THE UNIVERSITY OF MANITOBA

A CONSTRUCT VALIDATION STUDY OF PIAGET'S EQUILIBRATION PROCESS

by

© PATRICIA BRODSKY

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF PSYCHOLOGY

WINNIPEG, MANITOBA

MAY 1981

A CONSTRUCT VALIDATION STUDY OF PIAGET'S EQUILIBRATION PROCESS

BY

PATRICIA BRODSKY

A thesis submitted to the Faculty of Graduate Studies of
the University of Manitoba in partial fulfillment of the requirements
of the degree of

DOCTOR OF PHILOSOPHY

A CONSTRUCT VALIDATION STUDY OF PIAGET'S EQUILIBRATION PROCESS

PATRICIA BRODSKY

ACKNOWLEDGEMENTS

One who has the privilege of receiving a Ph.D. degree has many people to thank, for this degree cannot be attained without the help of many people.

I wish to thank the chairperson of my dissertation committee, Dr. Shirin Schludermann. Her generous support and help have always been extended to me, over the years. I wish also to thank Dr. Eduard Schludermann and Dr. Chebib who both assisted me with the statistical design and helped me clarify my thinking as regards the implications of this study. I also wish to thank Dr. Winston Rampaul for generously consenting to replace Dr. Chebib on my committee when Dr. Chebib left on sabbatical leave.

It has been a privilege for me to have Dr. Ira Iscoe consent to be the external examiner of my dissertation. His teaching, scholarship, research, friendship, and support over the twenty-four years since I was first his student in an undergraduate child development course has helped me persevere.

I wish also to thank Talmud Torah school and Leona Laxer the principal for allowing me to do my research in the school.

I wish to extend a special thanks to the teachers of the kindergarten and nursery classes. Also my thanks to their assistants. They reassured many a small child about the study. I have special acknowledgement to give to the fifty-seven children who participated in the study. I enjoyed them and they seemed to enjoy our games.

Judy Shane and Dale Rayter served as the other experimenters who monitored the Lafayette galvanometer. Without them I could not have run the study, so a special thanks to these good friends.

I wish also to thank Brenda Shane who trained half the children on the number task.

Michael Dreisel deserves special thanks from me because he helped me analyze my data and guided me through the mysteries of the computer.

Most important to me is the help and support of my family. My husband Marvin has been supportive of my aspirations all the way through our 27 years of marriage. Without his encouragement this degree would not have been possible for me to attain. The study would not have been run without Marvin, as it was, he helped me solve the apparatus problem and helped me construct the apparatus.

Maline and David my children also contributed. David spent many hours gluing Velcro and Maline came to the school and helped with the subjects.

Last, but not least, I would like to thank Lynne Dobbs who typed this dissertation. Lynne has over the years not only typed for me, but has also been a good friend.

ABSTRACT

A CONSTRUCT VALIDATION STUDY OF PIAGET'S EQUILIBRATION PROCESS

PATRICIA BRODSKY

Despite the large number of studies related to Piaget's theory there is little agreement as to the nature of his primary theoretical construct equilibration.

A review of the literature was done, contrasting Piaget's definition of equilibration and experimenters operational definitions of equilibration. Brainerd and the Genevan both use verbal conflict responses, differing as to which is the most appropriate verbal measure, judgments only or judgments, explanation and counter-suggestion. Piaget defines equilibration as a biological autoregulatory process which utilizes the energy of the organism.

Since activation theory and neurophysiological theories of attention conceptualize changes in psychophysiological arousal as reflecting utilization of energy by the organism it was hypothesized that a direct test of Piaget's hypothetical construct, equilibration would involve both a measure of change in arousal as well as verbal conflict.

The present study was designed to determine whether the Brainerd or Genevan or arousal criteria for transition, or non-transition were more predictive of subjects' performances on post-tests of number after training. The criterion for whether a child was transitional or non-transitional was the child's performance on the post-tests.

Five main hypotheses and two methodological hypotheses were tested in this study.

Fifty-seven nursery and kindergarten children were used as subjects. The Genevan criteria contains within it Brainerd's criteria. Therefore, all the children were given both criteria while their galvanic skin response was monitored during the pre-tests.

The pre-tests were adapted from Inhelder et al. 1974 and they were class inclusion, seriation and number.

The subjects were then given five training sessions on a number task adapted from Halford and Fullerton.

Two weeks after training ended the subjects were post-tested for class inclusion and on two tests of number conservation.

The hypotheses were tested by four stepwise multiple regression analyses, canonical correlation analysis, by chi-square analysis, and Pearson Product Moment correlations.

Four out of five of the major hypotheses of this study were confirmed by the results of the study. The two methodological hypotheses were confirmed.

The major hypothesis that increased arousal would reflect accommodation by transitional subjects and no increase would reflect assimilation by non-transitional subjects was confirmed by all the analyses done. The related hypotheses that the Genevan criteria for transition or non-transition would be more predictive of subjects post-test performance than the Brainerd criteria was also confirmed by the canonical correlation analyses and the chi-square analyses.

Two unanticipated results were that training was best predicted by the Genevan criteria and age, and that arousal level during pre-test class inclusion best predicted performance on post-test class inclusion and the number post-tests, despite the fact that none of the verbal pre-test class inclusion measures nor the post-test class inclusion performances were related to post-test number performance.

Two methodological hypotheses concerned with measurement of the GSR were also confirmed. Deflections (changes in ohms resistance) of the GSR correlation with base level resistance were essentially zero, thus disconfirming Wilder's Law of Initial Values as being relevant for the measurement of GSR deflections. Finally, conductance scores which correct for base level did not reach significance in any of the data analyses whereas the Arousal Interval Scale scores which did not correct for base level accounted for significantly more variance than any other pre-test measure. Use of conductance scores as a measure of arousal would have led one to accept the null hypothesis in this study when it was false.

The results of this study were interpreted as validation of Piaget's hypothetical construct, equilibration.

TABLE OF CONTENTS

Continued....

TABLE OF CONTENTS (CONTINUED)

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

1. Introduction and Literature Review

Despite the large number of studies related to Piaget's theory which have been conducted in recent years there is little agreement as to the nature of his primary theoretical construct, equilibration (Larsen, 1977, Miller, 1976).

Equilibration, defined by Piaget (1971a) as an autoregulatory process operating on all levels of functioning of the organism, consists of two subordinate processes; assimilation and accommodation. Assimilation is defined as the basic adaptational process of the organism. Accommodation occurs so that the organism may assimilate environmental elements while maintaining its basic biological integrity. Assimilation has occurred when external elements are integrated with previous structures which remain unaffected structurally. Accommodation is when external elements are integrated with previous structures which are modified to a greater or lesser degree by the integration. However, there is no break of continuity with the previous state (Piaget, 1971a).

Equilibration is a motivating force of all biological organisms in that its goal is the obverse of entropy. It is a force that maintains and increases the hierarchical organization of biological organisms (Piaget, 1971a, Mischel, 1971, von Bertlanffy, 1968, Waddington, 1962).

Investigators of the determinants of equilibration have typically investigated only the cognitive aspect of Piaget's equilibration construct as reflected in the child's verbal behavior during the performance of a Piagetian task (Brainerd, 1977, Larsen, 1977, Miller, 1976).

This study is designed to investigate the motivational aspect of Piaget's equilibration construct. Motivation utilizes energy and Piaget (1971a) has clearly stated that though the starting point for motivation is cognitive, the resulting process for cognitive change utilizes the energy resources of the organism (Piaget, 1971a, Mischel, 1971).

The following section examines the motivational aspect of Piaget's equilibration construct in detail. The presently used operational definitions of equilibration are examined in terms of their being appropriate measures of cognitive equilibration, as Piaget (1960, 1964, 1971a,b,1972) has defined it. The framework for the analysis of equilibration's theoretical definition and its operational definition is done from the point of view of construct validation of an hypothetical construct (Carnap, 1949, MacCorquodale and Meehl, 1956; Marx, Cronbach and Meehl, 1966; Campbell and Fiske, 1966).

1a. Piaget's Definition of Equilibration

The primary theoretical construct in Piaget's theory of cognitive development is the process of equilibration. The function of the process of equilibration is the maintenance of equilibrium or the integrity of the organism both physical and intellectual (Piaget, 1971a). A major problem in the cognitive developmental literature has been the operational definition of equilibration (Larsen, 1977, Miller, 1976; Overton and Reese 1973; Brainerd, 1973a, 1977). Since equilibration is the "cause" of cognitive development and it is supposedly involved in every cognitive "act", it is critical that experimenters be able to determine when it is operating and be able to measure it. Before one can develop measures of a construct it is necessary to define the construct (Carnap, 1949; Cronbach and Meehl, 1966; MacCorquodale and Meehl, 1956; Campbell and Fiske, 1966; Marx, 1964). At present, equilibration is operationally defined in terms of the child's verbal conflict responses, i.e., judgments and/or explanations and responses to challenging questions, when he is doing a cognitive task (Brainerd, 1977; Larsen, 1977). There may be no agreement among the various experimenters as to the method of analyzing and eliciting the child's verbal conflict responses, but all use the verbal conflict response as their basic measure of cognitive conflict, i.e., equilibration (Inhelder, Sinclair, and Bouvet, 1977; Brainerd, 1973a, 1977; Gelman, 1972; Larsen, 1977). The same class of responses are also used in measures of "surprise" or affective conflict (Gelman, 1972; Langer, 1969).

Larsen (1977) in an analytical review of the various methods of defining and measuring equilibration concludes that the quantitative approaches to analyzing verbal conflict scores (re Brainerd) have not shown themselves clearly superior to the qualitative dialectical approach of the Genevan school (p. 1165). When there is a problem in the operational definition of a construct there are several ways of conceptualizing the difficulty, but first one must examine the theoretical construct to determine whether the operational definitions are in fact measuring what the theorist in question meant by the construct. Analytically, the failure to "identify" the equilibration process experimentally may be due to 1)its non-existence, 2) inadequate definition by the theorist, or 3) inappropriate or inadequate measurement of the theoretical definition (Carnap, 1949; Campbell and Fiske, 1966; Cronbach and Meehl, 1966; MacCorquodale and Meehl, 1956; Marx, 1964). Discussion in the literature has focused on the first and last alternatives (Larsen, 1977). The following review of the literature will focus on an examination of Piaget's definition and the measurement of it by researchers. It cannot be determined whether equilibration exists if it is not defined and measured appropriately.

Piaget (1971a) defines equilibration as autoregulation (p.10, 1971a), which means that it is a self-regulatory process of the organism (i.e., it is an intrinsic process that arises out of the structure of the organism).

"...thus cognitive schemata are derived step by step from the preceding one, and in the last analysis they always depend upon coordination of the nervous system and the organic system, in such a way that knowledge is necessarily interdependent on the living organism as a whole. "...they constitute a special part of vast regulator systems by means of which the organism as a whole preserves its autonomy and at the same time, resists entropic decay (Piaget, 1971a, p.13)."

Cognitive development takes place because the process of equilibration is an intrinsic function of a living organism that operates on both the physical and intellectual levels (Piaget, 1971a). His is a biological interpretation of the acquisition of knowledge, and he relies heavily on biological systems theories in his explication of his theory of cognitive development. Piaget (1971a) makes extensive use of von Bertlanffy's (1968) General Systems theory and Waddington's (1962) Epigenetic Systems theory, in defining biological adaptation as an equilibrium between assimilation and accommodation (Piaget, 1971a,

p. 172).

Assimilation and accommodation are the dual processes of biological adaptation. Assimilation has occurred when external elements are integrated with previous structures which remain unaffected. There is no break of continuity with the former state. Accommodation is when external elements are integrated with previous structures which are modified to a greater or lesser degree by the integration. However, there is no break of continuity with the former state. Accommodation cannot, therefore, occur without assimilation (Piaget, 1971a, pp. 4, 172-173). "In fact, in biological terms, accommodation can be nothing but the accommodation of an organized structure, and so it can only be produced, under the influence of some external factor or element, according to whether there is temporary or lasting assimilation of such element or of its extension within the structure that it modifies (Piaget, 1971a, p. 173)."

Equilibration is the autoregulatory process of adaptation or the process of integrating external elements into the ongoing structure of the organism, either without altering structure (assimilation) or with alteration of structure (accommodation), but with no break of continuity with the former state of the organism. Therefore, equilibration is an intrinsic motivational process, motivating the organism at all times and at all levels of functioning (Piaget, 1971a). "...surely this means that the whole range of assimilatory processes which are physiologically typical of metabolism and material exchanges with the environment must also include a classificatory system (p.161)." Thus, equilibration is autoregulation, an intrinsic motivational force, and it utilizes physical energy.

As Mischel (1971) points out, this concept of motivation is radically different from need or drive theories of motivation, or reinforcement theories (Hull, 1943; Skinner, 1953; Spence, 1956; Amsel 1958, 1977; Berlyne, 1965). Motivation is included in the global concept of equilibration: "So understood, cognitive conflict (that is, "awareness of a momentary disequilibrium" in the system of schemas), the "need" for establishing cognitive consistency ("equilibration") between the schemas one has and the novel feature that induced disequilibrium would be the motivation for cognitive activities (Mischel, 1971, p.331)."

Piaget's theory is an analogy with biological open systems theory. Only Bowlby (1969), Werner (1957) and Piaget (1971a) among psychological theorists conceptualize the development of the organism in biological terms (Mischel, 1971; Langer, 1969b); Ainsworth, 1973). Hence, the biological conceptualization of energy source and utilization. The biological system mobilizes energy in terms of its needs which are ultimately determined by its need to maintain its structure and resist entropy (Mischel, 1971, von Bertlanffy, 1962).

Piaget's analogy leads him to the proposition that just as the physical organism "needs" to be internally consistent, i.e., maintain its integrity, the cognitive structure "needs" to be internally consistent (Flavell, 1963;Mischel, 1971; Piaget, 1971a). Energy is mobilized as the child is motivated to accommodate his schemas because his internal logical structure is not internally (logically) consistent until the formal operational stage is reached (Flavell, 1963; Langer, 1969a; Mischel, 1971; Inhelder and Piaget, 1958; Piaget, 1971a).

Another clear implication of Piaget's (1971a) biological interpretation of cognitive development is that his hypothesized cognitive structure on the level of thought and behavior (action) is isomorphic with central nervous system organization. The child at birth has sensorimotor schemas which are built into his central nervous system (Flavell, 1963; Mischel, 1971; Piaget, 1971a). These schemas are modified by his actions on the environment. Discrepancies or "ailments" to his sensory motor schemas result in restructuring of these schemas. Clearly, what is being organized at the physical level is the central nervous system. Such a hypothesis, while untested, is clearly tenable in light of the findings on the plasticity and growth of infants and young children's central nervous system (Sperry, 1975; Kinsbourne, 1978; Walsh and Cummings, 1975; Denenberg, Garbanati, Sherman, Yutzey, and Kaplan, 1978). Piaget (1971a) apparently holds that the nature of the possible organization of the schematic structure is limited by the possible organization of the brain during development. He states that the sequence and type of schemas possible are invariant. They are ultimately limited by the organization of adult logical structure and by implication adult brain structure(Piaget, 1971a; Mischel, 1971; Inhelder and Piaget, 1958; Inhelder, et al., 1974). Piaget does not deal extensively with motivation because he assumes energy utilization would be determined by the needs of the cognitive structure. Therefore, when a child is developmentally able to attend to an "ailment" to his schema (a conflict between his schema and reality) energy would automatically be utilized and the motivation would be cognitive arising from cognitive values or the parallel cognitive affective system (Mischel, 1971; Flavell, 1963). Some have thought that affect would be the motivation (Langer, 1969a), but Piaget (1971a, 1954) has clearly said that affects are determined by cognitive organization. Or in other words, there is affective systematization paralleling the organization of logical schemas (Flavell, 1963; Mischel, 1971; Piaget, 1954). "The child's affects are not, so Piaget argues against Freudian notions of libido, something "given" that can be cathected to different objects; rather, the construction of objects makes possible new feelings connected with these objects (Piaget, 1954, pp. 33-38)(Mischel 1974, p. 317)." The undifferentiated feelings of the infant become systematized parallel with cognitive structuring, and they also acquire an existence independent of the affective situation as do objects (Piaget, 1954). Present day theories of affect and emotions take a similar position. Indeed, cognitive interpretation of a situation as the determiner of emotion seems to be the only reasonable interpretation of the research results (Brodsky and Brodsky, 1978; Hale and Strickland, 1976; Mandler, 1975; Harris and Katkin, 1975; Schacter, 1964; Schacter and Singer, 1962; Valins, 1970; Lindsley, 1970; Berlyne, 1960, 1965; Davidson, 1978).

Clearly, Piaget's (1971a) definition of the equilibration process includes the utilization of energy. That must be part of its definition if it is an autoregulatory process (von Bertlanffy, 1968; Piaget, 1971a; Mischel, 1971). At the organismic level metabolism measures the utilization of energy by organismic processes. At the cortical level increases in "activation or arousal" reflect utilization of energy (Lindsley, 1970; Milner, 1970; Teyler, 1975). When Activation theory was first proposed reticulo-thalamo-cortical arousal was proposed as energizing, and arousal increases were seen as reflecting energy utilization (Duffy 1962; Lindsley, 1951; Malmo, 1959).

None of the operational definitions of equilibration discussed by Larsen (1977), Brainerd (1973), or Gelman (1969) and others include the energy utilization or motivational aspect of Piaget's definitions of the equilibration process, and yet it is relatively easy to measure arousal changes of subjects doing cognitive tasks. Indeed such measures are often more informative about cognitive conflict than verbal reports (Bruner and Postman, 1950). Despite the extensive research since Bruner and Postman's work on perceptual defense based on an arousal measure which clearly indicated conflict when the subjects reported none, only Sroufe (1977) has hypothesized that accommodation is accompanied by an "optimal" increase in arousal (Berlyne, 1960, 1965; Schacter, 1964; Lynn, 1966; Zelazo, 1972; Harris and Katkin, 1975; Mandler, 1975; Hale and Strickland, 1976; Beatty and Wagoner, 1978; Brodsky and Brodsky, 1978). Walsh and Cummins (1975) presented evidence in an extensive review of the literature that such "optimal" increases in arousal arising from environmental stimulation directly stimulate brain development in young animals and children.

1b. Operational Definitions of Equilibration

The use of verbal conflict responses alone as measures of cognitive conflict in children is questionable on empirical as well as theoretical grounds. Namely a number of researchers in the area of the relationship between language development and cognitive development have confirmed Inhelder and Piaget's (1964) findings confirming his theory that early cognitive development precedes language development. They found that the young child could understand certain relationships which he could not verbally express. Subsequent research found that the relationship between language development and cognitive development is complex, in that the child's verbal ability to express physical relationships is behind his cognitive ability to understand them until about age three.

At three or thereabouts, the child's verbal ability outstrips his cognitive ability and he is able to express relationships verbally which he does not understand (Inhelder and Piaget, 1964;Bruner, 1966; Bloom, 1970, 1975; Inhelder et al., 1974; Miller, S., 1976).

Another problem related to the definition of equilibration is found in the "training" or learning literature. It is the difficulty in determining the child's developmental level or whether a child is, for example, a conserver or in transition (e.g., in the process of reorganizing his schemas re: the concept)(Brainerd, 1973a, 1973b, 1974, 1977; Strauss, 1970, Strauss and Langer, 1972; Piaget, 1964, 1972; Inhelder, et al., 1974; Larsen, 1977). There is great difficulty in determining this transitional state using the child's verbal responses. A frequency count is usually resorted to, i.e., frequency of conservation statements or judgments during pre-testing (Brainerd,1974, 1977; Cooley 1978; Gelman, 1972). The Genevans use vacillation in the child's judgment,explanations,and inconsistency in responding to challenging questions, a qualitative measurement technique (Inhelder, et al., 1974).

This issue is important because presumably if Piaget's theory is correct only children in transition will respond to training (Inhelder et al., 1974). Both the qualitative and quantitative methodologies are questionable, because both are based upon the child's verbal ability to express relationships. Thus, some studies find "training" works (Gelman, 1969; Brainerd, 1973b, 1977), and others find that it does not work (Miller, P., 1973; Miller, S., 1976; Reese and Schack, 1974; Inhelder et al., 1974; Cooley, 1978). Put succinctly Brainerd's (1977) issue of false positives and false negatives is over the fact that the Genevan's "...require rather strong evidence. With tests for concrete-operational concepts, for example, children must give both correct judgments and logical explanations of these judgments. For some especially important concepts (e.g., conservation), further behavioral evidence (e.g., resistance to counter suggestion) is sometimes required. Less orthodox investigators will usually accept somewhat weaker evidence. With tests for concrete-operational concepts, they frequently require only correct judgments (Brainerd, 1977, p. 360)." The criteria for determining the level of cognitive development must be stringent, i.e., reliable and valid if the issue between 'learning' and cognitive development is to be resolved. Without a reliable and valid method of discriminating transitional children all training studies can be questioned (Inhelder, et al., 1974; Miller, 1976b). Brainerd (1977) argues for the use of judgment only measures because "empirically" there are fewer false rejections of sequences with judgments-only criteria than there are false acceptances of sequences with judgments and explanations. Larsen (1977) who is interested in determining if there is a "structure d'ensembles" (a cognitive structure) replies: "All of this may be so, but considering that error rate is determined by both false acceptance and false rejections (type I and type II error), it would seem that a better comparison would be in terms of both of these; and since it can be assumed that if one is high the other is low, they would add up about the same in each case. Thus, there does not seem to be any purely psychometric argument for one criterion over the other (p. 1162)." Brainerd (1973a, 1973b, 1977) has conceptualized the measurement of equilibration as a psychometric problem. Indeed, any psychological measurement of behavior is a psychometric problem. If one takes Brainerd's analysis further psychometric theory requires that a measure have construct validity (Anastasi, 1969). The issue being discussed by Larsen (1977), Miller, S. (1976), and others is the problem of construct validity. "Construct validation is involved whenever a test is to be interpreted as a measure of some attribute or quality which is not "operationally defined." "Construct validity must be investigated whenever no criterion or universe of content is accepted as entirely adequate to define the quality to be measured (Cronbach and Meehl, 1966, p.61)." Those using the Genevan criteria for developmental level find two groups of children, those who respond to training, and those who do not (Inhelder et al., 1974; Miller, S., 1976; Larsen, 1977). Piaget's (1971b) theory would predict that there would be a group of children who would not respond to training, because they were not in transition. Some of those using Brainerd's (1977) criteria find that a statistically significant number of children respond to training. It is highly likely that this group includes two groups of children, those in transition and those who were falsely identified as being non-conservers using judgments only criteria, who would be identified as two separate groups using the Genevan criteria. Construct validity requires that the pre-test measures of "cognitive developmental level" be measures of the hypothetical construct as defined by Piaget (1960, 1964, 1971a, 1971b, 1972). The theoretical construct to be measured is transitional status. The question is does Brainerd's (1977) Judgments-Only measure transitional status better than the Genevan method. From the point of view of psychometric theory and research the Genevan method would appear to be a better measure because it takes a larger sample of the relevant behavior. According to psychometric theory and research the more items (or behaviors) in a measure that measures the construct in question the higher validity will be (Campbell 1957, Fishbein and Ajzen, 1974; Green, 1978; and Epstein, 1980; Messick, 1980). It cannot be said that Brainerd (1977) meets this criteria as adequately as the Genevan school from a psychometric point of view. Further, since children in the age range used in the training studies can express physical relationships verbally that they do not understand (Bloom, 1975), then lack of consistency or confidence as indicated by judgments, explanations and responses to challenge in verbal responses meets the requirements of construct validity insofar as the relationship between language and cognition does not confound the determination of cognitive level by verbal means. It may require some ingenuity to scale explanations and responses to challenge so that they can be statistically analyzed but it has been done before in the development of creativity tests (Razik, 1970).

Another group of researchers have attempted to explain away cognitive developmental level taking another approach. "According to this view what the child does (e.g., his judgment on a conservation task) is determined by what he attends to (i.e., by what stimulates him) (Larsen, 1977, p. 1164)." This "simplification" approach as Larsen (1977) calls it has some of the same problems as the Brainerd (1977) and Genevan approach in that conservation judgments and/or explanations (verbal) are used to define the transitional state (Miller, S., 1976).

Clearly, a child must attend to a relevant stimulus dimension to notice a discrepancy with his schema if he is to accommodate his cognitive schema (Piaget, 1952, 1962). Indeed, cognitive development studies have utilized methods of inducing attention to relevant stimulus dimensions in efforts to train children on a more advanced schema than they had on pre-test (Gelman, 1969, 1972; Miller, P., 1973; Henry, 1976).

Henry (1976) assessed independently cardinal-ordinal abilities and attentional preferences of kindergarten conservers and non-conservers, and third grade conservers of number. He used judgments and explanations to eliminate transitional subjects and to define conservers and non-conservers. Henry (1976) found: "...that attentional preferences for number, as measured in this study (i.e., preference independent of situations which specifically call for numerosity judgments) are preceded by the acquisition of number conservation (p.756)." Lawson, Baron and Siegal (1974) found similar results in that their non-conserving subjects responded to number infrequently and significantly less often than did conservers even when specifically asked about number. Pick, Frankel, and Hess (1975) in a review of the determinants of children's attention say: "However, fundamental to any investigation of attention is the idea of selectivity in some aspect of cognitive functioning-selective perception, selective memory, even selective thought (p.370)." "We also find that there are important changes in age in the way children regulate and direct their own attention. One developmental trend is toward greater effeciency in strategies of search and exploration (p.370)." Or, as Piaget (1952, 1962) would say, his attention becomes decentered, and his is able to attend to two dimensions simultaneously when making a judgment (e.g., height and width for volume conservation).

In light of the above it would appear reasonable to conclude that the present evidence seems to indicate that attention to relevant stimulus dimensions is determined by the level of cognitive development. Again, developmental level and attention are measured in all these studies by verbal responses of children and as such are likely also to reflect the imperfect relationship between language development and cognitive development.

None of the investigators into the role attention plays in equilibration have used psychophysiological measures of attention despite the extensive body of research and theory on the psychological and neurophysiological determinants of attention (Lynn, 1966; Berlyne, 1960; Lindsley, 1970; Gardner, 1974; Kinsbourne, 1970, 1973, 1977; Milner, 1970; Kahneman, 1972; Wolff, 1965; Teyler, 1975).

In conclusion, this review of the research literature has found that equlibration is operationally defined as a child's verbal conflict response, which imperfectly reflects cognitive conflict because language development is not isomorphic with cognitive development. This leads inevitably to the conclusion that judgments, explanations, and response to challenges be used as a more reliable and valid measure of equilibration on the verbal level than judgments alone. Further, cognitive development level appears to determine how many and which stimulus dimensions a child will attend to in a conservation task as Piaget (1952, 1962) has predicted.

1c. Arousal Proposed as a Measure of Equilibration

An examination of Piaget's (1971) definition of equilibration finds that he includes more in his construct than the present operational definitions of it. Equilibration is a motivational process on at least two levels of organization of the organism. On the cognitive level it operates as a "need for internal consistency or logical consistency," and on the neurophysiological level it utilizes energy. A "need for logical consistency" is not easily measured operationally as it is not observable or at least it would be hard to get agreement as to an operational definition. One could conclude that the literature reviewed in this paper has so little agreement because they are attempting to measure the "need for logical consistency."

Equilibration is clearly a hypothetical construct, i.e., it involves the supposition of entities or processes not among the observed (MacCorquodale and Meehl, 1956). The process of validating a hypothetical construct is by the building of a nomological net (Cronbach and Meehl, 1960). A nomological net is an interlocking set of laws which constitute a theory. The laws may relate "... (a) observable properties or quantities to each other; or (b) theoretical constructs to observables; or (c) different theoretical constructs to one another." Learning more about "a theoretical construct is a matter of elaborating the nomological network in which it occurs, or of increasing the definiteness of the components." "We can say that "operations" which are qualitatively different "overlap" or "measure the same thing" if their positions in the nomological net tie them to the same construct variable (Cronbach and Meehl, 1966, p.77)."

It follows from the above that adequate measurement of a hypothetical construct depends upon measuring more than one of its defining attributes (Cronbach and Meehl, 1966; Campbell and Fiske, 1966). Given that the usual measure of cognitive equilibration, verbal responses is based upon two non-isomorphic developmental systems it is proposed that measurement of arousal changes might better reflect equilibration. The justification for using arousal changes is based upon (b) above, and the statement that we can say that qualitatively different operations measure the same thing if their positions in the nomological net tie them to the same construct variable (Cronbach and Meehl, 1966). Piaget's (1971a) definition of equilibration as a biological autoregulatory process clearly implies that the process of equilibration when operating on the cognitive level utilizes energy at the neurophsyiological level (Piaget, 1971a, Mischel, 1971, Flavell, 1963).

It also follows from Piaget's (1971a) definition of the dual processes subsumed under equilibration, assimilation and accommodation involve the utilization of different amounts of energy. Accommodation is when external elements are integrated (assimilated) with previous structures which are modified to a greater or lesser degree. Assimilation occurs when external elements are integrated with previous structures which remain unaffected. Accommodation involves restructuring so that the external elements can be assimilated (Piaget, 1971a). Therefore, it appears that accommodation, involving both restructuring and assimilation would utilize more energy than assimilation alone.

There are further grounds to believe that the theoretical constructs of psychophysiological arousal and equilibration should logically be related to each other. ([c] Cronbach and Meehl, 1966). These grounds come from the theory and research that defines the nomological net of the hypothetical construct arousal or activation. Arousal or activation is a hypothetical construct whose construct validity is based upon an extensive nomological network of empirical studies relating arousal to behavior, and arousal measures to the functioning of the reticulo-thalamo-cortical (RTC) brain system (Lindsley, 1970; Lynn, 1966; Malmo, 1959; Duffy, 1962; Berlyne, 1960; Lacey, 1967; Venables and Christie, 1973; Walsh and Cummins, 1975).

Changes in arousal as measured by peripheral or central nervous system measures of RTC activity have an ubiquitous relationship to stimulus change. Any change in a stimulus dimension results in a change in these measures, if the subject attends to the change (Milner, 1970;Lynn, 1966). It is for this reason that psychophysiological arousal is conceptualized as reflecting energizing of the organism (Lindsley, 1970; Woodworth and Schlossberg, 1954; Walsh and Cummins, 1975).

The relationship between attention and disinhibition of the reticular formation by the cortex is accounted for by the structure and function of the RTC (Brazier, 1968; Lindsley, 1970; Lynn, 1966; Netter, 1962). Theorists attempting to account for inhibition and disinhibition of the "orientation reaction" all accept the role of the "cortical analyzer in evoking or maintaining inhibition of the reticular formation (Gastaut, 1957; Grastyan, 1959; Hernandez-Peon, Sharpless and Jasper, 1956; Jouvet, 1961; Moruzzi, 1960; Sokolov, 1960). These theorists differ only as regards the subordinate role of lower inhibitory centers in the "orientation reaction" (Lynn, 1966). The "cortical analyzer's" role is to select the stimuli that will or will not be attended to, and thus will or will not evoke activation of the reticular formation (Lynn, 1966; Sokolov, 1960). Therefore, cortical or cognitive mediation determines whether a subject will behave, as well as exhibit a change in level of arousal (Beatty and Wagoner, 1978; Berlyne, 1960; Brodsky and Brodsky, 1975; Davidson, 1978; Hale and Strickland, 1976; Harris and Katkin, 1975; Lindsley, 1970; Lynn, 1966; Mandler, 1975; Schacter, 1964; Schacter and Singer, 1962; Sokolov, 1960; Valins, 1970; Zelazo, 1972). Even in neonates it has been found that swaddled infants give the same degree of arousal response to a change in intensity of a stimulus as when they are free to be active (Lipton, Stoneschneider and Richmond, 1960). Lipton et al. (1960) concluded from their study that both arousal and activity in neonates are controlled by a central process. In a recent study of the relationship between activation and cognitive difficulty of the task in adults Beatty and Wagoner (1978) found that hierarchically organized cognitive processing varies to the degree to which CNS activation is mobilized during execution. Or, in other words, arousal level varied according to task difficulty. Different levels of arousal occurred depending upon the amount of higher order cognitive processing the task required.

Therefore, given the structure and function of the RTC and Piaget's definition of equilibration as autoregulation, utilizing energy; and his differential definition of accommodation and assimilation which implies differential utilization of energy, it is hypothesized that peripheral and central measures of arousal should differentially reflect accommodation and assimilation during a cognitive task even when verbal measures do not reflect said processes.

The overall hypothesis of the present study is that the motivational aspect of cognitive equilibration will be differentially reflected by changes in psychphysiological arousal. If the above analysis of the relationship between the two hypothetical constructs, equilibration and arousal, is correct then accommodation should be reflected by a greater increase in arousal than would be reflected when assimilation occurs alone.

2. Hypotheses

Piaget's (1964, 1972) theory predicts that subjects in transition, (in the process of reorganizing their schema re: a concept) would accommodate their schema given training that presents a conflict between their schema and reality (Brainerd, 1973a, 1973b, 1974, 1977; Strauss, 1970; Strauss and Langer, 1972; Inhelder, et al., 1974; Larsen, 1977).

The present study was designed to test whether an increase in arousal reflects accommodation by transitional subjects and minimal or no increase in arousal reflects assimilation by non-transitional subjects, when both groups of subjects are given training that presents them with a conflict between their schema and reality. The criterion for whether a child was transitional or non-transitional was his or her performance on post-tests of the concept.

The study was also designed to determine whether the Brainerd (1977) criteria or the Genevan (Inhelder et al., 1974) criteria for transition, non-transition (developmental level) were more predictive of subjects' performance on post-tests of a concept.

Hypothesis 1: Subjects who perform as conservers on post-tests of number conservation after training will have significantly higher arousal scores during pre-testing, than subjects who do not benefit from training.

Hypothesis 2: Transitional subjects selected by Genevan criteria will have a significantly higher arousal level during the pre-testing than Genevan non-transitional subjects. Both the transitional and non-transitional subjects selected by Brainerd's criteria will be lowest in arousal.

Hypothesis 3: Subjects identified as transitional by the Genevan criteria and the Brainerd criteria should have higher arousal scores than Brainerd's non-transitional subjects followed by the Genevan non-transitional subjects who should have the lowest arousal scores.

Hypothesis 4: During pre-testing there will be a higher degree of correlation between the developmental level scores of the Genevan selected subjects and their arousal scores than a correlation between the judgment scores of the Brainerd selected subjects and their arousal scores.

Hypothesis 5: The number of false positives and false negatives identified by the three measures of disequilibrium will be lowest for the arousal scores followed by the Genevan criteria. The Brainerd criteria will result in the highest number of false negatives and false positives.

2a. Additional Hypotheses

Wilder's Law of Initial Values (LIV): It is possible within this study to test two hypotheses proposed by Brodsky and Brodsky (1978). That the GSR (Galvanic Skin Response) did not conform to Wilder's Law of Initial Values (LIV). "The LIV, an empirical-statistical rule, claims that the following is a general rule for the quantitative rule of response and stimulus. Given a standard dose of stimulus and a standard period of measurement the response that is the change from the initial (pre-stimulus) level, will tend to be smaller when the initial value (IV) is higher (p.1211)." Or, in other words, the correlation between GSR (deflections) and base level should be negative and large.

Brodsky and Brodsky (1978) also hypothesized that GSR scoring methods which correct for base level would, in any given experiment, increase the probability of accepting the null hypothesis when it is, in fact, false. If the GSR is not correlated with base level, then including a correction for the presumed influence of base level would in effect be adding error variance and variance due to individual differences to the variance due to the experimental effect (Brodsky and Brodsky, 1978).

Hypothesis 1a: The correlations between base level resistance and the GSR (deflections) resistance during class inclusion, seriation and number conservation pre-tests will be small and not significant.

Hypothesis 2a : The correction of change in GSR scores for the presumed relationship of change to base level (LIV) increases the probability of accepting the null hypothesis in any given experiment.

CHAPTER II

Method

The experiment was composed of three sessions, pre-testing, training, and post-tests. Post-testing was done two weeks after training. The pre-tests were Piagetian tests of cognitive developmental level; class inclusion, seriation, and number conservation. They were adapted from Inhelder et al. (1974). Each subject's GSR was monitored during the pre-tests.

The Galvanic Skin Reponse (GSR) was used as the measure of arousal. The operational definition of a specific measure of arousal within a specific experimental context depends upon the degree of correlation that measure has with the nomological network defining the hypothetical construct arousal. In the case of the GSR this correlation has been found by a number of investigators to be better than other measures of psychophysiological arousal (Lacey and Lacey, 1962; Lacey, Kagan, Lacey and Moss, 1963; Lacey, 1967; Lindsley, 1951; Woodworth & Schlosberg, 1954; Lynn, 1966).

During the second session all the subjects were trained on a conservation of number task adapted from Halford and Fullerton (1970) which they used in a study which resulted in a significant number of subjects achieving number conservation on post-test. The training method was a prediction-outcome method of inducing discrimination or cognitive conflict according to Brainerd's (1973) criteria for training methods.

Two weeks after training ended three post-tests were given, two for number conservation and one for class inclusion. One of the number post-tests was the conservation of number test given during the pre-testing, and the other was adapted from the number post-test used by

Halford and Fullerton (1970). The post-test for class inclusion was included because none of the 57 subjects passed the pre-test for class inclusion. Langer's (1969) class inclusion task was used. Performance on the post-tests was scored using the Genevan criteria.

1. Subjects

Fifty-seven nursery school and kindergarten children[1] were used as subjects. Parental permission was obtained for all the children in the nursery and kindergarten. Children, who indicated they did not wish to serve as subjects, were dropped. Nine nursery and three kindergarten children refused. One child was dropped because of prolonged absence from school due to illness.

The mean age was 5 years old; the standard deviation was 5.39 and the range 22 months (4 years and 3 months to 6 years and 1 month).

There were 29 boys and 28 girls. None of the analyses found any differences due to sex.

All 57 subjects were in all parts of the study. Six subjects who had a perfect score on the pre-test for seriation and number, and perfect scores on both post-tests for number were not used in the statistical analyses of the results of this study as they were classified as having attained number conservation before the study began.

All the subjects were given both the Brainerd (1974) selection criteria and the Genevan selection criteria (Inhelder et al., 1974).

The Genevan selection criteria has three levels that are administered to the child: judgments, explanations, and counter-suggestion

---

[1]The children attended Talmud Torah School in Winnipeg, Manitoba.

(Inhelder et al., 1974). Brainerd's (1974) selection criteria uses only the child's response to the first level of the Genevan criteria, judgments. Thus, the Genevan method contains within it the Brainerd selection criteria.

There were four female experimenters who served during the course of this study. One experimenter did all the pre-testing, post-testing; and trained half the subjects. While the first experimenter did the pre-testing another experimenter ran the psychogalvonometer. Two experimenters served in this capacity. The fourth experimenter trained the other half of the subjects. The two-week spring break made it necessary to use this last experimenter when the first half of the subjects were ready for post-testing. There were no differences in the training performances of the two groups who were trained by different experimenters.

2. Pre-Tests

2a. Apparatus

Class inclusion: A bunch of ten yellow daisies and three red roses [artificial flowers] (Inhelder et al., 1974).

Seriation: Ten little sticks ranging from 16 to 10.6 cm in length each differing in length from the next by 0.6 cm. and a screen (Inhelder et al., 1974).

Elementary number conservation: 10 red chips and 10 black chips (Inhelder et al., 1974).

A portable Lafayette Psychogalvonometer (model 7609A) was used to continually record the subjects' GSR while they were performing the pre-tests. The electrode used was a Palmer-Dorsal electrode with two cups, one at either end. The cups have a silver electrode at the bottom. The cups were filled with Grass Electrode paste which is non-toxic and is in common use in all hospitals. The electrode can be attached to the child's non-preferred hand so that the cups are in contact with the top and palm of the hand. This arrangement is comfortable for the child and allows freedom of movement.

2b. Procedure

The experiment was run in an empty classroom at the school. The subject and the experimenter sat across from each other at a table. The table and chairs were a size appropriate for four or five year old children. Only the materials for the specific skill being dealt with were on the table at any one time. The remaining materials were kept out of sight. The psychogalvonometer was on a table next to the table the subject was seated at.

The GSR electrode and equipment was explained to the child before the electrode was placed on the child's non-preferred hand. Handedness was tested by "Which hand do you color with?" Before beginning with the session proper, a resting base level was obtained, a parallel or declining recording. The GSR was continuously recorded. An event marker on the psychogalvonometer recorded when the subject was responding to the examiner's questions or statements.

Pre-tests: (Adapted from Inhelder et al., 1974). Class Inclusion: The experimenter named the flowers. To make sure the child knew the class name she asked: "Are the daisies flowers? . . . Are the roses flowers? . . . Do you know the names of some other flowers? . . . which ones?" Then the task began.

Trial 1. The experimenter made a bunch out of the flowers and asked: "Are there more daisies or more flowers in this bunch?" The child gave a judgment response. The experimenter asked for an explanation: "How do you know?" The child gave an explanation, if necessary the experimenter asked: "More than what?" Then a countersuggestion was given to the child's original judgment statement.

Trial 2. The experimenter put all the flowers together and said: "I am going to make up a bunch with all the daisies and you are going to make a bunch with all the flowers. Who will have the bigger bunch? The child responded with a judgment response. The experimenter asked for an explanation "How do you know?" The child gave an explanation. Then a countersuggestion was given to the child's original judgment statement.

Seriation: The experimenter asked the child, "Are all the sticks the same size or are they different sizes?" To make sure that the child knew that the sticks were of different sizes he or she was asked: "Hand me the biggest (or smallest) stick." Then the task began.

Trial 1: The experimenter asked the child, "Can you make a pretty staircase with all these sticks by putting them in order like this?" The experimenter put three in order. The child gave a judgment response. The experimenter then told the child, "you finish the staircase." After the child finished the experimenter asked, "Why is that the right order for a staircase?" The child gave an explanation, if the child had put them in the right order and had given a seriation response the experimenter moved a stick in the order and made a counter-suggestion simultaneously. "If I move this stick here won't the staircase be a better staircase?"

Trial 2. The experimenter placed a cardboard screen between herself and the child and said: "This time I am going to make a staircase, hand me the sticks one by one . . . in the right order for me to make a staircase." After doing this the child was asked for a judgment: "Is this the right order for a staircase?" and then an explanation, "Why is this the right order?" If the child had given a correct seriation explanation a counter-suggestion was posed: The experimenter removed the screen showing the staircase and said, "If I move this stick here wouldn't it make a better staircase?" while moving the stick out of order.

Elementary Number Conservation: The experiment laid out two rows of chips, one row of seven black chips and one row of seven red chips, equally spaced. The experimenter asked the child, "Are there as many black chips in this row as there are red chips in this row?", pointing to the two rows as she spoke. If the child said, "no" then the experimenter, if necessary, paired the chips one by one making sure the chils appreciated the equivalence of the two rows. Then the task began.

Trial 1. Preserving the layout of the two equally spaced rows the experimenter modified the layout by increasing the spaces between the chips of one row so that they formed a longer row. The experimenter then asked, "Are there as many . . . the same number of black chips as red chips now . . . or are there more?" The child gave a judgment response. The experimenter asked for an explanation. "How do you know they are the same? or "How do you know there are more (or less)?" Then a countersuggestion was given to the child's original judgment statement.

Trial 2. The experimenter collected all the chips and made two equally spaced circles, one of black chips and one of red chips. She asked, "Are there as many red chips in this circle as black chips in this circle?

pointing to each circle as she spoke. If the child said "no" the experimenter paired off the chips as in the previous task until the child appreciated the equivalence of the two circles. The experimenter then moved the chips in one of the circles closer together and asked, "Are there as many . . . the same number of black chips as red chips now . . . or are there more? The child gave a judgment response. The experimenter asked for an explanation, "How do you know they are the same?" or "How do you know there are more (or less)?" Then a countersuggestion was given to the child's original judgment statement.

The experimenter recorded all the subject's answers on a protocol sheet that had on it the questions to be answered and a space for the child's answer.

2c. Scoring

All the scoring of the pre-tests and the psychogalvanometer tapes were done after all the children were post-tested.

The psychogalvanometer tapes were scored first. Two scoring methods were used: the Arousal Interval Scale (AIS) developed by Brodsky and Brodsky (1976), and conductance scores, recommended by Venables and Christie (1973).

The AIS score did not correct for the hypothetical correlation of GSR (deflections) to base level (LIV). The theoretical and empirical rationale for the AIS, and the procedure for correcting deflections on the recording tape to AIS scores is to be found in Brodsky and Brodsky (1978). Three AIS scores were recorded for class inclusion, seriation, and number.

Reliability coefficients have not been computed for the Arousal Interval Scale (AIS), but it is a highly valid scale in that it has been significantly related to the particular behavioral criterion in every study that has used it (Brodsky and Brodsky, 1978). Therefore, it can be inferred that the scale is reliable because an unreliable measure cannot be highly correlated with a criterion (Messick, 1980).

The conductance scores were corrected for the hypothetical correlation of GSR (deflections). The recorded deflections on the tape were converted to ohms/resistance. This was done with a table provided with the Lafayette psychogalvanometer. Conductance is the reciprocal of resistance. All resistances were converted to conductances. The following formula was used to correct each conductance score for each subject, for each pre-test. (Three log conductance scores.)

$\log_e \left(\frac{X_i}{\text{base}} \right) + 10.$ X_i represents the number of ohms recorded for the particular pre-test, for a single subject. Base, is the base level also in ohms for the same subject. The constant 10 was added to remove negative numbers from the distribution (Woodworth and Schlossberg, 1954; Lykken and Venables, 1971; Venables and Christie (1973). Reliability coefficients are not given for the conductance scale (Venables and Christie, 1973). Nor does the conductance scale have high validity (Brodsky and Brodsky, 1978).

Pre-test protocols were scored first according to Brainerd's (1974) method. Correct judgments received a 1, and incorrect judgments received a zero. A subject could thus receive a top score of 2 for each pre-test (two trials per pre-test), and a total possible score of 6 for three pre-tests.

The pre-tests were then scored using an adaptation of Cooley's (1978) Developmental Level Scale. A scoring method for the Genevan selection criteria; judgment, explanation, and countersuggestion. In this study a subject was given a 1 for a correct judgment or a zero for an incorrect judgment; a 1 for a conservation explanation or a zero for a non-conservation explanation; a 1 for a conservation judgment not changed under countersuggestion and a zero for a non-conservation judgment not changed under countersuggestion. Conversely, a zero for a conservation judgment changed under countersuggestion and a 1 for a non-conservation judgment which changed to a conservation judgment under countersuggestion.

Cooley (1978) used percentages whereas this study used raw scores. Conservation explanations were considered inadequate if they failed to meet any of the following criteria adapted from Forsberg (1973).

Criteria for scoring explanations. Conservation explanations were considered adequate if they met one or more of the following criteria:

1. Compensatory relations: the subject states that changes in certain dimensions are compensated for by changes in other dimensions.

2. Reversibility: the subject states that the transformation would be cancelled by an inverse transformation.

3. Addition/subtraction: the subject states that nothing has been added or taken away.

4. Identical action: the subject states that the standard object could be transformed in a similar manner to the comparison object.

5. Initial equality and/or irrelevant transformation: the subject states that the two objects were initially equal and/or that the transformation makes no difference to the property in question (these two explanations generally occurred together).

6. Logical necessity: the subject states a general rule.

7. Quantitative equivalence based on another property.

The following conservation explanations were considered inadequate:

8. Irrelevant: included in this category are tautologies, unclear responses, and irrelevant remarks.

9. No answer or "don't know."

Thus, the total possible score for each pre-test was 6 (3 for each trial) and 18 was the total possible score for all three pre-tests.

Despite Brainerd's (1973a, b; 1977) extensive psychometric analysis of verbal measures of transitional status, he at no time reportes reliability coefficients for his judgments-only measure. Further, there are no reliability coefficients listed for the Judgments-only measure in the "Tests and Measurements in Child Development: Handbook II" (Johnson, 1976). On the other hand, a number of reliability coefficients are given for Genevan measures of transitional status. The split-half reliabilities, i.e., the internal consistency of the content, range from .63 to .96 (Johnson, 1976).

The lowest split-half reliability, .63, was reported for class inclusion content. The split-half reliability coefficients for seriation and number range from .80 to .96 (Johnson, 1976). The only interscorer reliability reported was for number tasks and that was .97 (Johnson, 1976). Cooley (1978) reports interrater reliabilities for the Development Level Scale used in this study ranging from .83% to 100% agreement of two raters. It was, therefore, thought to be unnecessary to use more than one rater in this study.

As was discussed in the introduction there is no agreement in the literature as to the construct and criterion validity of these measures of transitional status. This study was designed to determine which pre-test measure bests diagnosis transitional status and best predicts post-test performance after training. Prediction and diagnosis are the two necessary characteristics for criterion validity as discussed by Messick (1980). Using Messick's (1980) analysis of characteristics of validity, this study is also designed to investigate the following aspects of construct validity, convergent coherence, discriminate distinctiveness and nomological relatedness. The Genevan method clearly has been found to be valid in terms of another of Messick's (1980) characteristics of construct validity, namely, population generalizability. The Piagetian developmental stages as measured by the Genevan method have been verified cross-culturally (Dasen, 1972).

3. Training

3a. Apparatus

The apparatus was adapted from Halford and Fullerton (1970). There were 8 cardboard sheets, 60.96 x 50.8 cm. on which were placed either five or eight rectangular pieces of felt, 2.5 x 3.5 cm. These felt rectangles represented beds and were evenly spaced across the top third of the sheet. There were five beds on seven of the sheets, and eight on the remaining one. On each bed was a piece of colored cardboard, with a picture of a doll on it. This was the standard set of dolls. Each doll was backed with velcro, enabling it to be stuck to the beds or stuck to velcro in a row beneath the beds.

On each sheet the spacing of the beds and the standard set of dolls was varied in the following way. The spacing of the dolls was either wider, less than, or equal to, the spacing of the felt beds. A straight line below the first row of velcro separated the first row of velcro from five other rows of velcro. Five test series of dolls occupied these rows. At least one set of the test dolls was equal in number, length, and spacing to the standard set of dolls on the particular card. The four other test series of dolls on a particular card were varied as to number, length, color and spacing from the standard set of dolls on that card. Thus, the standard set could be used as a cue to select the correct test series of dolls to match the beds.

The seven cards with five beds and the card with eight beds were varied as to the order of presentation on each of the five days of training. The color of the test set of dolls and the standard set of dolls on the eight bed card was also varied on each of the five days of training.

3b. Procedure

The procedure was adapted from Halford and Fullerton (1970). The training procedure took place in one individual session per day over five consecutive days. Each training session consisted of 8 cards. The training sheets were placed a suitable distance away from the child. The child was given the following instructions. "On top of the sheet in front of you are some beds and lots of sets of different colored dolls. Now some of the dolls are in the beds and some are not. What is the color of the dolls in the beds? (This was to ensure that the child correctly perceived the arrangement.) Now, I am going to take the dolls out of the beds and put them underneath. Now, I want you to tell me which of the other sets of dolls will fit into the beds so that every bed will have a doll in it and there will be no dolls left without a bed?"

The experimenter then pointed to each set of dolls in turn except the standard set, and asked: "Will this set of dolls fit into the beds so that the beds and dolls are the same?" In the case of a positive response the experimenter placed the dolls in the beds until the beds were filled or the dolls in the set chosen were exhausted. The experimenter then asked: "Do all the dolls have beds? Do all the beds have dolls?" In the case of an error, the experimenter pointed to the doll (or bed) which did not have a bed (or doll) and asked: "Has this doll (or bed) have a bed (or doll)?" The child was not told whether or not his response was correct.

3c. Scoring

There was one score per card, zero for selecting the incorrect set and 1 for selecting the correct set. The highest score possible per session was eight. The highest possible score at the end of the five sessions was 40. The training protocols were also scored after all the children were post-tested.

4. Post Tests

4a. Apparatus

Two tests of number conservation. One was a repeat of the number conservation task given on pre-test, i.e., seven (7) black chips and seven (7) red chips. The second test was 6 coins and a pool of candies. The second number post-test was adapted from Halford and Fullerton's (1970) post-tests.

The third post-test was another class inclusion test adapted from Langer (1969a). It consisted of 10 beads, three blue beads and seven red beads, lined up in an open shallow cardboard box.

4b. Procedure

Two weeks after the training session ended the subjects were post-tested for class inclusion and for number conservation. Class inclusion was tested first.

Class inclusion. The experimenter pointed to the beads and asked: "What are these?" If the child did not answer beads he or she was given the correct name. Then the experimenter said: "They are round beads. What color at these?" The experimenter pointed to the red beads. The experimenter asked the question again pointing to the blue beads. The children all knew the colors.

Trial 1. The experimenter then asked: "Are there more round beads or more red beads?" The child gave a judgment response. The experimenter asked for an explanation. "How do you know?" The child gave an explanation. If necessary the experimenter asked: "More than what?" Then a countersuggestion was given to the child's original judgment statement.

Trial 2. The experimenter then asked: "Which will make the longer necklace, the round beads or the red beads?" The child gave a judgment. The experimenter asked for an explanation, "How do you know?" The child gave an explanation. Then a countersuggestion was given to the child's original judgment statement.

Number Conservation 1

The subjects were post-tested for number conservation using the same procedure as the pre-test for number. Using the black and red chips: a judgment, an explanation, and a countersuggestion were all elicited and recorded.

Number Conservation 2

The second post-test was adapted from the Halford and Fullerton (1970) post-test. The experimenter placed six coins in a row and the candies in a pool in front of the subject. The experimenter asked the child to: "Pick the same number of candies out of the bunch and make a row of the same number of candies as there are coins." If the number of candies were incorrect the experimenter paired them off until the child appreciated the equivalence of the rows. Then the experimenter proceeded with the same procedure as in post-test one. The words candies and coins were substituted for black and red chips.

The children were allowed to take a piece of candy for themselves at the end of this post-test.

4c. Scoring

All the post-tests were scored according to the Developmental Level Scale previously described in the pre-test section. The highest total score possible for a single post-test was 6. The highest total score possible for all three post-tests was 18.

CHAPTER III

Results

There were five main hypotheses and two additional hypotheses tested in this study. Of the five main hypotheses, 1, 2, 3, and 5 were confirmed by the statistical analyses. Both the additional hypotheses were also confirmed.

1. Data Analyses

The hypotheses of this study were tested by four stepwise multiple regression analyses, canonical correlation analysis, by chi-square analyses and simple Pearson Product Moment correlations.

The stepwise multiple regression technique provides measures of the extent to which all the independent variables predict the dependent variables, and the extent to which each independent variable is a unique predictor of the dependent variables. It has a capacity to mirror complex variables (Cohen and Cohen, 1975).

In the analyses done for this study, the first independent variables to be introduced into the regression equations were age and sex then the remaining independent variables were selected in terms of which of the independent variables at each stage had the largest semi-partial correlation coefficient, and hence made the largest contribution to the multiple correlation (R^2). The stepwise procedure defined an order based upon the relative uniqueness of the independent variables in the study.

The stepwise multiple regression procedure produces standardized regression coefficients (Beta weights) for each independent variable. The Beta weights values are partial regression coefficients and represent the unique influence of each individual independent variable with the influence of all other independent variables removed. Given that the multiple correlation coefficient (R^2) is significant, then a statistically significant Beta weight associated with a specific independent variable means that the variance associated with that independent variable makes a statistically significant contribution in accounting for the dependent variable (Cohen and Cohen, 1975).

The stepwise multiple regression analyses were chosen because according to Cohen and Cohen (1975) it is the appropriate analysis when the research goal is primarily predictive. One of the goals of this study was to determine which pretest measure or measures best predicted the performance on the dependent variables training and the post-tests. Only those analyses which had both a significant multiple correlation R^2 and significant Beta weights are presented here.

The first analysis consisted of the individual pre-test scores for each measure. They were judgments only: class inclusion, (JOC), seriation (JOS), and number (JON). Developmental level: class inclusion (DLC), seriation (DLS), and number (DLN). Arousal interval scores: class inclusion (ARC), seriation (ARS), and number (ARN). Conductance scores: class inclusion (CONC), seriation (CONS), and number (CONN).

The dependent variable was the total score for each of the post-tests class inclusion, number 1, plus number 2 (POT). (Table 1). This analysis was done to determine which of all the independent variables predicted post-test performance on class inclusion and the two number post-tests. The results of this analysis are presented in Table 1.

As can be seen the only significant Beta weight was for the arousal interval scale scores for class inclusion. The multiple R for

for this analysis was .87 and significant beyond the .001 level. The analysis accounted for the most variance of any of the analyses done 76%.

The second analysis was done to determine which of the individual pre-test independent variables, as listed above, best predicted post-test performance on the post-test class inclusion task (POC). The results of this analysis are presented in Table 2.

As can be seen there are three significant Beta weights. Arousal interval scores for class inclusion (ARC) was significant beyond the .01 level whereas arousal interval scores for seriation ARS and the developmental level scores for class inclusion were significant beyond the .05 level. The multiple R, .727 was significant at beyond the .05 level and 53% of the variance was accounted for.

The third analysis was done to determine which of the total scores of the four pre-test measures best predicted the total score on the three post-tests: class inclusion, number test 1 and number test 2 (POT). The pre-test total scores used as the independent variables were: judgments-only seriation plus number (JOSN), judgments-only class inclusion, seriation plus number (JOT), Developmental level seriation plus number (DLSN), Developmental level class inclusion, seriation plus number (DLT). Arousal Interval Scale scores for seriation plus number (ARSN) Arousal Interval Scale scores for class inclusion, seriation plus number (ART), Conductance scores for seriation plus number (COSN) and Conductance scores for class inclusion, seriation and number (CONT). The results of this analysis are presented in Table 3.

As can be seen there is only one significant Beta weight Arousal

Interval scale total scores for class inclusion, seriation and number (ART). The multiple R is significant beyond the .001 level and 64% of the variance is accounted for.

None of the multiple regression analyses that used the two number posttests as the dependent variable contained significant Beta weights despite highly significant multiple correlations.

The fourth analysis was done to determine which of the total scores of the four pretest measures best predicted the total scores for training. The pretest total scores used as the independent variables were the same as those used in the third analysis. The results of this analysis are presented in Table 4.

As can be seen there are three significant Beta weights; Age, at beyond the .001 level, Developmental level scores for seriation plus number (DLSN) and Developmental level scores for class inclusion, seriation and number (DLT) were significant at beyond the .05 level. The multiple R, .642 was significant at beyond the .05 level and 41% of the variance was accounted for.

To further delineate the relationship between the independent variables, the pretest measures, and the dependent variables training and the three posttests a cannonical correlation analysis was done. For this analysis the dependent variables were posttests: class inclusion (POC), number test 1 (PON) number test 2 (POP), and training (TRAIN). The independent variables were judgments-only; class inclusion (JOC); seriation (JOS), number (JON), Developmental level: class inclusion (DLC) seriation (DLS), number (DLN); Arousal Interval Scale: class inclusion (ARC), seriation (ARS), number (ARN); and Conductance: class inclusion

(CONC), seriation (CONS), number (CONN).

This study had four dependent variables: Training; Posttest class inclusion, Posttest number 1 and posttest number 2. The stepwise multiple regression analyses could determine which pretest measure or measures best predicted either training or the individual posttests or the total posttest score. However, it could not determine whether the dependent variables were related to each other and/or whether the independent variables which were significantly related to the individual dependent variables were also related significantly to all or some of the dependent variables. Cannonical correlation analysis provides for a multiplicity of dependent variables as well as multiple independent variables (Cohen and Cohen, 1975; Tatsuoka, 1971). It identifies the components of one set of variables that are most highly related (linearily) to components of the other set of variables. Canonical correlation analysis also includes a measure of the extent to which variables within a set are redundant. It was expected that the pretest measures would be redundant because the Developmental level measure included within it the Judgment-only measure. Thus, it differed only in that it required an explanation and a counter suggestion as well as an initial judgment. Similarly, the Conductance measure and the Arousal Interval Scale measure differ only as to the mathematical treatment of the same psychophysiological response. The Conductance score corrects for base level resistance whereas the Arousal Interval Scale score does not. Canonical correlation analyses produces a set of redundancy coefficients which are the squared multiple correlations of each set with every other variable in the set. Thus, the extent to which all of the pretest measures are measuring the same thing can be evaluated (Cohen and Cohen, 1975).

If the learning theory explanation of cognitive development is correct the dependent measures should also be redundant. That is, presumably training for the number concept should be related to posttest performance of the number concept.

If Piaget's theory that only children who experience disequilibrium or show evidence of accommodation will respond to training then one would not expect training performance to be related highly to posttest number concept performance. If class inclusion is necessary for understanding of the number concept then the posttest class inclusion test should be highly related to the number tests. Table 5 presents the redundancy coefficients for each independent variable with every other independent variable in the set. It also presents the redundancy coefficients of each dependent variable with every other dependent variable in the set.

As is apparent the pretest measures are highly redundant; the R squares range from .80 to .915. However, the dependent variables present a different picture. Training shares very little variance with either of the number tests which have the highest R squares. Class inclusion is essentially unrelated to the other three dependent variables.

The simple correlations between the independent variables expected to be redundant support the results reported in Table 6. Table 6 gives the simple correlation coefficients for the verbal pretest measures Developmental level and Judgment-only with each other for the individual pretests and for the total scores. It also gives the correlations for the two arousal measures; Conductance and Arousal Interval

Interval scores, with each other. The correlations for the verbal measures range from .89 to .93 whereas the arousal measures range from .56 to .64. Table 7 reports the correlations of the pretest measures and the posttest measures with training. It also presents the correlations of the three posttests with each other. The only correlation that is high enough to be meaningful is that of the two number posttests (number 1 and number 2 with each other, .85. The rest of the significant correlations here are so low that they share very little variance with training.

Canonical correlation analysis also yields pairs of canonical variates for each set. The first pair represents the linear composite of the independent variables which are maximally correlated with a linear composite of the dependent variables. The subsequent pairs are a linear composite of the residual left after each variate has been extracted. The subsequent variates account for less and less of the shared variance and are othogonal to each other and the first canonical variate pair (Cohen and Cohen, 1975).

The analysis yielded two significant canonical variates for each set (Table 8). The first pair of canonical variates (1) was highly significant, and accounts for 75% of the variance. The second pair of canonical variates (2) was significant and accounted for 47% of the variance. Canonical variates are best interpreted by means of the structure coefficients, which are zero order correlations of the variate with its constituent variables (Cohen and Cohen, 1975; Tatsuoka, 1971). The standardized coefficients (loadings) for each set of significant canonical variates are presented in Table 9. Table 10 contains the structure coefficients for each set of canonical variates. Comparing the structure coefficients for each independent and dependent variable (Table 10), with its corresponding canonical variate loading (Table 9), the first pair of canonical variates appears to have most of its variance accounted for on the independent variable side by the AR Arousal Interval Scale scores and the DL Developmental Level scores. (ARC) Interval Scale class inclusion was the largest structure coefficient of the arousal measures (Table 10), and the largest loading of all the independent variables (Table 9), appears to be accounting for more of the variance than what it apparently shares with Arousal Interval Scale (ARS and ARN) for seriation and number in this analysis. The simple correlation of Arousal Interval Scale score (ARC) with the total arousal Interval Scale scores (ART) is .91 and with the Arousal Interval Scale scores for seriation and number (ARSN), .82, whereas the correlations with Arousal Interval Scale scores (ARS and ARN) for number and seriation are .84 and .79, respectively.

Given the high correlation between Arousal Interval Scale scores for class inclusion (ARC and ART) the total Arousal Interval Scale scores for all three pretests (ART) are almost interchangeable. The difference between them could be accounted for by Arousal class inclusion (ARC) relationship with posttest class inclusion (POC) which is a member of the other half of this canonical variate pair. The other members of the dependent variables are the posttest for number test 1 (PON) and number test 2 (POP). Training is essentially independent of the first canonical variate pair. Indeed, the structure coefficients (Table 10) indicate that Training is part of the second pair of significant canonical variates

(Table 9). The dependent variables that appear to be accounting for most of the variance of the second pair of canonical variates, from an examination of the structure coefficients in Table 10 are training and posttest class inclusion (POC). The latter is negatively correlated with this variate. Post test class inclusions, lack of relationship with training is confirmed by the low simple correlation between it and training r = .018 and again by it's redundancy coefficient, .07 (Table 5). This leaves the relationship between (POP) the second number post test and TRAIN. The simple correlation here is .38. This side of the second canonical variate is essentially composed of training performances and its relationship to performance on the second number test (POP).

On the independent variable side the negative correlation among the structure coefficients (Table 10) can be accounted for by the negative or non-existent relationship between class inclusion measures and training. The correlation coefficients of training with developmental level class inclusion (DLC) and judgment-only class inclusion (JOC) are -.14 and -.06, respectively. If one considers coefficients below .30 as not significant, as is customary, this leaves Arousal Interval Scale number (ARN) with .486 and Conductance number (CONN) with .452 (Table 10) as accounting for most of the relationship with training. However, this relationship is not confirmed by either the correlation coefficients (Table 7), or the stepwise multiple regression analysis (Table 4). The Beta weights of the stepwise multiple regression analyses of the individual pretest measures with training as the dependent variable did not reach significance. Training is not highly correlated with either the pretests or the posttests (Table 7). Training also has a low redundancy coefficient with the other dependent variables (Table 5). The stepwise multiple regression analysis in Table 4 indicates that only 41% of the variance in training is accounted for by all the variables listed. Of the variables listed, only three reach significance: Age, DLSN, and DLT, of which age is clearly the most important.

Canonical correlation analysis maximizes the relationship between two sets of variables (Cohen and Cohen, 1975) and in this case in the absence of age as a variable in this analysis, it appears to have capitalized on the variance left. Thus, it is safest to conclude, that despite the significance reached by the second canonical variate pair, the relationship, though interpretable, is really not meaningful because of the small amount of variance involved. The stepwise multiple regression analysis and the simple correlations also do not confirm the variable relationships delineated by the second canonical variate pair.

The chi-square analyses were done to determine which pre-test measure had the lowest overall error in predicting which subjects would pass or fail the two number post-tests (PONP). According to Brainard (1977), "...The appropriate response criterion is the one with the lowest error rate (p.365)." The criterion used for the post-tests was a score of 10 or better out of a possible score of 12. Conservative cut off points were used for each pre-test. Figures 1, 2, and 3, present the scatter plots of each pre-test measure with the total score for the two number tests (PONP). The cut off points used for the chi-square analysis are indicated on the scatter plots. The results of the chi-square analyses are presented in Table 11. Table 11 also gives the number of subjects correctly identified and incorrectly identified as transitional and non-transitional, the error rates, the cut off points for the pre-tests and post-tests and the chi-squares for each pre-test's ability to correctly identify a subject.

The Arousal Interval Scale scores had the lowest overall error rate (24%), and their ability to discriminate transitional and non-transitional states was significant at beyond the .001 level. The developmental level scores had the next lowest error rate (31%) and the ability to discriminate was significant beyond the .01 level. The judgments-only scores had the highest error rate (41%) and the chi-square for discrimination of transitional and non-transitional states was not significant.

2. Arousal as a Measure of Accommodation and Assimilation

Each analysis consistently found that arousal as measured by the Arousal Interval Scale during pre-testing predicted subjects' post-test performance on the two number tests and post-test class inclusion. In the multiple regression analysis of the individual pre-test measures as the independent variables and the total scores of the post-tests only the unique variance associated with arousal during class inclusion pre-testing (the Beta weight) reached significance (Table 1). Again, in the multiple regression analyses with the individual pre-test measures as the independent variable and post-test class inclusion as the dependent variable, two arousal pre-test measures reached significance: arousal interval scores during pre-testing for seriation and class inclusion (Table 2). In the multiple regression analysis with the total scores for each pre-test measure as the independent variables and the total post-test scores as the dependent variable, only the total Arousal Interval scores for all three pre-tests had a significant Beta weight (Table 3). The chi-square analyses of the number of subjects correctly and incorrectly identified as transitional and non-transitional found that the total Arousal Interval Scale score during pre-testing discriminated at beyond the .001 level of significance, and the total Arousal Interval Scale score during pre-testing for seriation and number discriminated at beyond .01 level of significance, which subjects would and would not pass the post-tests for number (Table 11).

Finally, the results of the canonical correlation analyses are best summarized as the total Arousal Interval Scale scores during pre-testing and the total developmental level scores during pre-testing account for .866 of the variance in the relationship of the pre-tests to performance on the post-tests for number (Tables 8, 9, and 10). Figure 4 diagrams the relationship that best describes the first pair of canonical variates listed in Table 8 as significant beyond the .001 level with a canonical correlation of .866 and accounting for 75% of the total variance; as interpreted by assessing the loadings of the two sets of variables loading on the first canonical variate pair (Table 9) by each variable correlation with the canonical variate (Table 10).

Given these results the following hypotheses can be said to be confirmed by these results: Hypothesis 1, subjects who performed as conservers on post-tests of number conservation after training did have significantly higher arousal scores during pre-testing, than the subjects who did not benefit from training. Or, in other words, high levels of arousal reflected the equilibration process of accommodation of transitional subjects during pre-testing and low levels of arousal reflected the equilibration process of assimilation by non-transitional subjects during pre-testing thus predicting with a high level of accuracy which subjects would pass post-tests of number and which would not (Table 11).

However, it is the combination of the total developmental level scores with the total Arousal Interval scores during the pre-test which best predicts the performance of subjects on the two number post-tests. The first canonical variate pair is composed of just these variables and accounts for 75% of the variance in the relationship between pre-testing training and post-test number performance. Another way to account for the orthogonality of the developmental level scores and the Arousal Interval scores is that the arousal interval scores are extremely accurate in classifying transitional subjects whereas the developmental level scores tend to err on classifying transitional subjects as non-transitional(Table 11). They both have the same rate for classifying non-transitional subjects as transitional. Figures 1 and 2 give the scatter plots of the Arousal Interval scale total pre-test scores (ART) with the total scores for the two number post-tests and the total developmental level scores for the pre-tests with the two number post-tests. These scatter plots demonstrate why it is that the developmental level scores and Arousal Interval Scale scores are uncorrelated with each other but do combine together to give the best relationship to the post-tests of any of the other pre-test measures.

A number of children with low developmental level scores had very high Arousal Interval Scale scores and high post-test scores. Transitional subjects had higher Arousal Interval Scale scores whether or not they had high developmental level scores or judgments-only scores. Hypothesis 4 was not confirmed because of the orthogonality of the Genevan and arousal measure. Thus, there was essentially a zero order correlation between the Genevan developmental level scores, the Brainerd judgments-only scores and the Arousal Interval Scale scores.

Hypothesis 5. The number of false positives and false negatives (error rate) identified by the three measures of disequilibrium were lowest for the arousal scores followed by the Genevan criteria. The Brainerd criteria did result in the highest (error rate) false negatives and false positives. This hypothesis was clearly confirmed by the results summarized in Table 11. The chi-square analysis of subjects correctly and incorrectly identified by the three pre-test measures.

3. Class Inclusion

As previously mentioned the post-test class inclusion POC task was added to the post-testing during the running of the experiment because it was noted that even children who achieved a perfect score on the pre-tests, seriation and number, were not passing class inclusion. Of the six children who achieved perfect scores on seriation and number, as well as perfect scores on post-tests of number, who were not used in the statistical analysis because they were considered to have the number concept, all of them received a zero score on pre-test class inclusion. All but two received a zero score on post-test class inclusion. The various analyses performed confirmed that the pre-test verbal class inclusion measures were not correlated with the post-test number tests. The correlation of the pre-test development level class inclusion scores (DLC) plus the post-test class inclusion scores (POC) with the post-test scores for post-test numbers 1 and 2 (PONP) was .186.

Stepwise multiple regression analysis of the individual pre-test measures for class inclusion, seriation, and number were done with post-test class inclusion and total post-test scores as the independent variables (Tables 1 and 2). Table 2 shows that (ARC) arousal class inclusion, (ARS) arousal seriation, and (DLC) developmental level class inclusion account for a significant amount of the variance in the relationship between post-test class inclusion and the independent variables. The redundancy coefficient (Table 5) for post-test class inclusion (POC), .07, indicates that it is also not related to the post-tests of number.

However, arousal (ARC) scores during pre-testing for class inclusion are highly related to both post-test class inclusion and the total post-test scores (POT) (Tables 1 and 2). Arousal class inclusion (ARC) is also highly correlated with the total (ART) arousal score, r = .91, and with the total arousal score for seriation and number (ARSN), r = .82. Thus, the inclusion of the variance associated with arousal class inclusion (ARC) results in the total arousal score ART being a better predictor of the total post-tests scores (POT) than the total arousal score for the seriation and number (ARSN). The total arousal score (ART) classifies better than the total arousal score for seriation and number (ARSN) as well (Table 11). Indeed, the multiple R of the regression analysis of Table 1 is higher than that of Table 3. The multiple R for total pre-test scores against total post-test scores (Table 3) is .80 which is significant beyond the .01 level, whereas the multiple R for individual pre-test scores against total post-test scores (Table 1) is .87 which is significant beyond the .001 level. In both cases the arousal (AR) scores account for a significant amount of variance, but as it can be seen in Table 1, it is ARC alone which accounts for a significant amount of variance.

Therefore, it must be concluded that arousal increases during pre-testing for class inclusion are highly related to post-test performance on number and class inclusion, whereas performance on verbal measures of class inclusion pre-test or post-test are not related to post-test performance of number.

4. Training

If any of the independent variables predict training performance the most accurate ones found in this study are first, age, then, total developmental level score (DLT), followed by DLSN, the total score for seriation and number (Table 4).

A less comfortable conclusion to be drawn from these results is that training has very little effect upon post-test results for number. The simple correlations account for very little variance (Table 7). Both the redundancy coefficients and the structure coefficients (Tables 5 and 10) confirm the conclusion that training has very little effect upon post-test results for number. As for its effect upon post-tests results for class inclusion, training appears to have either no influence or a negative influence (Tables 7, 9, and 10).

5. Methodological Issues

It is apparent that the Genevan developmental level measure is a more accurate predictor of post-test performance than the Brainerd judgments-only measure despite the high correlations between the measures. The correlations are: class inclusion, .89; seriation, .89; number, .93; total scores for class inclusion, seriation, and number, .89; and total scores for seriation and number, .93. Despite the fact that these correlations are so high the developmental level scores classify transitional and non-transitional subjects significantly better than the judgments-only scores (Table 11). Since the judgments-only measure was the first question in the three item developmental level measure in this study it can be concluded in that it is the extra items in the developmental level measure, explanations and counter-suggestions which cause the developmental level measure to discriminate significantly better and have a lower overall error rate.

Brainerd (1977) presented a mathematical model "proving" that the false negative rate, for such a measure as the developmental level used in this study would be higher than the false positive error rate for judgments-only. In this study the false negative rate for developmental level is 6, the false positive rate for judgments-only is 18 (Table 11). Thus, in terms of overall error rate and Brainerd's (1977) mathematical hypothesis, the Genevan developmental level measure is the best verbal response criterion (Table 11).

Finally, it should be noted that the judgments-only measure is particularly inappropriate for class inclusion tests. Time and again subjects in this study gave the judgment that there were more flowers on pre-test and that there were more round beads on post-test thus receiving a correct judgment score. However, the explanations revealed that they equated daisies with flowers and they equated round with red beads. They had reduced the inclusive category to one of its sub-categories, rather than subsuming the sub-categories within the inclusive class. The judgments-only measure cannot reflect this phenomenon.

As was proposed previously it was possible within this study to test two additional methodological hypotheses related to the measurement of the galvanic skin response (GSR). Hypothesis 1a: was that correlations between base level resistance and GSR (deflections) resistance during class inclusion, seriation and number conservation pre-tests will be small and not significant. This hypothesis is a test of Wilder's Law of Initial Values which holds that GSR deflections are correlated highly with base level resistance. The correlations between base level (ohms resistance) and deflections (ohms resistance) for each pre-test were class inclusion, $r = .09$, seriation, $r = .07$, and number $r = .16$. The number of subjects was 57.

The second hypothesis (2a) that could be tested was that correction of change in GSR scores for the presumed relationship of change to base level would increase the probability of accepting the null hypothesis in a given experiment, because including a correction for base level in the change scores would essentially be adding error variance if the degree of change was not related to base level resistance. The conductance scores (CON) used in this experiment correct for base level resistance/conductance as recommended by Venables and Christie (1973).

These conductance scores (CON) did not reach significance level in any of the analyses done despite a reasonably good correlation with the arousal interval scores (Table 6). The arousal interval scores reached significance many times, in fact, the results indicate that arousal as measured by the Arousal Interval Scale accounts for more of the variance associated with post-test results than any other pre-test measure. Thus, it can be concluded that in this study the correlations with the base level of changes in resistance under stimulation were essentially zero, and that the conductance scores which corrected for base level resistance would have led one to accept the null hypothesis when it was false.

6. Summary

Four out of five of the major hypotheses of this study were confirmed by the results of the study. Two additional methodological hypotheses were confirmed.

The major hypothesis that increases in arousal would reflect accommodation by transitional subjects and no increase in arousal would reflect assimilation by non-transitional subjects was confirmed by all the analyses done. The related hypotheses that the Genevan developmental level (Inhelder et al., 1974) criteria for transition or non-transition would be more predictive of subjects post-test performance than the Brainerd (1977) criteria was also confirmed by the canonical correlation analyses and the chi-square analyses.

Two unanticipated results were that training was best predicted by Developmental level measures and age, and that arousal level during class inclusion best predicted performance on post-test class inclusion and the number post-tests despite the fact that none of the verbal pre-test class inclusion measures nor the post-test class inclusion performance were related to post-test number performance.

The two additional hypotheses concerned with the measurement of the GSR were also confirmed. The deflections (changes in ohms resistance) of the GSR correlations with base level resistance were esentially zero thus disconfirming Wilder's Law of Initial Values as being relevant for the measurement of GSR deflections. Finally, conductance scores which correct for base level did not reach significance in any of the data analyses whereas the Arousal Interval Scale scores which did not correct for base level accounted for significantly more variance than any other pre-test measure. Use of conductance scores as a measure of arousal would have led one to accept the null hypothesis in this study when it was false.

CHAPTER IV

Discussion

This study tested five hypotheses all of which were derived from a hypothetical integration of Piaget's (1971a) hypothetical construct, equilibration, and Arousal theory (Duffy, 1962; Lindsley, 1951; Malmo, 1959). Hypotheses 1, 2, 3, and 5 received strong confirmation from the results of the analyses performed.

This section will discuss the results in terms of the theoretical implications of the non-confirmed hypothesis, number 4, and the confirmed hypotheses. Additional findings that will be discussed in terms of their theoretical implications are the relationship of training and class inclusion to the post-test results.

1. Arousal as a Measure of Equilibration

Hypothesis 1: The results of this study provided strong confirmation that subjects who performed as conservers on post-tests of number conservation after training, did have significantly higher arousal scores during pre-testing than subjects who did not benefit from training. This hypothesis was deduced from Piaget's (1971a) definition of equilibration as a biological autoregulatory process, which clearly implied that the process of equilibration at the cognitive level utilizes energy at the neurophysiological level (Piaget, 1971a; Mischel, 1971; Flavell, 1963). It also followed that the dual processes subsumed under equilibration, assimilation, and accommodation, would involve the utilization of different amounts of energy. Piaget (1971a) defines accommodation as involving a restructuring of a schema so that external elements can be assimilated. Thus, accommodation would involve the utilization of more energy than assimilation.

By definition, subjects are understood to have accommodated their schemas, if they do not pass pre-tests for number (class inclusion, seriation and number), but do receive high scores on post-tests of number after training.

In this study, high pre-test arousal scores were highly associated with high scores on number post-tests. Conversely, low arousal scores during pre-test were associated with low scores on number post-tests. The arousal scores classified subjects who had accommodated (transitional subjects) and subjects who had assimilated (non-transitional subjects) with a lower error rate than either verbal measure of accommodation and assimilation. This finding replicates the results of a study done by Brodsky and Brodsky (1979). That study also found that the arousal measure predicted accommodation and assimilation at a highly significant level on a conflict inducing class inclusion task.

As further evidence that the increase in arousal recorded in this study reflected accommodation, and assimilation there were six subjects dropped from the statistical analysis in this study because they achieved perfect pre-test and post-test scores. These children would have been predicted to have shown little or no increase in arousal during pre-test because they would be assimilating the pre-test tasks to their schema for number invariance. Their mean Arousal Interval Scale scores for the three pre-tests was 4.33, range 2-6. The median Arousal Interval Scale score used to classify transitional and non-transitional subjects was 10.56. These subjects would have been classified as non-transitional which in fact they were. They already had the number schema in their repetoire.

Thus, it can be concluded that arousal as measured by the Arousal Interval Scale (AIS) while the subject is performing a Piagetian cognitive task will reflect to a high degree of accuracy, accommodation and assimilation as well as transitional and non-transitional status.

Hypothesis 4: The results which did not confirm hypothesis 4 are particularly relevant here. Hypothesis 4 proposed that there would be a high correlation between the Genevan developmental level scores and arousal scores. The results indicated that arousal was probably orthogonal to both verbal pre-test measures.

This can be interpreted to mean that the "cortical analyser" (Sokolov, 1960) controls both the arousal response and the verbal response. The verbal response does not have to reflect the cognitive recognition of a discrepancy between the subject's schema and the task requirements. However, if there is internal cognitive accommodation the arousal response must reflect it (Beatty and Wagoner, 1978; Bloom, 1970; 1975; Brodsky and Brodsky, 1978; Bruner, 1966; Hale and Strickland, 1976; Harris and Katkin, 1975; Inhelder et al., 1974; Inhelder and Piaget, 1964; Lynn, 1966; Mandler, 1975; Miller, S., 1976; Schacter, 1964; Schacter and Singer, 1962; Sokolov, 1960; Valins, 1970). The findings that the developmental level measure misidentifies more transitional subjects than the arousal measure supports this interpretation of the probable internal occurrences. Support also comes from the many studies that have found that arousal measures reflect discrimination even when the subject gives no behavioral evidence that he has made a discrimination (Beatty and Wagoner, 1978; Berlyne, 1960, 1965; Brodsky and Brodsky, 1979; Bruner and Postman, 1950; Hale and Strickland, 1976; Harris and Katkin, 1975; Lynn, 1966; Mandler, 1975).

Further support for this interpretation of the role of the "cortical analyser" is the finding that the accuracy of predicting post-test number scores is increased by using both the arousal measure and the Genevan developmental level measure. Arousal does very well on its own, but the combination of two of the hypothetical functions of the "cortical analyser" gives the best prediction rate for post-tests, even though these variables are apparently independent of each other. The performance on the post-tests was three weeks after pre-testing when the arousal response was recorded. The post-test was also two weeks after training. Therefore, the subjects had plenty of time and practice in reorganizating their schema for number. It is assumed with some justification, that post-test results are a more reliable and valid measure of the true state of a subject's knowledge than either pre-tests or training performance (Hilgard and Marquis, 1968). Therefore, from a traditional learning point of view, it could be said that the post-tests reflect more accurately the "true" state of the schemas of the "cortical analyser," and that is why both the arousal and developmental level pre-test measures correlate highly with post-tests, but not with each other.

The results also constitute considerable support for the construct validity of Piaget's (1971a) hypothetical construct, the equilibration process. According to Cronbach and Meehl (1960) validation of a hypothetical construct is done by the process of building a nomological net. They define a nomological net as an interlocking set of laws. The laws may relate "... (a) observable properties or quantities with each other; or (b) theoretical constructs to one another or (c) different theoretical constructs to one another (1966, p.77)."

(a) Three observable variables were found to be related to each other as predicted from the theoretical integration of Arousal theory and Piaget's (1971a) hypothetical construct, equilibration. Subjects' post-test scores on number tasks were highly related to subjects' developmental level scores and arousal (AIS) scores for the pre-tests.

(b) Piaget's (1971a) theoretical conceptualization of the relationship of cognitive motivation to energy utilization received confirmation from the results of this study. Subjects who failed the pre-tests and who had large increases in arousal during pre-tests were found to perform significantly better on post-tests than subjects who failed the pre-tests and evidenced little or no increase in arousal during pre-tests. Presumably, arousal increases reflecting accommodation resulted in the high arousal subjects' "learning" during training. Thus, there is support from these results that the "need for internal consistency or logical consistency (Mischel, 1971)," at the cognitive level results in the mobilization of energy for "learning."

(c) This study tested five hypotheses derived from a hypothetical integration of Piaget's (1971a) hypothetical construct, equlibration and Arousal theory [another hypothetical construct] (Duffy, 1962; Lindsley, 1951; Malmo, 1959). The hypothesized relationship of the two hypothetical constructs from different theories received strong support from the results of this study. Therefore, the results of this study can be said to meet all three of Cronbach and Meehl's (1966) criteria for construct validation.

2. Methodological Issues

Hypotheses 2 and 3 concerned the validity of the two verbal response measures of transitional status. They hypothesized a closer relationship between the more valid measure and arousal, and transitional status as determined by post-testing. The Genevan developmental level measure was hypothesized to be a more valid measure of transitional status than the Brainerd judgment-only measure. As stated previously, Arousal and developmental level were together the most accurate predictors of post-test scores. Also, as the results indicated, developmental level scores reached significance alone in their ability to predict post-test scores, whereas the judgments-only measure never reached significance. Using Brainerd's (1977) criteria for the best response measure, the overall error rate, i.e., the combined rate of false negatives and false positives, the developmental level scores were clearly better than judgments-only at classifying subjects as transitional and non-transitional. Brainerd (1977) also presented a mathematical hypothesis that the false negative rate, for such a measure as the developmental score used in this study, would be higher than the false positive error rate for judgments-only. In absolute numbers the false positive rate in this study for judgments-only was three times the false negative rate of the developmental level scores.

In this study the same subject was rated on both measures. As noted before, judgment is the first step of the Genevan method. The problem with the judgment-only method is that one question per task is not a large enough sample of the subjects' behavior. In this study the question requiring a judgment was asked twice for each task, and even then, the judgment-only scores were not reliable or valid. The developmental level method utilized six samples of verbal behavior on each task. From a psychometric point of view the larger the number of items in a test that samples a behavioral domain of interest, the greater the reliability of the test (Anastasi, 1968; Green, 1978). Validity in this case appears to turn on having a measure of the child's ability to explain his judgment and resist countersuggestion. It was the variance associated with explanations and countersuggestions that discriminated between transitional and non-transitional status significantly better than judgments-only by itself. During the experiment time and again the judgments were correct, but often the child either could not explain or give a correct explanation. Not a single subject gave an incorrect judgment with a correct explanation.

The developmental level scores do tend to classify more children as false negatives than either the arousal or the judgments-only measure, but the difference was not great. Also false negatives discriminate against the Genevan hypothesis, that only transitional subjects will respond to training, since false negatives are children who are in fact transitional. In particular this tendency to have greater false negatives would increase the probability of accepting the null hypothesis. In this study the Genevan developmental level measure did so well because it is extremely accurate in identifying non-transitional subjects.

The problem of which method is appropriate for assessing transitional status came about because of the controversy over whether children learn by accretion (the learning hypothesis) or whether they learn because they have the appropriate cognitive schemas to integrate the information presented to them (Brainerd, 1973a, 1973b, 1974, 1977; Inhelder et al., 1974; Larsen, 1977). Thus, not only are the methods different for assessing transitional status, but the criteria used for determining whether the subject "learned" or not are different (Brainerd, 1977; Gelman, 1969; Halford and Fullerton, 1970; Henry, 1976; Lawson et al., 1974). For example, this study used three post-tests and scored them with the developmental level measure. Studies which use judgment-only to assess transitional status typically use judgment-only to score post-tests, the criterion performance.

The question is which verbal response method is a more valid measure of the child's understanding of the number concept. Given the present dichtomy of studies in the literature that question cannot be answered. However, the results of this study can begin to answer that question. Three methods of assessing transitional status were used in this study. The most accurate measure was the arousal measure. Arousal level during pre-testing was highly predictive of post-test performance as measured by developmental level scores. If the developmental level measure was not accurately assessing the child's understanding of the number concept there is no reason why arousal level during pre-testing should be related to post-test results. Furthermore, the arousal level during pre-testing did not correlate with developmental level scores during pre-testing, but was correlated with developmental level scores during post-testing. Therefore, the level of arousal during pre-testing cannot be attributed to the differences between the two methods of assessment during pre-testing.

Finally, arousal level during pre-testing predicted post-test developmental level scores much more accurately than the pre-test developmental level scores. Thus, it can be concluded that the developmental level measure is a more valid measure of the child's level of knowledge both at pre-test and post-test, because arousal does not reflect the cognitive content of cortical activity; it reflects an increase in cortical activity, which signifies that the subject has made an orientation reaction response, the meaning of which can only be interpreted in terms of the content of the behavioral response and the stimulus situation.

The results of this study comparing three measures of transitional status strongly support the conclusion that of the two verbal measures Brainerd's judgment-only and the Genevan developmental level; the Genevan developmental level is a more reliable and valid measure of a child's level of understanding of number.

3. Training

The results of this study found that age accounted for most of the variance in training performance, followed by DLT, total developmental level score and then DLSN total developmental level score for seriation and number. However, none of these variables accounted for very much variance as the highest simple correlation was .43 for DLSN with training. Furthermore, training performance was found to be essentially unrelated to post-test performance on the number post-tests.

It would be difficult to interpret the task as not teaching number, as it is tested by Piagetian tasks for number conservation. Indeed, half the subjects were asked at the end of the training, how they were able to pick the right set of dolls. Every child who had achieved an errorless performance on training by the last day answered by saying "you count." The children who did not achieve an errorless performance usually gave as an answer, "You have to find the dolls that fit in the beds;" a variation of the training directions. The only exception was one little girl, who told the experimenter that to get them right you must count, but to count was cheating! The other half that were not asked could be observed to be counting the sets of dolls, usually out loud. Furthermore, the task was originally chosen because it was used successfully to teach children number before (Halford and Fullerton, 1970).

The only explanation that apparently can account for these results is that in any learning experiment, even with learning to a criterion, when retention is tested after a period of time a certain proportion of subjects "forget" what they learned (Hilgard and Marquis, 1968). To translate this into Piagetian terms, Piaget (1964) would predict that children who were transitional would accommodate their schemas, and if trained they would learn the new concept and retain it over time. Children who were non-transitional could learn to count, and learn that counting was necessary to get the "right" set of dolls. However, what these subjects in this study obviously did not notice was that spacing and length of the rows of the dolls was not related to how many dolls there were in a row. The Inhelder et al. (1974) tasks for number utilize spatial transformations of equal number rows or circles of objects. The non-transitional subject interprets the spatial transformations as changes in number.

Thus, transitional children as identified by the arousal measure and the developmental level measure learned the concept of number invariance, whereas the non-transitional children learned how to count the dolls, but did not learn number invariance. Another small group of children did not even learn to count the dolls. If the learning by accretion hypothesis is correct there should have been a good correlation between training and post-test performance. If the cognitive learning hypothesis of the Genevan's is correct, there should have been a low correlation between training and post-test performance. The latter was the case in this study.

Finally, after age, the only pre-test measures to account for a significant amount of total variance of the independent variables with training, were the developmental level scores. These results indicate that, even for training, the developmental level measure is a more valid measure than the judgment-only measure. Presumably, the developmental level scores do better here because they are consistently better at identifying non-transitional subjects than the judgment-only measure.

The results of this study support the hypothesis of the Genevans that only children in transition will respond to training (Inhelder et al., 1974; Piaget, 1971b).

4. Class Inclusion

"All that the operations constituting number, i.e., one to one correspondence (with conservation of the resulting equivalence despite differences in shape) or simple repetition of unity (1 + 1 = 2; 2 + 1 = 3; etc.) require are the additive groupings of classification and the serialization of asymetrical relations (ordering); but these are blended into a single operational whole, so that unit 1 is simultaneously an element in a class (1 included in 2, 2 in 3, etc.) and in a series (first 1 preceding the second 1, etc.) (Piaget, 1960, p.144)."

In other words, the understanding of Piaget's number concept depends upon the child's understanding of class inclusion and seriation. The results of this study confirmed that the number concept tested was related to seriation, but the verbal measure (developmental level) of class inclusion was not related to either seriation or number. Nor was class inclusion related to the post-tests of number. Not one subject out of the original 57 subjects passed class inclusion during pre-testing. There were four subjects who received a passing score on post-test class inclusion. Two of these subjects were dropped from the statistical analysis because they had perfect scores on both pre-tests seriation and number, and perfect scores on both number post-tests. The other two were among the 19 transitional subjects. These subjects also had perfect scores on the two number post-tests. The most frequent score was zero for all subjects. Thirty-seven subjects out of 57 scored zero on post-test class inclusion. A recent review of the research on class inclusion concluded that relationship of class inclusion to other cognitive developmental schemas is unclear (Winer, 1980). The unanticipated results of this study may shed some light on this problem. Arousal level during pre-testing for class inclusion is the only pre-test measure that is significantly related to the total post-test results. As mentioned in the results it is apparently the contribution of arousal class inclusion (ARC) to the total score for total arousal (ART) that results in the total arousal score (ART) classifying transitional and non-transitional subjects on the two number tests better than the total arousal score for seriation and number, ARSN (Table 11).

There are two possible explanations for this phenomenon. (1) Piagetian number tests are inappropriate tests of Piaget's (1960, p.144) number concept. This may be because as Piaget (1960) says: "Furthermore, as soon as these additive groupings are acquired multiplicative groupings are immediately understood in the form of correspondence (p.143)." The Inhelder et al., (1974) number tasks can be done if the child only understands number invariance, i.e., that spatial transformations do not change number. In other words, only addition and subtraction need to be understood in order to do the task. The understanding of the principle of multiplication is not necessary. However, it is clear, from Piaget's above definition, that class inclusion is necessary for the understanding of multiplication. (2) The other possibility is that the understanding of Piaget's full number concept is dependent upon the child being in decalage for both class inclusion and number invariance. If this is the case then attainment of the full number concept as Piaget (1960) describes it, would be dependent upon the understanding of seriation, followed by number invariance (addition and subtraction), followed by class inclusion, and then the full number concept (multiplication).

The results show that the subjects, who showed large increases in arousal during pre-testing for class inclusion also showed increases in arousal during pre-testing for seriation and number. These were the subjects who understood the concept of number as defined by the two number post-tests.

Thus, there are two questions which arise from this study that deserve further investigation. Are the Inhelder et al. (1974) tests for number appropriate tests for Piaget's (1960) theoretical formulation of the number concept? Does the attainment of class inclusion occur before or after number invariance? The results of this study tentatively indicate that class inclusion occurs after number invariance and that the child who will learn number invariance is in decalage for both class inclusion and number invariance. Given training for number invariance such children in this study attained number invariance but not class inclusion.

5. Conclusion

The results of this study of the relationship of the hypothetical construct arousal to Piaget's (1971a) hypothetical construct, equilibration, provide strong confirmation that arousal changes reflect the different amount of energy utilization hypothesized for the dual processes of equilibration, accommodation, and assimilation; during the performance of a Piagetian cognitive task. These results were also found to fulfill Cronbach and Meehl's (1966) criteria for construct validation.

It was also found that the Genevan developmental level measure was a more reliable and valid measure of transitional status than the Brainerd (1977) judgment-only measure. The Genevan method met statistical significance criteria, as well as Brainerd's (1977) criteria for the best verbal response measure. The most accurate measure for classifying transitional status was the arousal measure.

Training was found to have a low correlation with post-test performance. It was concluded that this was because both transitional and non-transitional subjects performed well on the training task, but the transitional subjects learned number invariance, and thus performed well on the post-tests. However, the non-transitional children learned only to count the dolls, thus doing poorly on the post-tests. As further evidence for this interpretation it was found that only the developmental level measure, which accurately classified non-transitional subjects, predicted training performance.

Finally, the findings, that arousal during pre-test class inclusion was highly predictive of post-test performance; whereas the verbal measure for both class inclusion tests were not; were interpreted to mean: either the Inhelder et al. (1974) tests of number were inappropriate measures of Piaget's (1960) number concept, or that class inclusion occurs after number invariance. A hypothesis was proposed integrating these two interpretations; that the development of Piaget's (1960) number concept proceeds from understanding seriation, to number invariance (addition and subtraction), followed by class inclusion and then the full number concept (multiplication).

References

Ainsworth, M.D.S. The developmenta of infant-mother attachment. In B.M. Caldwell and H.N. Ricciuti (Eds.), Review of child development research, vol. 3. Chicago: University of Chicago Press, 1973.

Amsel, A. The role of frustrative non-reward in non-continuous reward situations. Psychologican Bulletin, 1958, 55, 102119.

Amsel, A., Radek, C., Graham, M., & Letz, R. Ultrasound emission in infant rats as an indicant of arousal during appetitive learning and extinction. Science, 1977, 197, (4305), 786-788.

Anastasi, A. Psychological testing. New York: MacMillan, 1968.

Anastasi, A. Psychological testing. New York: MacMillan, 1969.

Beatty, J., & Wagoner, B.L. Pupillometric signs of brain activation vary with level of cognitive processing. Science, 1978, 199, (4334), 1216-1218.

Berlyne, D.E. Conflict arousal and curiosity. New York: McGraw-Hill, 1960.

Berlyne, D.E. Structure and direction in thinking. New York: Wiley, 1965.

Bloom, L. Language development review. In F.D. Horowitz (Ed.), Review of child development research (Vol. 4). Chicago: University of Chicago Press, 1975.

Bloom, L. Language development: Form and function in emerging grammars. Cambridge, Mass.: M.I.T. Press, 1970.

Bowlby, J. Attachment and loss. I: Attachment. New York: Basic Books, 1969.

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

Brainerd, C.J. Neo-Piagetian training experience revisited. Is there any support for the cognitive-developmental stage hypothesis? Cognition, 1973, 2, 349-370 (b).

Brainerd, C.J. Response criteria in concept development research. Child Development, 1977, 48, 360-366.

Brazier, M.A.B. The electrical activity of the nervous system. Baltimore: Williams and Wilkins, 1968.

Brodsky, P., & Brodsky, M. An arousal interval scale: A psychophysical scale for GSR analysis. Perceptual and Motor Skills, 1978, 47, 747-756.

Brodsky, P., & Brodsky, M. The galvanic skin response as a separate measure of Piaget's equilibration process. Proceedings of the Biennial Meeting of the Society for Research in Child Development, 1979 (Summary).

Bruner, J. On cognitive growth: II. In J. Bruner, R. Oliver and P. Greenfield (Eds.), Studies in cognitive growth. New York: John Wiley, 1966.

Bruner, J.S., & Postman, L. Perceptual defence. In J.S. Bruner and D. Kretch (Eds.), Perception and personality: A symposium. Durham: Duke University Press, 1950, p.38.

Campbell, D.T. Factors relevant to the validity of experiments in social settings. Psychological Bulletin, 1957, 54, 297-312.

Campbell, D.T., & Fiske, D.W. Convergent and discriminant validiation by the multitrait-multimethod matrix. In E. Megargee (Ed.), Research in clinical assessment. New York: Harper & Row, 1966.

Carnap, R. Logical foundations of the unity of science. In H. Feigl and W. Sellars (Eds.), Readings in philosophical analysis. New York: Appleton-Century-Crofts, 1949.

Cohen, J., & Cohen, P. Applied multiple regression/correlation analysis for the behavioral sciences. New York: John Wiley, 1975.

Cooley, L. An experimental analysis of the role of equilibration in the development of conservation of continuous quantity. Unpublished doctoral dissertation. University of Manitoba, 1978.

Cronbach, L.J., & Meehl, P.E. Construct validity in psychological tests. In E. Megaree (Ed.), Research in clinical assessment. New York: Harper & Row, 1966.

Dasen, P.R. Cross-cultural Piagetian research: A summary. Journal of Cross-Cultural Psychology, 1972, 3, 23-40.

Davidson, R. Specificity and patterning in biobehavioral systems: Implications for behavioral change. American Psychologist, 1978, 33, (5), 430-436.

Denenberg, V., Garbanati, J., Sherman, B., Yutzey, D., & Kaplan, R. Infantile stimulation induces brain lateralization in rats. Science, 1978, 201, 1150-1151.

Duffy, E. Activation and behavior. New York: John Wiley, 1962.

Epstein, S. The stability of behavior. II. Implications for psychological research. American Psychologist, 1980, 35, 790-806.

Fishbein, N., & Ajzen, I. Attitudes toward objects of single and multiple behavioral criteria. Psychological Review, 1974, 81, 59-74.

Forsberg, L.A. The role of field dependence-field independence in the conservation of mass, weight, and volume. Unpublished doctoral dissertation, University of Manitoba, 1973.

Gastaut, H. The role of the reticular formation in establishing conditioned reactions. In W.H. Jasper (Ed.), Reticular formation of the brain. Boston: Little-Brown, 1957, Pp. 561-579.

Gelman, R. Conservation acquisition: A problem of learning to attend to relevant attributes. Journal of Experimental Child Psychology, 1969, 7, 167-187.

Gelman, R. The nature and development of early number concepts. In H.W. Reese (Ed.), Advances in child development and behavior (Vol. 7). New York: Academic Press, 1972.

Grastyan, E. The hippocampus and higher nervous activity. In M. Brazier (Ed.), Transactions of the second conference on the central nervous system and behavior. New York: Josiah Macey Jr. Foundation, 1959, Pp. 119-205.

Green, B.F. In defense of measurement. American Psychologist, 1978, 33, 664-670.

Flavell, J.H. The developmental psychology of Jean Piaget. New York: Van Nostrand Reinhold, 1963.

Hale, W.D., & Strickland, B.R. Induction of mood states and their effect on cognitive and social behaviors. Journal of Consulting and

Clinical Psychology, 1976, 44, (1), 155.

Halford, G.S., & Fullerton, T.J.A. A discrimination task which induces conservation of number. Child Development, 1970, 41, 205-213.

Harris, V.A., & Katkin, E.S. Primary and secondary emotional behavior: An analysis of the role of autonomic feedback on affect, arousal, and attribution. Psychological Bulletin, 1975, 82 (6), 904-916.

Henry, D.E. Interrelationships among attentional preferences, cardinal-ordinal ability, and conservation of numbers. Child Development, 1976, 47, 750-758.

Hernandez-Peon, R. Neurophysiological correlates of habituation and other manifestations of plastic inhibition (internal inhibition). In H.H. Jasper (Ed.), The Moscow Colloquium on Electroencephalography and Clinical Neurophysiology. EEG Clinical Neurophysiology, 1960 (Suppl. 13).

Hilgard, E.R., & Marquis, D.G. Conditioning and learning. New York: Appleton-Century-Crofts, 1968.

Hull, C.L. A behavior system. New Haven: Yale University Press, 1952.

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

Inhelder, B., & Piaget, J. The early growth of logic in the child. New York: Harper & Row, 1964.

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

Inhelder, B., Sinclair, H., & Bovet, M. Learning and development of cognition. Cambridge, Mass.: Harvard University Press, 1974.

Johnson, O.G. Tests and measurements in child development: Handbook II. Vols. 1 and 2. San Francisco: Jossey-Bass, 1976.

Jouvet, M. Recherches sur les mecanismes neurophysiologiques du sommeil et de l'apprentissage negatif. In A. Fessard (Ed.), Brain mechanism and learning. Oxford: Blackwells, 1961.

Kahneman, D. Attention and effort. New York: Appleton-Century-Crofts, 1972.

Kinsbourne, M. The cerebral basis of lateral asymmetries in attention. Acta Psychologica, 1970, 33, 193-201.

Kinsbourne, M. The control of attention by interaction between the cerebral hemispheres. In S. Kornblum (Ed.), Attention and performance IV. New York: Academic Press, 1973.

Kinsbourne, M. Hemi-neglect and hemisphere rivalry. In E. Weinstein (Ed.), Hemi-attention and hemisphere specialization. New York: Raven Press, 1977.

Kinsbourne, M. Biological determinants of functional bysymmetry and asymmetry. In M. Kinsbourne (Ed.), Asymmetrical function of the brain. London: Cambridge University Press, 1978.

Lacey, J.I. Somatic response patterning and stress. Some revision of activation theory. In M.H. Appley and R. Turnball (Eds.), Psychological stress: Some issues in research. New York: Appleton-Century-Crofts, 1967. Pp. 14-37.

Lacey, J.I., Kagan, J., Lacey, B.C., & Moss, H.A. The visceral level.

Situational determinants and behavioral correlates of autonomic response patterns. In P.H. Knapp (Ed.), Expression of emotions in man. New York: International Universities Press, 1963.

Lacey, J.I., & Lacey, B.C. The law of initial value in the longitudinal study of automatic constitution: Reproducibility of autonomic response patterns over a four-year interval. Annals of new York Academy of Science, 1962, 98, 1257-1290.

Langer, J. Disequilibrium as a source of development. In P.H. Mussen, J. Langer and M. Covington, (Eds.), Trends and issues in developmental psychology. new York: Holt, Rinehart & Winston, 1969 (a).

Langer, J. Theories of development. New York: Holt, Rinehart & Winston, 1969 (b).

Larson, G.G. Methodology in developmental psychology: An examination of research on Piagetian theory. Child Development, 1977, 48 (4), 1160-1166.

Lawson, G., Baron, J., & Siegal, L. The role of number and length cues in children's quantitative judgments. Child Development, 1974, 45, 731-736.

Lindsley, D.B. Emotion. In S.S. Stevens (Ed.), Handbook of experimental psychology. New York: Wiley, 1951. Pp. 473-516.

Lindsley, D.B. The role of nonspecific reticulo-thalmo-cortical systems in emotion. In P. Black (Ed.), Physiological correlates of emotion. New York: Academic Press, 1970.

Lipton, E.L., Steinschneider, A., & Richmond, J.B. Autonomic function in the neonate. II. Physiologic effects of motor restraint. Psychosomatic Medicine, 1960, 22, (1), 57-65.

Lykken, D.T., & Venables, P.H. Direct measurement of skin conductance: A proposal for standardization. Psychophysiology, 1971, 8 (5), 656-672.

Lynn, R. Attention, arousal and orientation reaction. London: Pergamon Press, 1966.

MacCorquodale, K., & Meehl, P. Operational validity of intervening constructs. In M. Marx (Ed.), Psychological theory. New York: MacMillan, 1956.

Malmo, R.B. Activation: A neuro-psychological dimension. Psychological Review, 1959, 66, 367-386.

Mandler, G. Mind and emotion. New York: Wiley, 1975.

Marx, M.H. The general nature of theory construction. In M.H. Marx (Ed.), Theories in contemporary psychology. New York: MacMillan, 1964.

Messick, S. Test validity and the ethics of assessment. American Psychologist, 1980, 35, 1012-1026.

Miller, P.H. Attention to stimulus dimensions in conservation of liquid quantity. Child Development, 1973, 44, 129-136.

Miller, S. Nonverbal assessment of Piagetian concepts. Psychological Bulletin, 1976, 83, 405-430 (a).

Miller, S. Nonverbal assessment of conservation of number. Child Development, 1976, 47, 722-728 (b).

Milner, P.M. _Physiological psychology_, New York: Holt, Rinehart & Winston, 1970.

Mischel, T. Cognitive conflict and the motivation of thought. In T. Mischel (Ed.), _Cognitive development and epistemology_. New York: Academic Press, 1971.

Moruzzi, G. Synchronizing influences of the brain stem and the inhibitory mechanisms underlying the production of sleep by sensory stimulation. In H.H. Jasper and G.D. Smirnov (Eds.), Moscow Colloquium on Electroencephalography of Higher Nervous Activity. _EEG Clinical Neurophysiology_, 1960, (Suppl. 13).

Netter, F.H. _The CIBA collection of medical illustrations: 2. Nervous system_. New Jersey: CIBA, 1962.

Overton, W.F., & Reese, R.W. Models of development, methodological implications. In Nesselroade and H.W. Reese (Eds.), _Life span developmental psychology. Methodological issues_. New York: Academic Press, 1973. Pp. 65-68.

Piaget. J. _The child's conception of number_. New York: Humanities, 1952.

Piaget, J. _Les relations entre l'affectivite et l'intelligence dans le development mental de l'enfant_. Paris: Centre de Documentation Universitaire, 1954.

Piaget, J. Equilibration and the development of logical structures. In J.M. Tanner and B. Inhelder (Eds.), _Discussions on child development. Vol. 4_. New York: International Universities Press, 1960. Pp. 98-115. (a)

Piaget, J. _Psychology of intelligence_. New Jersey: Littlefield, Adams

& Co., 1960 (b).

Piaget, J. Development and learning. Journal of Research in Science Teaching, 1964, 2, 176-186.

Piaget, J. Biology and knowledge: An essay on the relations between organic regulations and cognitive processes. Chicago: University Press, 1971 (a).

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

Piaget, J. Problems of equilibration. In C.F. Nodine, J.M. Gallagher, & Humphreys, G. (Eds.), Piaget and Inhelder: On equilibration. Philadelphia: The Jean Piaget Society, 1972.

Pick, A., Frankel, D., & Hess, V. Children's attention: The development of selectivity. In E. Hetherington (Ed.), Review of child development reseach. Chicago: University of Chicago Press, 1975.

Razik, T. Psychometric measurement of creativity. In P.E. Vernon (Ed.), Creativity. New York: Penguin, 1970.

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

Schachter, S. The interaction of cognitive and physiological determinants of emotional state. In L. Berkowitz (Ed.), Advances in experimental social psychology. (Vol. 1). New York: Academic Press, 1964. Pp. 49-80.

Schachter , S., & Singer, J. Cognitive, social, and physiological determinants of emotional state. Psychological Review, 1962, 69 379-399.

Sharpless, S., & Jasper, H. Habituation of the arousal reaction. Brain, 1956, 79, 655-680.

Skinner, B.F. Science and human behavior. New York: The Free Press, 1953.

Sokolov, E. Neuronal models and the orienting reflex. In M. Brazier (Ed.), Transactions of the third conference on the central nervous system and behavior. New York: Josiah Macey Jr. Foundation, 1960. Pp. 187-276.

Spence, K.W. Behavior theory and conditioning. New Haven: Yale University Press, 1956.

Sperry, R.W. How the developing brain gets itself properly wired for adaptive functioning. In J. Sants and H.J. Butcher (Eds.), Developmental psychology, New York: Penguin, 1975.

Sroufe, A.L. Wariness of strangers and the study of infant development. Child Development, 1977, 48 (3), 731-746.

Strauss, S. Inducing cognitive development and learning: A review of short term training experiments. 1. The organismic development approach. Cognition, 1972, 1, 329-357.

Strauss, S., & Langer, J. Operational thought inducement. Child Development, 1970, 41, 163-175.

Tatsuoka, M. Multivariate analysis. New York: John Wiley & Sons, 1971.

Teyler, T. A primer of psychobiology: Brain and behavior. San Francisco: W.H. Freeman, 1975.

Valins, S. The perception and labeling of bodily changes as determinants of emotional behavior. In P. Black (Ed.), Physiological correlates of emotion. New York: Academic Press, 1970.

Venables, P.H., & Christie, M.J. Mechanisms, instrumentation, recording techniques, and quantification of responses. In W.F. Prokasy and D.C. Rasken (Eds.), Electroderman activity in psychological research. New York: Academic Press, 1973.

von Bertlanffy, F. General systems theory. New York: George Braziller, 1968.

Waddington, C.H. New patterns in genetics and development. New York: Columbia University Press, 1962.

Walsh, R.N., & Cummins, R.A. Mechanisms mediating the production of environmentally induced brain changes. Psychological Bulletin 1975, 82, (6), 986-1000.

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

Winer, G.A. Class inclusion reasoning in children: A review of the empirical literature. Child Development, 1980, 51, 309-328.

Wolff, P.H. The development of attention in infants. The Annals of the New York Academy of Sciences, 1965, 118, 815-830.

Woodworth, R.S., & Schlosberg, H. Experimental psychology. New York: Holt, 1956.

Zelazo, P. Smiling and vocalizing: A cognitive emphasis. Merrill-Palmer Quarterly, 1972, 18, 349-366.

Table 1

Multiple Regression Analysis of Post-Test Class Inclusion and Number Tests (POT) with the Individual Pre-Test Measures

Independent Variables	SEX	AGE	ARC[a]	DLN	CONN	ARS	DLC	JOC	ARN	JON	DLS	JOS	CONS	CONC
Beta weight	-.005	.154	.610	.154	.224	-.435	.410	-.379	.158	.221	-.349	.331	.101	-.028
F	.004	2.453	6.700*	.289	.879	3.987	3.276	2.815	.690	.589	2.994	2.644	.248	.029
df	1,36	1,35	1,34	1,33	1,32	1,31	1,30	1,29	1,28	1,27	1,26	1,25	1,24	1,23
Multiple R	.870													
Overall F	8.029**													
df	14,36													
Total percent variance accounted for 76%														

*$p < .05$

**$p < .001$

[a]ARC=Arousal Interval Scale scores for Class Inclusion. ARS=Arousal Interval Scale scores for Seriation. ARN=Arousal Interval Scale scores for Number. CONC=Conductance scores for Class Inclusion. CONS=Conductance scores for Seriation. CONN=Conductance scores for Number. DLC=Developmental scores for Class Inclusion. DLS=Developmental Level scores for Seriation. DLN=Developmental Level scores for Number. JOC=Judgments Only scores for Class Inclusion. JOS=Judgments Only scores for Seriation. JON=judgments Only scores for Number.

Table 2

Multiple Regression Analysis of the Post-Test Class Inclusion

Test with the Individual Pre-Test Measures

Independent Variables	SEX	AGE	CONC[a]	DLC	DLN	JOC	CONN	ARC	ARS	JON	JOS	DLS	ARN	CONS
Beta weight	-.064	.021	-.027	.728	.486	-.634	.267	1.022	-.729	-.306	.397	-.327	-.207	-.219
F	.219	.025	.014	5.317*	1.476	4.062	.641	9.689**	5.764*	.580	1.959	1.354	.612	.603
df	1,36	1,35	1,34	1,33	1,32	1,31	1,30	1,29	1,28	1,27	1,26	1,25	1,24	1,23
Multiple R	.727													
Overall F	2.8950*													
df	14,36													
Total percent of variance accounted for 53%														

*$p < .05$

**$p < .01$

[a]CONC=Conductance scores for Class Inclusion. CONS=Conductance scores for Seriation. CONN=Conductance scores for Number. ARC=Arousal Interval Scale scores for Class Inclusion. ARS=Arousal Interval Scale scores for Seriation. ARN=Arousal Interval Scale scores for Number. DLC=Developmental Level scores for Class Inclusion. DLS=Developmental Level scores for Seriation. DLN=Developmental Level scores for Number. JOC=Judgments-Only scores for Class Inclusion. JOS=Judgments-Only scores for Seriation. JON=Judgments-Only scores for Number.

Table 3

Multiple Regression Analysis of Post-Test Total Score (POT)

With Total Scores of the Pre-Test Measures

Independent Variables	SEX	AGE	JOSN[a]	JOT	DLSN	DLT	ARSN	ART	COSN	CONT
Beta weight	.032	.147	.947	-.725	-1.369	1.397	-1.207	1.565	.265	-.011
F	.090	1.934	3.841	2.813	3.473	4.035	3.675	5.933*	.414	.001
df	1,40	1,39	1,38	1,37	1,36	1,35	1,34	1,33	1,32	1,31

Multiple R .80

Overall F 7.136**

df 10,40

R^2 percent variance accounted for: 64%

*p .05

**p .001

[a]JOSN = Judgment-only total score for seriation and number. JOT = Judgment-only total score for class inclusion, seriation and number. DLSN = Developmental level total score for seriation and number. DLT = Developmental level total score for class inclusion, seriation and number. ARSN = Arousal interval scale total score for seriation and number, ART = Arousal interval scale total score for class inclusion, seriation and number. COSN = conductance total score for seriation and number. CONT = conductance total score for class inclusion, seriation and number.

Table 4

Multiple Regression Analysis of Training Scores (TRAIN)

with Total Scores of the Pre-Test Measures

Independent Variables	SEX	AGE	JOSN[a]	JOT	DLSN	DLT	ARSN	ART	COSN	CONT
Beta weight	-.209	.369	-.753	1.038	2.026	-1.935	.798	-.827	-.293	.548
F	2.270	7.439**	1.485	3.529	4.639*	4.730*	.983	1.014	.309	1.008
df	1,39	1,40	1,38	1,37	1,36	1,35	1,34	1,33	1,32	1,31
Multiple R	.642									
Overall F	2.809*									
df	10,40									
Total percent of variance accounted for 41%										

*p .05

**p .01

[a]JOSN = Judgment-only total score for seriation and number. JOT = Judgment-only total score for class inclusion, seriation and number. DLSN = Developmental level total scores for seriation and number. DLT = Developmental level total score for class inclusion, seriation and number. ARSN = Arousal interval scale score for seriation and number. ART = Arousal interval scale total score for class inclusion, seriation and number. COSN = Conductance total score for seriation and number. CONT = Conductance total score for class inclusion, seraition and number.

Table 5

Canonical Correlation Analysis: Redundancy Coefficients:

Squared Multiple Correlations of Each Independent Variable With

Every Other Variable in the Set

	Variable	R Squared
Judgment-only class inclusion	JOC	.867
seriation	JOS	.836
number	JON	.913
Developmental Level class inclusion	DLC	.867
seriation	DLS	.831
number	DLN	.915
Conductance class inclusion	CONC	.740
seriation	CONS	.806
number	CONN	.869
Arousal Interval Scale class inclusion	ARC	.876
seriation	ARS	.849
number	ARN	.801

Squared Multiple Correlations of Each Dependent
Variable with Every Other Variable in the Set

	Variable	R Squared
Training	TRAIN	.23
Post-tests class inclusion	POC	.07
first number test	PON	.75
second number test	POP	.78

Table 6

Correlations of Redundant Independent Variables

Arousal Measures		Verbal Measures	
ARC - CONC	.64	DLC - JOC	.89
ARS - CONS	.56	DLS - JOS	.89
ARN - CONN	.61	DLN - JON	.93
ARSN - COSN	.61	DLSN - JOSN	.93
ART - CONT	.63	DLT - JOT	.89

ARC = Arousal interval scale class inclusions scores.
ARS = Arousal interval scale seriation scores.
ARN = Arousal interval scale number scores.
ARSN = Arousal interval scale total scores for seriation and number.
ART = Arousal interval scale total scores for class inclusion, seriation and number.
CONC = Conductance class inclusion scores.
CONS = Conductance seriation scores.
CONN = Conductance number scores.
CONS = Conductance total scores for seriation and number.
CONT = Conductance total scores for class inclusion, seriation and number.
DLC = Developmental level class inclusion scores.
DLS = Developmental level seriation scores.
DLN = Developmental level number scores.
DLSN = Developmental level total scores for seriation and number.
DLT = Developmental level total scores for class inclusion seriation and number.
JOC = Judgments-only class inclusion scores.
JOS = Judgments-only seriation scores.
JON = Judgments-only number scores.
JOSN = Judgments-only total scores for seriation and number.
JOT = Judgments-only total scores for class inclusion, seriation and number.

Table 7

Correlation of Pre-Tests and Post-Tests with Training Scores

Pre-Tests		Post-Tests	
Judgments-Only		Class inclusion	.01
class inclusion	-.06	Number test 1	.18
seriation	.24	Number test 2	.38**
number	.33*		
Developmental Level			
class inclusion	-.14		
seriation	.29*		
number	.32*		
Arousal Interval Scale			
class inclusion	.14		
seriation	.20		
number	.22		
Conductance			
class inclusion	.20	Age .414**	
seriation	.14		
number	.25		

Pre-Tests		Post=Tests	
Judgments-Only		Number tests 1 and 2	.30*
seriation and number	.40**	Class inclusion,	
class inclusion, seriation and number	.32*	number tests 1 and 2	.273*
Developmental Level			
seriation and number	.43**		
class inclusion, seriation and number	.37**		
Arousal Interval Scale			
seriation and number	.22		
class inclusion, seriation and number	.20		
Conductance			
seriation and number	.20	Age .414**	
class inclusion, seriation and number	.21		

Correlation of Post-Tests with Each Other

Post-Tests			
Class inclusion	-	Number test 1.	.24
Class inclusion	-	Number test 2.	.26
Number test 1.	-	Number test 2.	.85***

Table 8

Canonical Correlation Analysis Between Two Sets of Variables: Pre-Test Measures (Indpendent Variables), and Dependent Variables, Training and Post-Tests

Pair of Canonical Variates	Canonical Correlation	Eigen Value	Chi-Square	df
1	.866	.750	107.98***	48
2	.697	.418	50.41*	33
3	.549	.301	23.42	20
4	.430	.185	8.52	9

*$p < .05$

**$p < .01$

***$p < .001$

Table 9

Canonical Correlation Analysis: Standardized Coefficients (Loadings)

First Pair of Canonical Variates

Independent Variables		Dependent Variables	
Judgments-only class inclusion (JOC)	-.544	Training (TRAIN)	-.051
Judgments-only seriation (JOS)	.428	Post-test class inclusion (POC)	.431
Judgments-only number (JON)	.186		
Developmental Level class inclusion (DLC)	.647	Post-test 1 number	
Developmental Level seriation (DLS)	-.410	(PON)	.401
Developmental Level number (DLN)	.272	Post-test 2 number	
Conductance class inclusion (CONC)	-.073	(POP)	.438
Conductance seriation (CONS)	-.025		
Conductance number (CONN)	.385		
Arousal Interval Scale class inclusion (ARC)	.857	R_c = .866**	
Arousal Interval Scale seriation (ARS)	-.566		
Arousal Interval Scale number (ARN)	.089	Variance accounted for 75%	

Second Pair of Canonical Variates

Independent Variables		Dependent Variables	
Judgments-only class inclusion (JOC	1.445	Training (TRAIN)	.366
Judgments-only seriation (JOS)	-.172	Post-test class inclusion (POC)	-.719
Judgments-only number (JON)	.613		
Developmental Level class inclusion (DLC)	-1.290	Post-test 1 number (PON)	-.442
Developmental Level seriation (DLS)	.193		
Developmental Level number (DLN)	-.273	Post-test 2 number (POP)	.895
Conductance class inclusion (CONC)	.189		
Conductance seriation (CONS)	.164		
Conductance number (CONN)	.202		
Arousal Interval Scale classinclusion (ARC)	-1.366	R_c = .697*	
Arousal Interval Scale seriation (ARS)	1.069		
Arousal Interval Scale number (ARN)	.287	Variance accounted for 48%	

*p < .05

**p < .001

Table 10

Canonical Correlation Analysis Structure Coefficients:
Correlation of Each Variable in a Set with its Canonical Variate

First Pair of Canonical Variates

Dependent Variables		Independent Variables	
Judgments-only class inclusion (JOC)	-.069	Training (TRAIN)	.203
Judgments-only seriation (JOS)	.056	Post-test class inclusion (POC)	.644
Judgments-only number (JON)	.639	Post-test 1 number (PON)	.878
Developmental Level class inclusion (DLC)	.116	Post-test 2 number (POP)	.868
Developmental Level seriation (DLS)	.020		
Developmental Level number (DLN)	.730		
Conductance class inclusion (CONC)	.651		
Conductance seriation (CONS)	.598		
Conductance number (CONN)	.621		
Arousal Interval Scale class inclusion (ARC)	.694		
Arousal Interval Scale seriation (ARS)	.470		
Arousal Interval number (ARN)	.567		

Second Pair of Canonical Variates

Dependent Variables		Independent Variables	
Judgments-only class inclusion (JOC)	-.086	Training (TRAIN)	.616
Judgments-only seriation (JOS)	.051	Post-test class inclusion (POC)	-.605
Judgments-only number (JON)	.334	Post-test 1 number (PON)	.204
Developmental Level class inclusion (DLC)	.317	Post-test 2 number (POP)	.480
Developmental Level seriation (DLS)	.192		
Developmental Level number (DLN)	.236		
Conductance class inclusion (CONC)	.140		
Conductance seriation (CONS)	.327		
Conductance number (CONN)	.452		
Arousal Interval Scale class inclusion (ARC)	.118		
Arousal Interval Scale seriation (ARS)	.371		
Arousal Interval Scale number (ARN)	.486		

Table 11

Chi-Square Analysis of the Number of Subjects Correctly and Incorrectly Identified by the Three Pre-test Measures

Independent Variables	Arousal (ART)			Developmental Level (DLT)			Judgment-Only (JOT)		
	Transi-tional	Non-trans-tional	Error rate	Transi-tional	Non-trans-tional	Error rate	Transi-tional	Non-trans-tional	Error rate
Number correctly identified	17	22		13	22		16	14	
Number incorrectly identified	10	2	24%	10	6	31%	18	3	41%
Cut off Points Post-test[1] PONP	ART 10.56 median ARSN 6 median			DLT 5 or more out of 18 DLSN 4 or more out of 12			JOT 2 or more out of 6 JOSN 1 or more out of 4		
Cut-off Points Post-test[1] PONP	10 out of 12			10 out of 12			10 out of 12		
X^2 df 1	ART	14.29**		DLT	7.08*		JOT	1.06	
	ARSN	10.37*		DLSN	7.08*		JOSN	.02	

[1] 19 subjects received a score of 10 or better on post-tests number 1 and 2 (PONP)

* p .01

** p .001

[a] ART, Arousal Interval Scale total score for Class Inclusion, Seriation and Number. ARSN, Arousal Interval Scale total score for Seriation and Number. JOT, Judgments-only total score for Class Inclusion, Seriation and Number. JOSN, Judgments-only total score for Seriation and Number. DLT, Developmental Level total score for Class Inclusion, Seriation and Number. DLSN, Developmental Level total score for Seriation and Number.

Table 12

The Correlation Coefficients of Arousal Scores (AR) with Genevan Developmental Level Scores (DL) and with Brainerd's Judgment-Only Scores (JO)

	Jugments-Only		Developmental Level	
Pre-Tests				
Class Inclusion (C)	JOC[a]	-.15	DLC	-.13
Seriation (S)	JOS	-.08	DLS	-.05
Number (N)	JON	.29*	DLN	.35*
C + S + N	JOT	-.08	DLT	-.03
S + N	JOSN	.13	DLSN	.13

*p .05

[a]JOC, Judgments-only class inclusion; JOS, Judgments-only seriation; JON, Judgments-only number; JOT, Judgments-only total score for class inclusion, seriation and number; JOSN, Judgments only total score for seriation and number; DLC, Developmental Level class inclusion; DLS, Developmental Level seriation; DLN, Developmental Level number; DLT, Developmental Level total score for class inclusion, seriation and number; DLSN, Developmental Level seriation and number

Table 13

Descriptive Statistics of the Independent Variables and Dependent Variables

Independent Variables	Mean	Standard Deviation	Range
Sex	1.491	.504	1.000
Age	60.491 months	5.386	22.000
Judgments-Only			
Class inclusion (JOC)	.439	.627	2.000
Seriation (JOS)	.965	.801	2.000
Number (JON)	.807	.854	2.000
Total score for seriation and number (JOSN)	1.772	1.310	4.000
Total score for class inclusion, seriation and number (JOT)	2.211	1.333	4.00
Developmental Level			
Class inclusion (DLC)	.581	.887	4.000
Seriation (DLS)	2.368	2.127	6.000
Number (DLN)	2.018	2.303	6.000
Total score seriation and number (DLSN)	4.386	3.707	12.000
Total score for class inclusion, seriation and number (DLT)	4.947	3.642	12.000
Arousal Interval Scale			
Class inclusion (ARC)	4.951	4.525	17.000
Seriation (ARS)	4.974	5.107	21.000
Number (ARN)	5.012	4.728	19.000
Total score for seriation and number (ARSN)	9.986	9.349	40.000
Total score for class inclusion, seriation and number (ART)	14.937	13.371	53.500
Conductance Scale			
Class inclusion (CONC)	12.765	1.476	5,230
Seriation (CONS)	12.930	1.351	5.840
Number (CONN)	12.832	1.280	5.300
Total score for seriation and number (COSN)	25.762	2.525	10.000
Total score for class inclusion, seriation and number (CONT)	38.562	3.731	15.210

Continued...

Table 13 (Continued)

Descriptive Statistics of the Independent Variables and Dependent Variables

	Mean	Standard Deviation	Range
Dependent Variables			
Training	23.912	11.397	35.000
Post-Tests			
Class inclusion (POC)	.789	1.820	6.000
Number test 1 (PON)	3.088	2.721	6.000
Number test 2 (POP)	3.351	2.761	6.000
Total score for number tests 1 and 2 (PONP)	6.439	5.288	12.000
Total score for class inclusion, number tests 1 and 2 (POT)	7.228	6.077	18.000

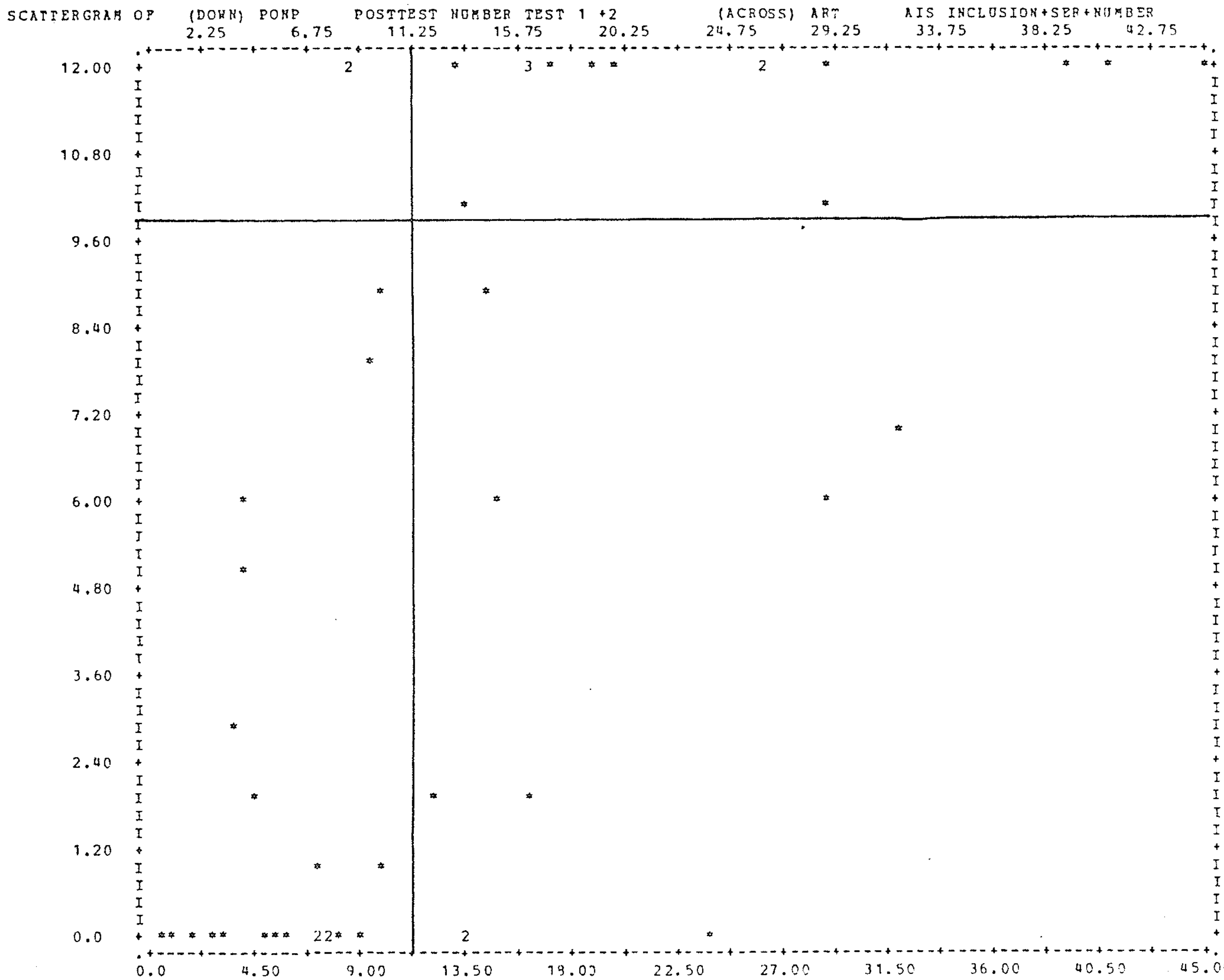


Figure 1

AIS Arousal Interval Scale

ART Arousal Interval Scale total scores for class inclusion, seriation and number.

PONP Post-tests total score for number post-tests 1 and 2.

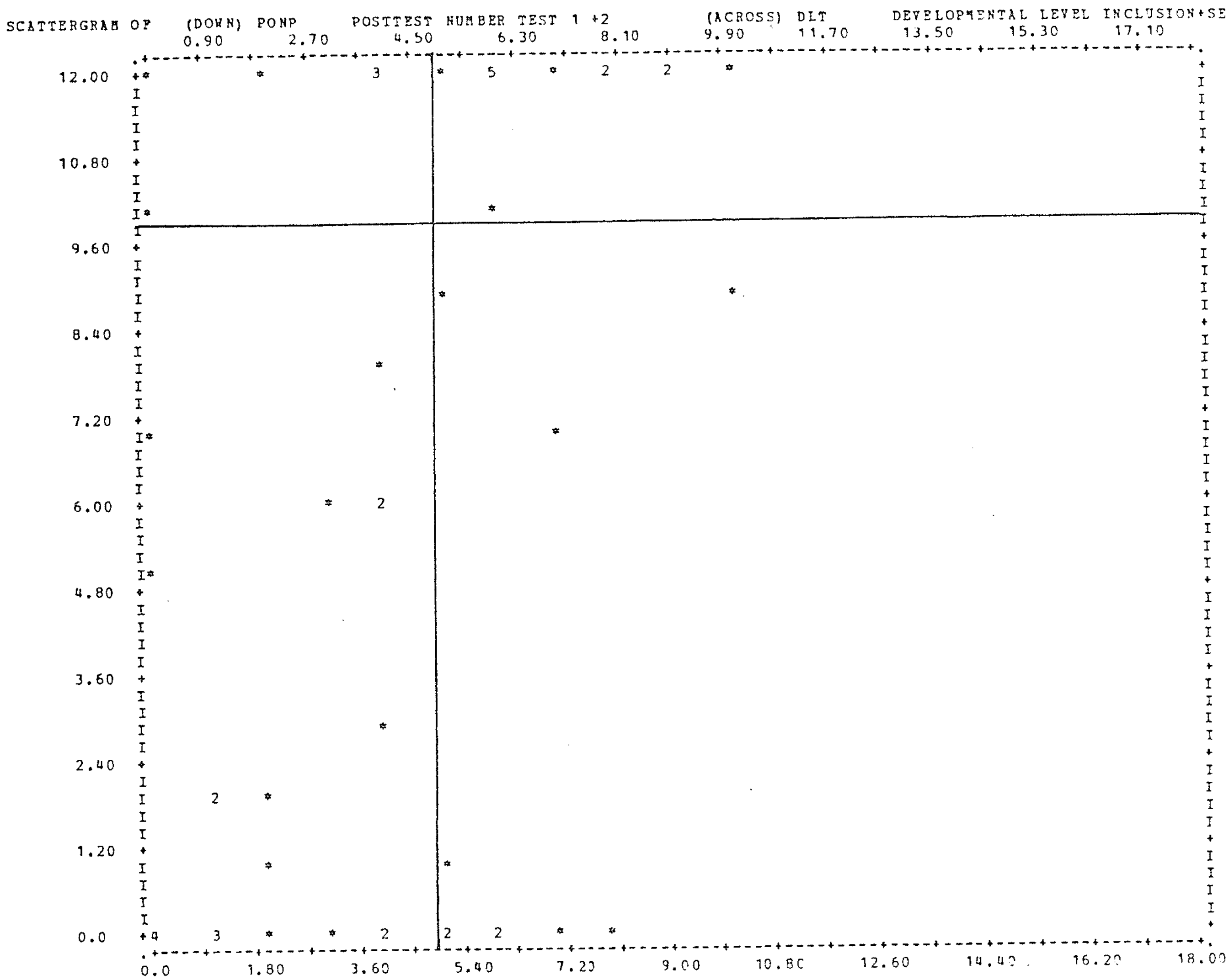


Figure 2

DLT Developmental Level total score for Class Inclusion, Seriation and Number.

PONP Post-tests total score for Number Post-tests 1 and 2.

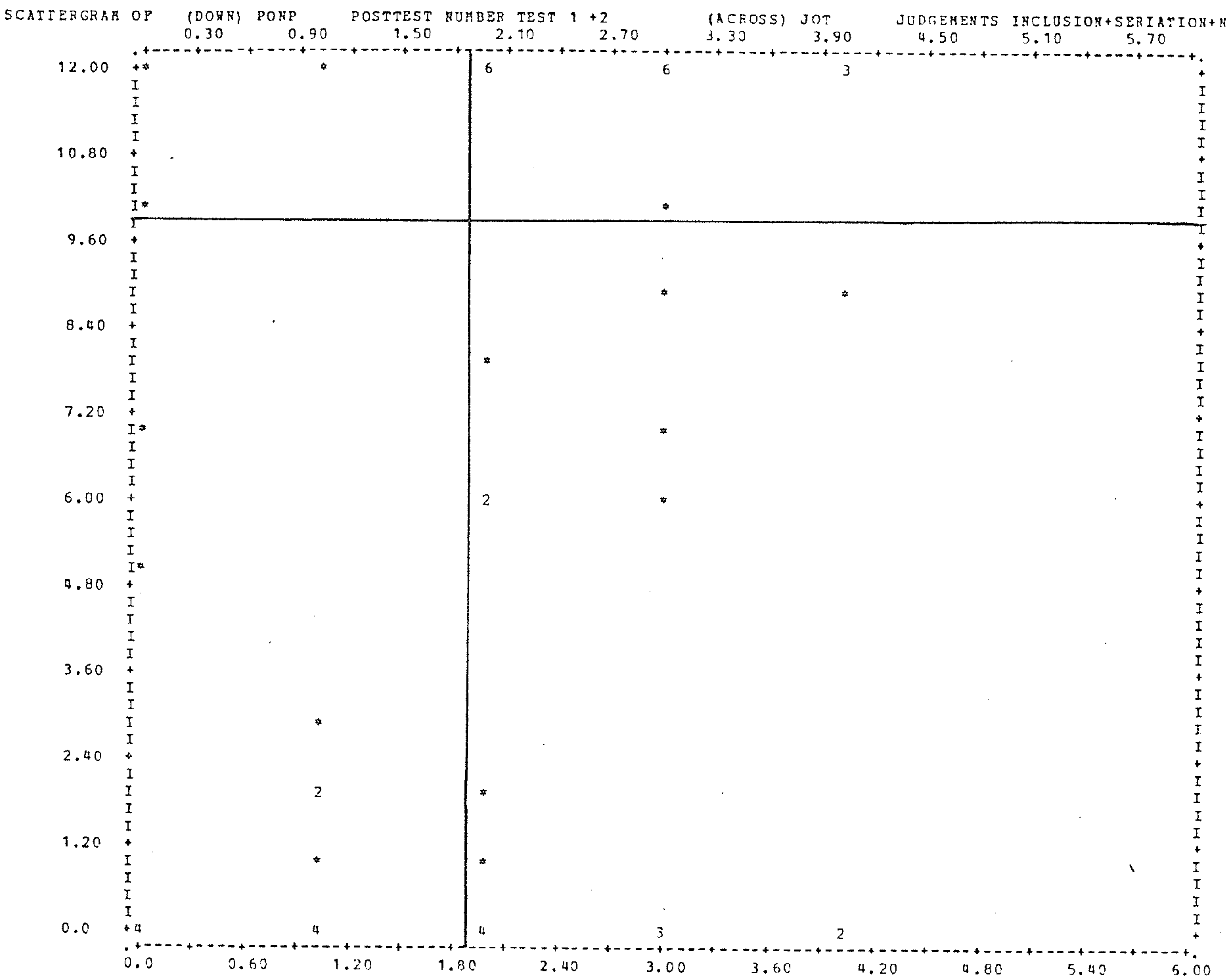
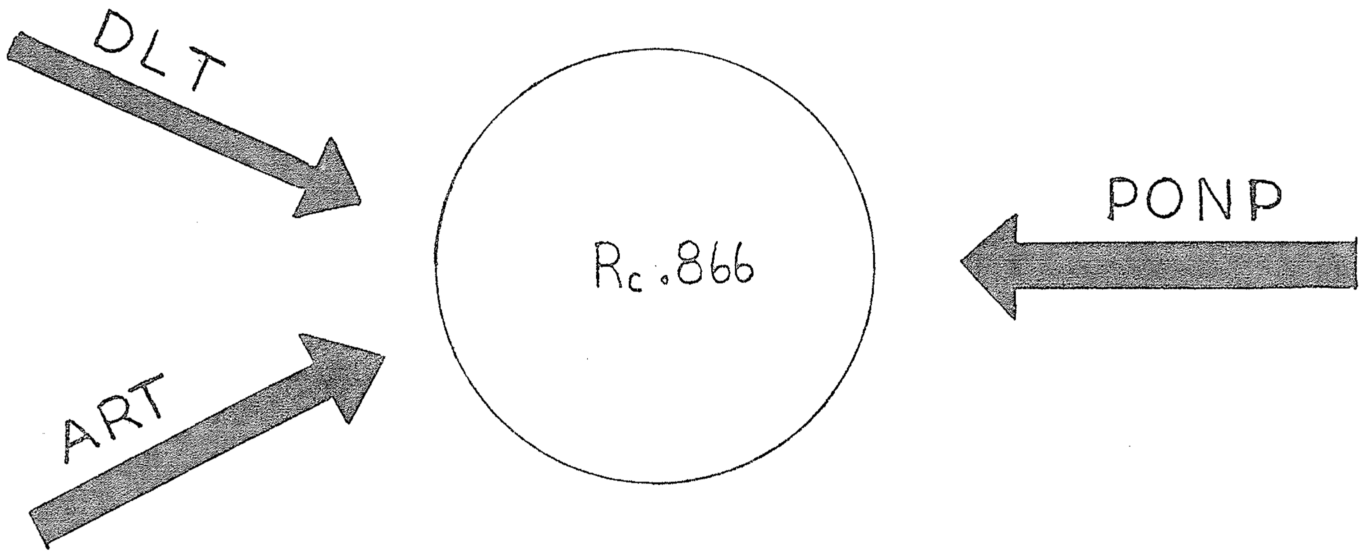


Figure 3

JOT Judgments Only total score for class inclusion, seriation and number.

PONP Post-tests total score for number Post-tests 1 and 2



FIRST PAIR OF CANONICAL VARIATES

Figure 4

DLT Developmental level total score for class inclusion, seriation and number.

ART Arousal Interval Scale total score for class inclusion, seriation and number.

PONP Post-tests total score for Number tests 1 and 2