this category reported using absolute length or some other kind of measurement to make their judgments on inference pairs. It can be seen that category 3 explanations were given by 37.92% of the subjects age 4-5 to 5-0, but only 6.67% by subjects age 6-0 to 6-7, and only 4.17% by subjects age 7-6 to 8-0. It seems plausible to assume that, in spite of the visual feedback, subjects in the middle and older age groups infrequently used "nontransitive" perceptually based hypotheses while subjects in the youngest age group often employed these strategies. Whether or not the younger subjects' frequent use of perceptually based explanations is dependent upon visual feedback is, as yet, a difficult and unanswered question. Piagetians would continue to argue that the young child's dependency on perceptual cues is more a matter of structural level than the type of training feedback.
Future Directions

The results of this study have initially demonstrated the importance of response criteria for future research on transitivity. Further examination of this dependent variable is still needed for children of varying ages. Given the extreme differences on both criteria between the subjects age 4-5 to 5-0 and subjects age 6-0 to 6-7 it is suggested that future research include, subjects age 5-0 to 6-0 (kindergarteners). Data from this group of subjects should provide a clearer and more complete picture of the development of transitivity.

Studies on transitivity have usually employed age as an independent variable rather than as a dependent variable. Piagetian researchers often use age as a dependent variable in order to determine the age of emergence and developmental sequence for various cognitive skills (e.g. conservation, seriation). It is thus suggested that Trabasso's transitivity task (Bryant & Trabasso, 1971) or a variation of that task should be used in a study where age is considered a dependent variable. Data from this type of study using children of many age levels would also yield valuable information on the natural course of development of the transitive inference skill.

The influence of methodological variations in type of training or type of task is still undetermined. It is necessary to maximize and hold constant the number of training trials while examining active and passive training. An extension of the active/passive distinction into the
testing phases of transitivity tasks would also allow a more complete investigation of this independent variable.

Reaction time data during testing was not collected in this study. The unpublished findings of Riley (1977) would seem to indicate that differences which are not significant using an inference pair error score criterion may be significant using reaction time data as a dependent variable. Future studies of transitivity should be careful to include analyses of reaction time data.

It does seem clear that often held assumptions of the equivalence of methodology used in studies on transitivity are unwarranted. These assumptions appear to confound both the arguments over the age of emergence of this skill and the theoretical arguments over the cognitive processes involved. Transitivity skills may best be investigated by recognizing that choice of methodology is often dependent upon one's position concerning the nature of cognitive development and that choice of methodology may also bias results in one direction or another. Methodological issues and the effects of specific methodological variations should no longer be ignored in the examination of transitivity nor in the examination of all cognitive developmental skills.
BIBLIOGRAPHY

Achenbach, T. M. & Weisz, J. R. A longitudinal study of developmental synchrony between conceptual identity, seriation, and transitivity of color, number, and length. *Child Development, 1975, 46*, 840-848.

Boneau, C. A. The effects of violations of assumptions underlying the t test. *Psychological Bulletin, 1960, 47* (1), 49-64.

Box, G. E. P. & Anderson, S. L. Permutation theory in the derivation of robust criteria and the study of departures from assumption. *Journal of the Royal Statistical Society, Series B, 1955, 17*, 1-26.

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

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

Brainerd, C. J. Judgments and explanations as criteria for the presence of cognitive structures. *Psychological Bulletin, 1973, 79*, 172-179.

Brainerd, C. J. Postmortem on judgments, explanations, and Piagetian cognitive structures. *Psychological Bulletin, 1974, 81* (1), 70-71.

Brainerd, C. J. Cognitive development and concept learning: an interpretative review. *Psychological Bulletin, 1977, 88* (5), 919-939.
Brainerd, C. J. Response criteria in concept development research. *Child Development*, 1977, *48*, 360-366.

Brainerd, C. J. & Hooper, F. H. A methodological analysis of developmental studies of identity conservation and equivalence conservation. *Psychological Bulletin*, 1975, *82*(5), 725-737.

Bryant, P. E. What the young child has to learn about logic. In B. A. Hinde & J. Stevenson-Hinde (Eds.) *Constraints on Learning*. New York: Academic Press, 1973.

Bryant, P. E. *Perception and Understanding in Young Children: An Experimental Approach*. New York: Basic Books, Inc. 1974.

Bryant, P. E. & Kopytynska, H. Spontaneous measurement by young children. *Nature*, 1976, 260, 773.

Bryant, P. E. & Trabasso, T. Transitive inferences and memory in young children. *Nature*, 1971, *232*, 456-458.

DeBoysson-Bardies, B. & O'Regan, K. What children do in spite of adults' hypotheses. *Nature*, 1973, *246*, 531-534.

Elkind D. Children's discovery of the conservation of mass, weight, and volume: Piaget replication study II. *Journal of Genetic Psychology*, 1961, *98*, 219-227.

Flavell, J. H. *The Developmental Psychology of Jean Piaget*. New York: D. Van Nostrand Company, 1963.

Flavell, J. H. *Cognitive Development*. New Jersey: Prentice-Hall, Inc., 1977.
Flavell, J. H. & Wohwill, J. F. Formal and functional aspects of cognitive development. In D. Elkind & J. H. Flavell (Eds.) Studies in Cognitive Development. New York: Oxford University Press, 1969.

Glass, G. V. & Stanley, J. C. Statistical Methods in Education and Psychology. Englewood Cliffs, New Jersey: Prentice-Hall Inc., 1970.

Grove, P. B. (Ed. in Chief) Webster's Seventh New Collegiate Dictionary. Springfield, Mass.: G. & C. Merriam Co., 1967.

Harris, P. L. & Bassett, E. Transitive inferences by four-year-old children? Developmental Psychology, 1975, 11 (6), 875-876.

Lindquist, E. F. The Norton study of the effects of non-normality and heterogeneity of variance. In E. F. Lindquist Design and Analysis of Experiments in Psychology and Education. New York: Houghton Mifflin Company, 1953.

Lutkus, A & Trabasso, T. Transitive inferences by preoperational, retarded adolescents. American Journal of Mental Deficiency, 1974, 78, 599-606.

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

Murray, J. P. & Youniss, J. Achievement of inferential transitivity and its relation to serial ordering. Child Development, 1968, 39, 1259-1268.
Osherson, D. N. *Logical Abilities in Children.* Potomac, Maryland: Lawrence Erlbaum Associates, 1974.

Pearson, E. S. The analysis of variance in cases of non-normal variation. *Biometrika*, 1931, 23, 114-133.

Piaget, J. *The Child's Conception of Number.* New York: W. W. Norton & Company, Inc., 1965.

Piaget, J. *Judgment and Reasoning in the Child.* Totowa, New Jersey: Littlefield, Adams, & Company, 1972.

Piaget, J., Inhelder, B., & Szeminska, A. *The Child's Conception of Geometry.* London: Routedge & Kegan Press, 1960.

Reese, H. W. & Schack, M. L. Comment on Brainerd's criteria for cognitive structures. *Psychological Bulletin*, 1974, 81(1), 67-69.

Riley, C. A. Representation and use of comparative information and inference making by young children (Doctoral dissertation, Princeton University, 1975). *Dissertation Abstracts International*, 1976, 36, 4210-4211B.

Riley, C. A. The representation of comparative relations and the transitive inference task. *Journal of Experimental Child Psychology*, 1976, 22, 1-22.

Riley, C. A. Personal communication, November, 1977.

Riley, C. A. & Trabasso, T. Comparative, logical structures, and encoding in a transitive inference task. *Journal of Experimental Child Psychology*, 1974, 17, 187-203.

Roodin, M. L. & Gruen, G. E. The role of memory in making transitive judgments. *Journal of Experimental Child Psychology*, 1970, 10, 264-275.
Smedslund, J. The problem of "what is learned." Psychological Review, 1953, 60, 157-158.

Smedslund, J. Transitivity of preference patterns as seen by pre-school children. Scandinavian Journal of Psychology, 1960, 1, 49-54.

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

Smedslund, J. The development of transitivity of length: a comment on Braine's reply. Child Development, 1965, 36, 577-580.

Smedslund, J. Psychological diagnostics. Psychological Bulletin, 1969, 71, 237-248.

Stetson, P. C. Verbal transitivity in children (Doctoral dissertation, University of Delaware, 1974). Dissertation Abstracts International, October, 1974, 4-A, 2064-2065 (University Microfilms No. 74-22, 218).

Trabasso, T. Representation, memory, and reasoning: how do we make transitive inferences? In A. D. Pick (Ed.) Minnesota Symposia on Child Psychology (vol. 9). Minneapolis: University of Minnesota Press, 1975.

Trabasso, T. The role of memory as a system in making transitive inferences. In R. V. Kail & J. W. Jagen (Eds.) Perspectives on The Development of Memory and Cognition. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1976.

Trabasso, T., Riley, C. A., & Wilson, E. G. The representation of linear order and spatial strategies in reasoning: a developmental study. In R. J. Palmagne (Ed.) Reasoning: Representation and Process in Children and Adults. Hillsdale, New Jersey: Erlbaum Associates, 1975.
Winer, B. J. *Statistical Principles in Experimental Design.* New York: McGraw Hill, 1971.

Wohlwill, J. F. From perception to inference: a dimension of cognitive development. In I. E. Sigel & F. H. Hooper (Eds.) *Logical Thinking in Children.* New York: Holt, Rinehart, & Winston, Inc., 1968.

Youniss, J. Inference as a developmental construction. In R. J. Falmagne (Ed.) *Reasoning: Representation and Process in Children and Adults.* Hillsdale, New Jersey: Erlbaum Associates, 1975.

Youniss, J. & Furth, H. G. Reasoning and Piaget. *Nature,* 1973, 244, 314-316.

Youniss, J. & Murray, J. P. Transitive inference with non-transitive solutions controlled. *Developmental Psychology,* 1970, 2(2), 169-175.
| Subject | Sex | Color Order | Order of Start | J only or J plus E | Age | Random/Ordered |
|---------|-----|-------------|----------------|--------------------|-----|----------------|
| 001     | 1   | 1           | 1              | 1                  | 1   | 1              |
| 002     | 2   | 1           | 1              | 1                  | 1   | 1              |
| 003     | 1   | 2           | 1              | 1                  | 1   | 1              |
| 004     | 1   | 1           | 1              | 1                  | 1   | 1              |
| 005     | 2   | 2           | 1              | 1                  | 1   | 1              |
| 006     | 2   | 1           | 2              | 1                  | 1   | 1              |
| 007     | 2   | 2           | 2              | 1                  | 1   | 1              |
| 008     | 1   | 1           | 2              | 1                  | 1   | 1              |
| 009     | 2   | 2           | 2              | 1                  | 1   | 1              |
| 010     | 1   | 2           | 2              | 1                  | 1   | 1              |
| 011     | 2   | 1           | 1              | 1                  | 1   | 2              |
| 012     | 1   | 1           | 1              | 1                  | 1   | 2              |
| 013     | 1   | 2           | 2              | 1                  | 1   | 2              |
| 014     | 2   | 2           | 2              | 1                  | 1   | 2              |
| 015     | 2   | 2           | 1              | 1                  | 1   | 2              |
| 016     | 2   | 1           | 2              | 1                  | 1   | 2              |
| 017     | 1   | 1           | 1              | 2                  | 1   | 2              |
| 018     | 1   | 1           | 2              | 2                  | 1   | 2              |
| 019     | 1   | 2           | 1              | 2                  | 1   | 2              |
| 020     | 2   | 2           | 2              | 2                  | 1   | 2              |
| 021     | 1   | 1           | 1              | 1                  | 1   | 1              |
| 022     | 2   | 1           | 1              | 1                  | 1   | 1              |
| 023     | 2   | 1           | 1              | 1                  | 1   | 1              |
| 024     | 1   | 2           | 1              | 1                  | 1   | 1              |
| 025     | 2   | 2           | 1              | 1                  | 1   | 1              |
| Subject | Sex | Color Order | Order of Start | J only or J plus E | Age | Random/Ordered |
|---------|-----|-------------|----------------|-------------------|-----|---------------|
| 026     | 1   | 1           | 2              | 1                 |     | 1             |
| 027     | 2   | 1           | 2              | 1                 |     | 1             |
| 028     | 2   | 2           | 2              | 1                 |     | 1             |
| 029     | 1   | 2           | 2              | 1                 |     | 1             |
| 030     | 1   | 2           | 2              | 1                 |     | 1             |
| 031     | 2   | 1           | 2              | 1                 |     | 2             |
| 032     | 1   | 1           | 1              | 1                 |     | 2             |
| 033     | 2   | 2           | 2              | 1                 |     | 2             |
| 034     | 1   | 2           | 2              | 1                 |     | 2             |
| 035     | 1   | 2           | 1              | 1                 |     | 2             |
| 036     | 2   | 1           | 1              | 2                 |     | 2             |
| 037     | 2   | 1           | 1              | 2                 |     | 2             |
| 038     | 1   | 1           | 2              | 2                 |     | 2             |
| 039     | 1   | 2           | 2              | 2                 |     | 2             |
| 040     | 2   | 2           | 1              | 2                 |     | 2             |
| 041     | 2   | 1           | 1              | 2                 |     | 1             |
| 042     | 2   | 1           | 1              | 2                 |     | 1             |
| 043     | 1   | 1           | 1              | 2                 |     | 1             |
| 044     | 2   | 2           | 1              | 2                 |     | 1             |
| 045     | 1   | 2           | 1              | 2                 |     | 1             |
| 046     | 1   | 1           | 2              | 2                 |     | 1             |
| 047     | 2   | 1           | 2              | 2                 |     | 1             |
| 048     | 1   | 2           | 2              | 2                 |     | 1             |
| 049     | 1   | 2           | 2              | 2                 |     | 1             |
| 050     | 2   | 2           | 2              | 2                 |     | 1             |
| Subject | Sex | Color Order | Order of Start | J only or J plus E | Age | Random/Ordered |
|---------|-----|-------------|----------------|-------------------|-----|---------------|
| 051     | 2   | 1           | 1              | 1                 | 2   | 2             |
| 052     | 1   | 1           | 1              | 1                 | 2   | 2             |
| 053     | 1   | 1           | 2              | 1                 | 2   | 2             |
| 054     | 1   | 2           | 1              | 1                 | 2   | 2             |
| 055     | 2   | 2           | 2              | 1                 | 2   | 2             |
| 056     | 2   | 1           | 2              | 2                 | 2   | 2             |
| 057     | 1   | 1           | 2              | 2                 | 2   | 2             |
| 058     | 1   | 2           | 2              | 2                 | 2   | 2             |
| 059     | 2   | 2           | 1              | 2                 | 2   | 2             |
| 060     | 2   | 2           | 1              | 2                 | 2   | 2             |
| 061     | 1   | 1           |                 | 1                 | 2   | 1             |
| 062     | 2   | 1           |                 | 1                 | 2   | 1             |
| 063     | 2   | 2           |                 | 1                 | 2   | 1             |
| 064     | 1   | 2           |                 | 1                 | 2   | 1             |
| 065     | 1   | 2           |                 | 1                 | 2   | 1             |
| 066     | 1   | 1           |                 | 2                 | 2   | 1             |
| 067     | 2   | 1           |                 | 2                 | 2   | 1             |
| 068     | 2   | 1           |                 | 2                 | 2   | 1             |
| 069     | 2   | 2           |                 | 2                 | 2   | 1             |
| 070     | 1   | 2           |                 | 2                 | 2   | 1             |
| 071     | 2   | 1           | 2              | 1                 | 2   | 2             |
| 072     | 1   | 1           | 1              | 1                 | 2   | 2             |
| 073     | 2   | 2           | 2              | 1                 | 2   | 2             |
| 074     | 2   | 2           | 1              | 1                 | 2   | 2             |
| 075     | 1   | 2           | 2              | 1                 | 2   | 2             |
| Subject | Sex | Color Order | Order of Start | J only or J plus E | Age | Random/Ordered |
|---------|-----|-------------|----------------|-------------------|-----|----------------|
| 076     | 2   | 1           | 2              | 2                 | 2   | 2              |
| 077     | 1   | 1           | 1              | 2                 | 2   | 2              |
| 078     | 1   | 1           | 2              | 2                 | 2   | 2              |
| 079     | 1   | 2           | 1              | 2                 | 2   | 2              |
| 080     | 2   | 2           | 1              | 2                 | 2   | 2              |
| 081     | 1   | 1           |                | 1                 | 3   | 1              |
| 082     | 2   | 1           |                | 1                 | 3   | 1              |
| 083     | 1   | 2           |                | 1                 | 3   | 1              |
| 084     | 2   | 2           |                | 1                 | 3   | 1              |
| 085     | 2   | 2           |                | 1                 | 3   | 1              |
| 086     | 1   | 1           |                | 2                 | 3   | 1              |
| 087     | 1   | 1           |                | 2                 | 3   | 1              |
| 088     | 2   | 1           |                | 2                 | 3   | 1              |
| 089     | 1   | 2           |                | 2                 | 3   | 1              |
| 090     | 2   | 2           |                | 2                 | 3   | 1              |
| 091     | 1   | 1           |                | 1                 | 3   | 2              |
| 092     | 2   | 1           |                | 1                 | 3   | 2              |
| 093     | 2   | 2           |                | 2                 | 3   | 2              |
| 094     | 1   | 2           |                | 1                 | 3   | 2              |
| 095     | 1   | 2           |                | 2                 | 3   | 2              |
| 096     | 2   | 1           |                | 2                 | 3   | 2              |
| 097     | 2   | 1           |                | 2                 | 3   | 2              |
| 098     | 1   | 1           |                | 2                 | 3   | 2              |
| 099     | 1   | 2           |                | 2                 | 3   | 2              |
| 100     | 2   | 2           |                | 2                 | 3   | 2              |
| Subject | Sex | Color Order | Order of Start | J only or J plus E | Age | Random/Ordered |
|---------|-----|-------------|----------------|--------------------|-----|----------------|
| 101     | 1   | 1           | 1              | 1                  | 3   | 2              |
| 102     | 1   | 1           | 1              | 1                  | 3   | 1              |
| 103     | 2   | 1           | 1              | 1                  | 3   | 1              |
| 104     | 1   | 2           | 1              | 1                  | 3   | 1              |
| 105     | 2   | 2           | 1              | 1                  | 3   | 1              |
| 106     | 2   | 1           | 2              | 2                  | 3   | 1              |
| 107     | 1   | 1           | 2              | 2                  | 3   | 1              |
| 108     | 1   | 2           | 2              | 2                  | 3   | 1              |
| 109     | 2   | 2           | 2              | 2                  | 3   | 1              |
| 110     | 2   | 2           | 2              | 2                  | 3   | 1              |
| 111     | 2   | 1           | 2              | 2                  | 3   | 1              |
| 112     | 1   | 1           | 2              | 2                  | 3   | 1              |
| 113     | 2   | 1           | 1              | 1                  | 3   | 1              |
| 114     | 1   | 2           | 1              | 1                  | 3   | 1              |
| 115     | 2   | 2           | 1              | 1                  | 3   | 1              |
| 116     | 2   | 1           | 2              | 2                  | 3   | 1              |
| 117     | 1   | 1           | 2              | 2                  | 3   | 1              |
| 118     | 2   | 2           | 1              | 2                  | 3   | 1              |
| 119     | 1   | 2           | 2              | 2                  | 3   | 1              |
| 120     | 1   | 2           | 2              | 2                  | 3   | 1              |
| Subject | Passive/Active* | J score* | E score* | TP score* |
|---------|----------------|----------|----------|-----------|
| 001     | 1              | 5        |          | 2         |
| 002     | 1              | 3        |          | 4         |
| 003     | 1              | 7        |          | 4         |
| 004     | 1              | 4        |          | 5         |
| 005     | 1              | 7        |          | 4         |
| 006     | 1              | 3        | 12       | 3         |
| 007     | 1              | 5        | 12       | 4         |
| 008     | 1              | 4        | 12       | 2         |
| 009     | 1              | 4        | 11       | 5         |
| 010     | 1              | 6        | 12       | 5         |
| 011     | 1              | 7        |          | 3         |
| 012     | 1              | 4        |          | 3         |
| 013     | 1              | 5        |          | 6         |
| 014     | 1              | 4        |          | 4         |
| 015     | 1              | 11       |          | 4         |
| 016     | 1              | 7        | 12       | 2         |
| 017     | 1              | 4        | 12       | 6         |
| 018     | 1              | 2        | 12       | 4         |
| 019     | 1              | 7        | 12       | 4         |
| 020     | 1              | 8        | 12       | 7         |
| 021     | 2              | 3        |          | 1         |
| 022     | 2              | 7        |          | 3         |
| 023     | 2              | 6        |          | 4         |
| 024     | 2              | 8        |          | 6         |
| 025     | 2              | 3        |          | 5         |
| Subject | Passive/Active | J_score | E_score | TP_score |
|---------|---------------|---------|---------|----------|
| 026     | 2             | 2       | 12      | 3        |
| 027     | 2             | 6       | 12      | 3        |
| 028     | 2             | 7       | 12      | 3        |
| 029     | 2             | 6       | 12      | 5        |
| 030     | 2             | 9       | 12      | 5        |
| 031     | 2             | 0       |         | 3        |
| 032     | 2             | 5       |         | 5        |
| 033     | 2             | 6       |         | 5        |
| 034     | 2             | 8       |         | 3        |
| 035     | 2             | 6       |         | 3        |
| 036     | 2             | 6       | 12      | 3        |
| 037     | 2             | 5       | 12      | 2        |
| 038     | 2             | 3       | 12      | 1        |
| 039     | 2             | 9       | 12      | 5        |
| 040     | 2             | 6       | 12      | 4        |
| 041     | 1             | 3       |         | 1        |
| 042     | 1             | 0       |         | 2        |
| 043     | 1             | 0       |         | 2        |
| 044     | 1             | 5       |         | 1        |
| 045     | 1             | 3       |         | 3        |
| 046     | 1             | 0       | 9       | 0        |
| 047     | 1             | 0       | 4       | 3        |
| 048     | 1             | 1       | 4       | 3        |
| 049     | 1             | 1       | 0       | 2        |
| 050     | 1             | 7       | 7       | 3        |
| Subject | Passive/Active | J score | E score | TP score |
|---------|----------------|---------|---------|----------|
| 051     | 1              | 1       |         | 1        |
| 052     | 1              | 8       |         | 2        |
| 053     | 1              | 2       |         | 4        |
| 054     | 1              | 2       |         | 1        |
| 055     | 1              | 0       |         | 0        |
| 056     | 1              | 1       | 7       | 2        |
| 057     | 1              | 4       | 12      | 2        |
| 058     | 1              | 6       | 6       | 4        |
| 059     | 1              | 2       | 2       | 0        |
| 060     | 1              | 1       | 1       | 1        |
| 061     | 2              | 2       |         | 3        |
| 062     | 2              | 8       |         | 4        |
| 063     | 2              | 7       |         | 4        |
| 064     | 2              | 5       |         | 3        |
| 065     | 2              | 4       |         | 3        |
| 066     | 2              | 4       | 8       | 2        |
| 067     | 2              | 6       | 6       | 4        |
| 068     | 2              | 4       | 12      | 3        |
| 069     | 2              | 4       | 4       | 4        |
| 070     | 2              | 6       | 12      | 7        |
| 071     | 2              | 5       |         | 3        |
| 072     | 2              | 3       |         | 3        |
| 073     | 2              | 6       |         | 3        |
| 074     | 2              | 1       |         | 4        |
| 075     | 2              | 6       |         | 4        |
| Subject | Passive/Active | J score | E score | TP score |
|---------|---------------|---------|---------|----------|
| 076     | 2             | 6       | 12      | 3        |
| 077     | 2             | 2       | 11      | 3        |
| 078     | 2             | 4       | 10      | 4        |
| 079     | 2             | 0       | 0       | 1        |
| 080     | 2             | 6       | 12      | 2        |
| 081     | 1             | 1       |         | 1        |
| 082     | 1             | 2       |         | 3        |
| 083     | 1             | 3       |         | 1        |
| 084     | 1             | 0       |         | 1        |
| 085     | 1             | 5       |         | 2        |
| 086     | 1             | 2       | 6       | 3        |
| 087     | 1             | 4       | 4       | 2        |
| 088     | 1             | 0       | 5       | 2        |
| 089     | 1             | 3       | 4       | 3        |
| 090     | 1             | 1       | 1       | 1        |
| 091     | 1             | 1       |         | 1        |
| 092     | 1             | 3       |         | 2        |
| 093     | 1             | 3       |         | 5        |
| 094     | 1             | 0       |         | 2        |
| 095     | 1             | 0       |         | 2        |
| 096     | 1             | 0       | 2       | 2        |
| 097     | 1             | 0       | 0       | 1        |
| 098     | 1             | 1       | 11      | 1        |
| 099     | 1             | 1       | 1       | 0        |
| 100     | 1             | 1       | 1       | 1        |
| Subject | Passive/Active | J score | E score | TP.score |
|---------|----------------|---------|---------|----------|
| 101     | 2              | 3       |         | 3        |
| 102     | 2              | 5       |         | 3        |
| 103     | 2              | 3       |         | 4        |
| 104     | 2              | 8       |         | 3        |
| 105     | 2              | 4       |         | 4        |
| 106     | 2              | 5       | 5       | 4        |
| 107     | 2              | 7       | 7       | 3        |
| 108     | 2              | 7       | 11      | 4        |
| 109     | 2              | 4       | 3       | 3        |
| 110     | 2              | 3       | 3       | 5        |
| 111     | 2              | 6       |         | 1        |
| 112     | 2              | 0       |         | 2        |
| 113     | 2              | 5       |         | 2        |
| 114     | 2              | 5       |         | 4        |
| 115     | 2              | 2       |         | 0        |
| 116     | 2              | 0       | 7       | 3        |
| 117     | 2              | 2       | 2       | 1        |
| 118     | 2              | 9       | 12      | 5        |
| 119     | 2              | 3       | 7       | 3        |
| 120     | 2              | 6       | 6       | 3        |
Raw Data (Cont'd.)

*Sex
  1 = Female
  2 = Male

*Color Order
  1 = Green>Brown>Blue>Red>Yellow
  2 = Yellow>Brown>Green>Red>Blue

*Order of Start
  1 = Pair AB
  2 = Pair DE

*J only or J plus E
  1 = Judgment only
  2 = Judgment plus Explanation

*Age
  1 = 4-5 to 5-0
  2 = 6-0 to 6-7
  3 = 7-6 to 8-0

*Random/Ordered
  1 = Random Pair Training
  2 = Ordered Pair Training

*Passive/Active
  1 = Passive Training
  2 = Active Training

*J score = Judgment Score (# of errors on inference pairs)

*E score = Explanation Score (# of errors on inference pairs)

*TP score = Training Pair Score (# of errors on training pairs)
Table 7
Summary of Mean Error Scores and Standard Deviations on Judgments for Age, Order of Training Pairs, and Type of Training

| Age    | Order of Training Pairs | Type of Training | M   | SD   |
|--------|-------------------------|------------------|-----|------|
| 4-5 to 5-0 | Random                  | Passive          | 4.80| 1.48 |
|        |                         | Active           | 5.70| 2.31 |
|        | Ordered                 | Passive          | 5.90| 2.60 |
|        |                         | Active           | 5.40| 2.50 |
| 6-0 to 6-7 | Random                  | Passive          | 2.00| 2.45 |
|        |                         | Active           | 5.00| 1.76 |
|        | Ordered                 | Passive          | 2.70| 2.54 |
|        |                         | Active           | 3.90| 2.28 |
| 7-6 to 8-0 | Random                  | Passive          | 2.10| 1.66 |
|        |                         | Active           | 4.90| 1.85 |
|        | Ordered                 | Passive          | 1.00| 1.16 |
|        |                         | Active           | 3.80| 2.90 |
| Age       | Order of Training Pairs | Type of Training | M    | SD  |
|-----------|-------------------------|-----------------|------|-----|
| 4-5 to 5-0| Random                  | Passive         | 11.80| .45 |
|           |                         | Active          | 12.00| 0.00|
|           | Ordered                 | Passive         | 12.00| 0.00|
|           |                         | Active          | 12.00| 0.00|
| 6-0 to 6-7| Random                  | Passive         | 4.80 | 3.42|
|           |                         | Active          | 8.40 | 3.58|
|           | Ordered                 | Passive         | 5.60 | 4.39|
|           |                         | Active          | 9.00 | 5.10|
| 7-6 to 8-0| Random                  | Passive         | 4.00 | 1.87|
|           |                         | Active          | 5.80 | 3.35|
|           | Ordered                 | Passive         | 3.00 | 4.53|
|           |                         | Active          | 6.80 | 3.56|
Table 9  
Summary of Mean Training Pair Error Scores  
and Standard Deviations During Testing for Age,  
Order of Training Pairs, and Type of Training

| Age     | Order of Training Pairs | Type of Training | M   | SD  |
|---------|-------------------------|------------------|-----|-----|
| 4-5 to 5-0 | Random                  | Passive         | 3.80| 1.14|
|         |                         | Active          | 3.80| 1.48|
|         | Ordered                 | Passive         | 4.30| 1.57|
|         |                         | Active          | 3.40| 1.35|
| 6-0 to 6-7 | Random                  | Passive         | 2.00| 1.05|
|         |                         | Active          | 3.70| 1.34|
|         | Ordered                 | Passive         | 1.70| 1.42|
|         |                         | Active          | 3.00| 0.94|
| 7-6 to 8-0 | Random                  | Passive         | 1.90| 0.88|
|         |                         | Active          | 3.60| 0.70|
|         | Ordered                 | Passive         | 1.70| 1.34|
|         |                         | Active          | 2.40| 1.51|
Explanation Categories

Category 1

Don't Know and Miscellaneous Answers - gave no reason or said "I don't know." Also includes subjects who merely restate their judgment responses for the test pair (e.g. Subject says "The blue is longer" for a blue/yellow inference pair where he/she has already responded "blue" to the test question "Which one is longer?"). Also includes any responses which do not fit the other categories (e.g. Subject merely lists individual colors of sticks without stating any relationship between the colors named. "The red one and the green one and the brown one helped me remember.").

Category 2

Romantic Explanations - made some attempt to justify choice by reference to criteria based upon his/her own needs and desires or upon some other external characteristic.

a) Refers to reasons which are in no way related to the task at hand (e.g., "My Daddy told me so.").

b) Also includes reasons of egocentrism and wish fulfillment (e.g., "I'm always right" or "I just remembered it" or "I'm thinking" or "I just know it. I know everything.").

Category 3

Perceptually Based Explanations - a) Refers to having seen or played with the sticks in a certain way during training (e.g., "I saw it that way" or "I played with it that way."). b) Refers to verbal feedback by examiner during training "You told me it was that way." (Verbal feedback during training was always given at the same time as the visual feedback was being given).
c) Actually refers to sticks as they are hanging on the rack during testing. This includes references to the heights of the colored tops and/or the side on which the sticks are placed (e.g. "The red is on that side" or "The red one is longer than the blue" (subject actually points to the tops of the sticks trying to show that one actually looks longer (shorter) than the other)).

d) Refers to measurements of sticks in an attempt to recall the actual stick lengths (e.g., "I measured them before and the red one was longer.").

Category 4

Use of Irrelevant Relationships Only - Gives at least one irrelevant inference pair and/or irrelevant training pair and/or incorrect color pair relationship as an answer to the test question. This subject's explanation does not contain any correct training pair relationships relevant to the test question.

a) Irrelevant Inference Pairs - uses other correct inference pair relationships to explain an inference test pair (For A>D, subject says A>C or B>D, etc.).

b) Irrelevant Training Pairs - Uses other correct training pairs to explain inference test pair. These are not the correct training pairs needed to make the test inference in question (For A>D, subject says D>E).

c) Incorrect Pairs - Uses a training or inference pair but states the "longer than" or "shorter than" relationship incorrectly (e.g. Subject says "Red is longer than blue" when actually blue is longer than red for his/her color
order).

d) Correct Pairs - Uses training pairs which are relevant to the inference test pair in question (For A>B, subject says A>B or B>C or C>D).

Category 5

Use of Correct Relationships Only - The subject only cites correct training pairs as explanations for inference test pair relationship. This subject cites no other irrelevant training pairs, irrelevant inference pairs, or incorrect pairs in his/her response.

Category 6

Use of Correct Relationships and/or Irrelevant Relationships - The subject cites at least one correct training pair to explain the inference test pair relationship but he/she also cites at least one irrelevant training pair, or at least one irrelevant inference pair, or at least one incorrect pair in the same explanation for the same inference test pair.